SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di scienze matematiche fisiche e naturali
Dipartimento di chimica

Mass spectrometry based metabolomics approaches for food safety, quality and nutrition.

PhD student: Giorgia La Barbera

Tutor: Prof. Chiara Cavaliere

PhD Coordinator: Prof. Osvaldo Lanzalunga

PhD thesis in Chemistry
XXX cycle
2017

« C'est une folie de haïr toutes les roses parce qu'une épine vous a piqué,
d'abandonner tous les rêves parce que l'un d'entre eux ne s'est pas réalisé,
de renoncer à toutes les tentatives parce qu'on a échoué.
C'est une folie de condamner toutes les amitiés parce qu'une vous a trahi,
de ne croire plus en l'amour juste parce qu'un d'entre eux a été infidèle,
de jeter toutes les chances d'être heureux
juste parce que quelque chose n'est pas allé dans la bonne direction.
Il y aura toujours une autre occasion,
un autre ami, un autre amour, une force nouvelle.
Pour chaque fin il y a toujours un nouveau départ. »

Antoine de Saint-Exupéry

Table of contents

PhD thesis publications

"Mycoestrogen determination in cow milk: Magnetic solid-phase extraction followed by liquid chromatography and tandem mass spectrometry analysis."
A.L. Capriotti, C. Cavaliere*, P. Foglia, G. La Barbera, R. Samperi, S. Ventura, A. Laganà.
Journal of Separation Science. Volume 39, Pages 4794–4804. 26 October 2016.
doi: 10.1002/jssc.201600879.

"**A rapid magnetic solid phase extraction method followed by liquid chromatography-tandem mass spectrometry analysis for the determination of mycotoxins in cereals".**
G.La Barbera, A.L. Capriotti, C.Cavaliere, P. Foglia, C.M. Montone, R. Zenezini Chiozzi*, A. Laganà.
Toxins, Volume 9, Issue 9. 2017
doi: 10.3390/toxins9040147

"Comprehensive polyphenol profiling of a strawberry extract (Fragaria × ananassa) by ultra high performance liquid chromatography coupled to high resolution mass spectrometry"
G. La Barbera, A.L.Capriotti, C. Cavaliere*, S. Piovesana, R. Samperi, R. Zenezini Chiozzi, A. Laganà.
Analytical and Bioanalytical Chemistry, Volume 409, Issue 8, Pages 2127-2142, March 2017.
doi:10.1007/s00216-016-0159-8

"**Evaluation of column length and particle size effect on the untargeted profiling of a phytochemical mixture by means of ultra-high performance liquid chromatography coupled to high resolution mass spectrometry"**
R. Zenezini Chiozzi, A.L. Capriotti, C.Cavaliere, F. Ferraris, G.La Barbera *, S. Piovesana, A. Laganà.
Journal Separation Science. Volume 40, Issue 12, Pages 2541-2557, June 2017.
doi: 10.1002/jssc.201700135

"Chromatographic column evaluation for the untargeted profiling of glucosinolates in cauliflower by means of ultra-high performance liquid chromatography coupled to high resolution mass spectrometry"
A. L. Capriotti, C. Cavaliere, G. La Barbera*, C. M. Montone, S. Piovesana, R. Zenezini Chiozzi, A. Laganà
Talanta. Volume 179C, Pages 792-802, January 2018.
doi.org/10.1016/j.talanta.2017.12.019

"Liquid chromatography-high resolution mass spectrometry for the analysis of phytochemicals in vegetal-derived food and beverages"
G. La Barbera, A. L. Capriotti, C. Cavaliere, C. M. Montone, S. Piovesana*, R. Samperi, R. Zenezini Chiozzi, A. Laganà
Food Research International, Volume 100, Part 1, Pages 28-52, October 2017.
doi.org/10.1016/j.foodres.2017.07.080

Abbreviations

ACN	Acetonitrile
AFB1	Aflatoxin B1
AFB2	Aflatoxin B2
AFG1	Aflatoxin G1
AFG2	Aflatoxin G2
AIF	All ion Fragmentation
APCI	Atmospheric Pressure Chemical Ionization
CID	Collision-Induced Dissociation
DDA	Data Dependent Acquisition
DIA	Data Independent Acquisition
EFSA	European Food Safety Authority
ESI	Electrospray
ETs	Ellagitannins
FAO	Food and Agriculture Organization
FWHM	Full Width at Half Maximum
FT	Fourier Transform
GCB	Graphitized Carbon Black
GLS	Glucosinolate
HCD	Higher-Energy Collisional Dissociation
HHDP	Hexahydroxydiphenic acid
HR	High Resolution
HRMS	High-Resolution Mass Spectrometry
IS	Internal Standard
IT	Ion trap

LC-MS	Liquid Chromatography-Mass Spectrometry
LLE	Liquid-Liquid Extraction
LOD	Limit of Detection
LOQ	Limit of Quantification
LR	Low Resolution
MALDI	Matrix Assisted Laser Desorption Ionization
ME	Matrix effect
MeOH	Methanol
mGCB	magnetic Graphitized Carbon Black
ML-PLSDA	Multilevel Partial Least-Squares Discriminant Analysis
MNP	Magnetic Nanoparticles
MRL	Maximum Residue Level
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
mSPE	magnetic Solid Phase Extraction
m/z	mass to charge ratio
NCE	Normalized Collision Energy
OTA	Ochratoxin A
PAs	Proanthocyanidins
PIS	Product Ion Scan
PLS-DA	Partial Least Square-Discriminant Analysis
PE	Process Efficiency
QC	Quality Control
QqQ,	Triple quadrupole
qTOF	quadrupole-TOF

QuEChERS	Quick, Easy, Cheap, Effective, Rugged, and Safe
RE	Recovery
RSD	Relative Standard Deviation
RT	Retention Time
SIM	Single Ion Monitoring
SID	source ion dissociation
SLE	Solid-Liquid Extraction
SPE	Solid-Phase Extraction
SPME	Solid-Phase Microextraction
SRM	Selected Reaction Monitoring
TFA	Trifluoroacetic Acid
TOF	Time-of-Flight
UHPLC	Ultra-High Performance Liquid Chromatography
WHO	World Health Organization
XIC	Extracted Ion Chromatogram
ZAN	Zearalanone
α-ZAL,	α-zearalanol
β -ZAL,	β-zearalanol
α-ZEL,	α-zearalenol
ZEN	Zearalenone

Abstract

In the last years, the research in food science and nutrition has grown parallel to the consumers' concern about the quality and safety of the food they eat. As a result, ensuring food safety, food quality and investigating on nutrition has never been more necessary than today. More powerful analytical procedures are now required to detect and identify undesired and toxic compounds. Moreover, consumers demand is now moving toward the valorization of food with potential health benefits. The identification of bioactive compounds is crucial to provide customers a healthy diet. Finally, one of the main challenges is to improve our limited understanding of the roles of nutritional compounds at molecular level. Food scientists have to face a large number of challenges to adequately answer the new emerging questions of food science. In this context, mass spectrometry showed to be a powerful tool for the investigation on food quality, safety and nutrition.

The technical developments of chromatography coupled to mass spectrometry allowed several progresses for the detection of both endogenous compounds and contaminants in food. However, the main progress was due to high resolution mass spectrometry, a promising technique that has opened new horizons in screening and identification of a wide range of unknowns compounds. Mass spectrometry-based metabolomics is a key tool that has been involved, nowadays, in the study of the food and nutrition domains. As per definition, metabolomics includes the exhaustive study of the whole small metabolite composition of a particular system, the food metabolome is not an exception, being rich of endogenous and exogenous compounds with different properties and abundances. Nowadays, three main approaches for the screening of substances can be pointed out: targeted screening (analysis of few known compounds), suspected screening (analysis of a class of expected compounds) and untargeted screening (analysis of a wide range of unknown and unexpected compounds).

In this thesis, the three approaches have been studied in details to understand their difficulties, advantages and disadvantages for the analysis of food metabolome. Several steps have been proved to be necessary for the development of a reliable analysis such as sample preparation, chromatographic and mass spectrometric method, data analysis and identification. The thesis is then divided in three main sections: targeted approach for food safety, suspected approach for food quality, untargeted approach for nutrition. In each section, one of the three main approaches used in food metabolomics has been studied for

its application on one of the three different food analysis fields. Moreover, each topic has been treated focusing the attention on a specific step of the method development.

The topic of metabolomics for food safety has been studied in a targeted approach by developing a method for the analysis of secondary metabolites of fungi, namely mycotoxins, in food (Paper I and II). A particular material for the clean-up of mycotoxins for their quantification in milk and cereals was presented.

The topic of metabolomics for food quality has been treated in a suspected approach by developing a method for the analysis of several classes of secondary metabolites of plants such as phenolic acids, flavonoids, and glucosinolates in food (Paper III, IV and V). The development of a chromatographic and mass spectrometric method for the metabolic profiling of strawberry and cauliflower was presented.

The topic of metabolomics for nutrition has been treated in an untargeted approach by developing a method for the analysis of the human urinary metabolome after the consumption of meat and dairy products. Several data analysis approaches have been shown for the investigation of the whole urinary metabolome.

Ultimately, it has been shown how several aspects should be taken into account when analyzing complex matrices like food through different approaches. Sample preparation, chromatographic and mass spectrometric methods and data analysis have to be treated in different ways based on the used approach. In each case, the advantages of each technique can be exploited based on the purpose of the study. However, aside from the challenges that have to be faced, mass spectrometry-based metabolomics can definitely represent a powerful tool for food safety, quality and nutrition.

Chapter 1:

Background

1 Food safety, quality and nutrition

In the last years, the awareness of the impact of food on health has led to an increased interest on food safety and quality assessment. Research in food science and nutrition has grown parallel to the consumers' concern about the safety of the food they eat. The complexity of this issue is due to the globalization and the movement of food and related raw materials worldwide generating several contaminations. Moreover, many products contain multiple and processed ingredients, which are often shipped from different parts of the world, and share common storage spaces and production lines. As a result, ensuring the safety, quality, and traceability of food has never been more complicated and necessary than today.

The first goal of food science has traditionally been, and still is, to ensure food safety. The new European regulations in the European Union countries (e.g., Regulation EC 258/97 or EN 29000 and subsequent issues), the Nutrition Labeling and Education Act in the USA, and the Montreal Protocol have had a major impact on food laboratories. Consequently, more powerful, cleaner, and cheaper analytical procedures are now required by food chemists to detect and identify undesired compounds such as toxic or dangerous contaminants. Several classes of contaminants of both natural and synthetic origin can be detected in food. Detection and quantification at low levels of these contaminants is crucial for the assessment of the safety of the food available on the market [1].

Aside from the detection of toxic compounds, consumers demand is now moving toward the valorization of food with potential health benefits. Today, food is considered not only a source of energy but also an affordable way to prevent future diseases. The importance of the relationship between food and health is demonstrated by the development of several scientific fields such as nutraceutics. The characterization of food and the identification of compounds believed to benefit human health is crucial for both the food and the dietary integrators market, to provide customers with a healthy, balanced diet. To give an adequate answer to the rising consumer demands, food and nutrition researchers around the world are facing increasingly complex challenges that require the use of the best available science and technology. The new methodologies and the new generated knowledge derived from this trend is impressive and it includes, for example, the possibility to account for food products tailored to promote the health and well-being of population [1, 2].

Food scientists and nutritionists have to face a large number of challenges to adequately answer the new questions emerging from this new field of research. One of the main challenges is to improve our limited understanding of the roles of nutritional compounds at molecular level. Interaction of modern food science and nutrition with disciplines such as pharmacology, medicine, or biotechnology provides impressive new challenges and opportunities. Researchers in food science and nutrition are moving from classical methodologies to advanced analytical methodologies such as “omics” approaches and bioinformatics, frequently together with *in vitro*, *in vivo*, and/or clinical assays to investigate topics in food science and nutrition that were considered unapproachable few years ago [1].

2 Food safety

The Food and Agriculture Organization (FAO) and the World Health Organization (WHO) define the concept of "food security" as "physical and economic access by all people at all times to sufficient, safe and nutritious food to cover their dietary needs and food preferences for an active and healthy life" [3].

In developed countries, where access to enough food seems guaranteed, it is the aspect related to food safety the matter of concern. At the beginning of the century, the safety of the food became a priority for both consumers and public administrations. This was mainly due to the incredible change in production systems, technologies and eating habits. Intensive land use and industrial production led to the presence of chemical contaminants in food due to both environmental pollution and industrial food processing.

Although the main source of food contamination is due to abiotic pollution, resulting from the presence of chemicals and their metabolites, biotic contamination can also strictly affect food safety. Bacteria or fungus can indeed produce some dangerous toxins that could contaminate several kinds of foods.

In the next paragraphs, a brief overview of the typical contaminants that are often investigated in the area of food safety will be presented.

2.1 Abiotic contaminants

Scientific and technological advances have led to the development of new chemical compounds that, despite offering benefits in various areas, could also be potentially toxic. All these chemical compounds may enter into the environment in different ways and, being able to resist to photochemical and biological degradation, they persist in the environment and they bioaccumulate in living systems.

Environmental contaminants can be present in air, water or soil, then easily entering the food chain, posing an acute health risk if present at high concentrations. The major concern related to the presence of environmental contaminants in foods is their potential endocrine disruption, carcinogenic and other chronic effects.

Traditionally, attention to chemical contamination has focused on conventional pollutants, i.e., those regulated due to the large volume in which they are issued. However, it is noteworthy that another large group of chemical contaminants, not being present in the environment at high concentrations, have harmful effects on it and on the health of living systems. These pollutants are included within the so-called "emerging contaminants", a term that refers to all those pollutants that are not currently regulated by law in full, but might be candidates for a future regulation since it is believed or intuited that could be potentially harmful for environment or health [4, 5]. Within the chemical pollutants we can find different groups such as for example pharmaceuticals, personal care products, drugs of abuse, hormones, surfactants, heavy metals, polyfluorinated compounds, polycyclic aromatic hydrocarbons, brominated flame retardants, organochlorine compounds, plasticizers and additives[6]. Among all these classes of chemical pollutants, pesticides and veterinary drugs are two of the classes of major concerns when dealing with food safety.

The use of pesticides has become an integral part of modern agriculture to increase crop yields and quality by controlling various pests, diseases and weeds. Although pesticides are crucial for fighting hunger and disease, it cannot be ignored their toxicity to living organisms that can undergo their harmful effects through the environment or food treated with these compounds. The toxic effects of pesticides depend on the dose, route of exposure and exposure time. Approved uses of pesticides following good agricultural practices should result in pesticide residues below maximum residue limits established in a given country.

Several compounds belonging to the class of veterinary drugs can be detected in animal derived food. These drugs, which are administered to live animals, can remain as residues in

animal tissues. The major classes of veterinary drugs include antibiotics, anthelmintics, coccidiostats, nonsteroidal anti-inflammatory drugs, sedatives, corticosteroids, beta-agonists and anabolic hormones. Some of these drugs are carcinogenic, some other show endocrine-disrupting effects, certain antibiotics can cause severe allergic reactions in sensitive individuals and increase the risk that pathogenic microorganisms become resistant to antibiotics. Therefore, parent drugs and in some instances, their metabolites are either banned in animal-derived foods or regulated as maximum limits when the drugs are approved for use.

Finally, other contaminants in food could be derived from processes like food processing, migration of contaminants from packaging to food or food additives and food adulteration.

2.2 Biotic contaminants

Toxins are naturally occurring substances that can be classified in several groups based on the organisms they are produced by. They can be either plant toxins, bacterial toxins, marine biotoxins (phycotoxins) or fungal toxins (mycotoxins). While the bacterial/fungal contamination can be eliminated with heat treatment, the toxins can remain permanently in the food product as contaminants, representing one of the major concerns in foods.

Among plant toxins there is a large number of structurally similar compounds such as the about 500 pyrrolizidine alkaloids. Plant toxins are still analytically interesting because several studies showed their presence in several foods[7]. Also, marine biotoxins, highly toxic compounds produced by phytoplankton, are monitored and regulated by certain limits because of their toxicity to humans. Finally, mycotoxins are secondary metabolites produced by filamentous fungi (molds) that can colonize various crops, resulting to be pose a health concern for humans.

Mycotoxins are formed frequently in field crops during the growing season and are unavoidable as a result of our inability to control the environmental and crop susceptibility factors that allow for their production by toxigenic, plant pathogenic fungi. They are of concern mainly in cereals, nuts, infant formula, milk, dried fruit, coffee, fruit juice and wine.[8] Fungal toxins comprise one of the major sources of food poisoning in humans, agricultural animals and wildlife. The impacts of these naturally occurring toxins range from acute disease and deaths to chronic disease with production of tumors and significant

immunosuppression underlying the occurrence of other infectious bacterial and fungal diseases. The risks associated with these mycotoxins occurring in field crops, foods, feeds and animal products impacts both food safety and economic aspects of the marketing of grain for food and feed (Figure 1.1). Thus, food and feed manufacturers have a great concern for the occurrence of mycotoxins in the commodities used in their products. There are many mycotoxins, but only a few are currently regulated, with the European Union, having a more comprehensive list than most other countries, which includes aflatoxins, ochratoxin A, patulin, deoxynivalenol, zearalenone, fumonisins and T-2/HT-2 toxins[9]. Despite the limited number of regulated mycotoxins, there is considerable interest in a more widespread screening, as there is evidence to suggest that a wide range of fungal metabolites and conjugates can occur naturally in foods.

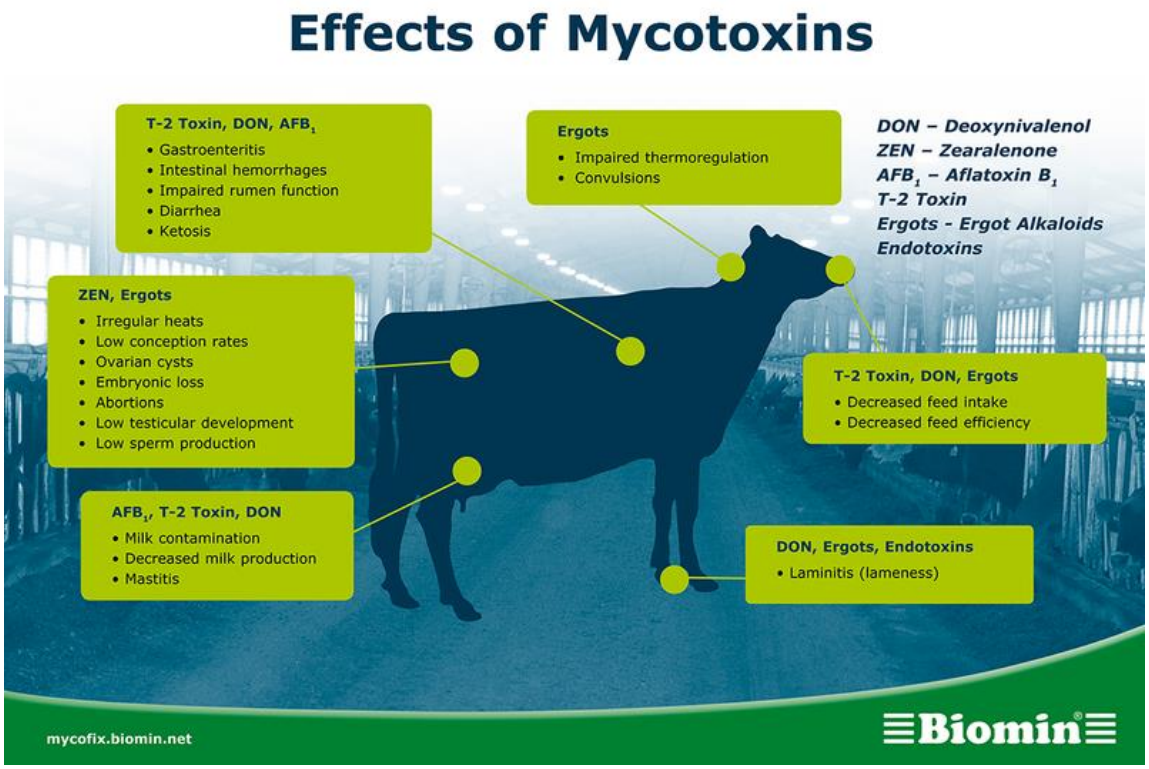


Figure 1.1: Mycotoxins effects on cows when fed with contaminated feedings [10].

3 Food quality

The term “food quality” includes multiple aspects of food such as composition, aroma, flavour, taste, or food properties like bioactivity. One of the key aspects in assessing food

composition is aroma determination, due mostly to volatile compounds, and taste determination, due mostly to non-volatile compounds. Compounds contained in foods are of great importance as they are highly relevant for the perceived overall quality of a particular product. Besides the sensorial aspects, assessing the composition of food is important to obtain a comprehensive profiling of a specific food product, that it is essential in order to assess food variety and origin. Finally, the comprehensive profiling of a food product, together with an investigation on the relationship between its composition and the effect on human health, has an important role in the assessment of food quality. The detection and identification of bioactive compounds can be really helpful for nutraceuticals isolation or production.

3.1 Bioactive compounds

In the recent years, several studies showed that food does not only provide adequate nutrients to meet the metabolic requirements of the body, but it can also contribute to the improvement of human health. The identification of compounds believed to benefit human health is crucial to provide a healthy balanced diet.

These compounds are usually referred to as bioactive compounds or nutraceuticals. Biesalski *et al.* [11] defined as bioactive those "compounds that are essential and nonessential (e.g., vitamins or polyphenols) that occur in nature, are part of the food chain, and can be shown to have an effect on human health". More specifically, the compounds showing positive effects instead should be designated by the term "nutraceuticals" [12]. Nutraceuticals can be found in foods of both animal and vegetal origin (the latter being currently referred to as "phytochemicals"). The term "phytochemicals" refers to plant-derived compounds that are not considered essential nutrients, but have been demonstrated to have beneficial health effects. Plants synthesize a vast range of organic compounds that are traditionally classified as primary and secondary metabolites. Primary metabolites are compounds that have essential roles associated with photosynthesis, respiration, and growth and development. These include phytosterols, acyl lipids, nucleotides, amino acids and organic acids. Other phytochemicals are referred to as secondary metabolites and their function appears to have a key role in protecting plants from herbivores and microbial infection, as attractants for pollinators, as UV protectants, etc. Secondary metabolites are

also of interest because of their use as dyes, fibers, glues, oils, waxes, flavoring agents, drugs and perfumes, and they are viewed as potential sources of new natural drugs, antibiotics, insecticides and herbicides [13] [14]. In recent years, the role of some secondary metabolites as protective dietary constituents has become an increasingly important area of human nutrition research. There is an increasing evidence that modest long-term intakes of some phytochemicals can have favorable impacts on the incidence of cancers and many chronic diseases, including cardiovascular disease and Type II diabetes. Based on their biosynthetic origins, plant secondary metabolites can be divided into three major groups: flavonoids and allied phenolic and polyphenolic compounds, nitrogen-containing alkaloids and sulphur-containing compounds and terpenoids [15].

3.1.1 Phenols, phenolic acids, polyphenols

Plant phenolics are biosynthesized almost exclusively via the shikimate pathway (Figure 1.2). The shikimate pathway is an essential metabolic route by which plants, fungi and bacteria synthesize the aromatic amino acids phenylalanine, tyrosine and tryptophan, as well as a number of other aromatic compounds which are critical to sustaining the primary functions of living organisms. The intermediates of the shikimate pathway also serve as starting points for biosynthesis of a vast number of plant and microorganism secondary metabolites which, although not critical to the immediate survival of the producing organism, however help it in different ways to maintain its position in ecological environment [16]. The majority of these shikimate-derived substances are formed from the end products of the shikimate pathway, i.e. the aromatic amino acids. Phenolics are one of the most important groups of secondary metabolites in plants. From a structural point of view, phenolic compounds include a wide range of substances: simple phenols, phenolic acids, phenylpropanoids, coumarins, quinones, flavonoids, tannins, and other miscellaneous phenols [17]. Phenolics having a basic structure of a three-carbon side chain on an aromatic ring (C6-C3 skeleton) are known collectively as phenylpropanoids and are probably the most common shikimate metabolites. In plants, phenylalanine ammonia lyase converts phenylalanine to *trans*-cinnamic acid, which is the precursor of thousands of phenolic compounds, including simple phenolics, phenylpropanoids, and flavonoids. Hydroxylation of cinnamic acid produces *p*-coumaric acid, which can be further hydroxylated and

methylated to form caffeic, ferulic, and sinapic acids. These simple phenylpropanoids are the building blocks of the complex polymers such as suberins and lignins. Stilbenes and flavonoids are formed by condensation of a phenylpropane unit with a unit derived from acetate via malonyl-coenzyme A. Flavonoid metabolism (Figures 1.2 and 1.3) is composed of a number of branch pathways leading to isoflavonoids, flavonols, proanthocyanidins (condensed tannins), and anthocyanins [18].

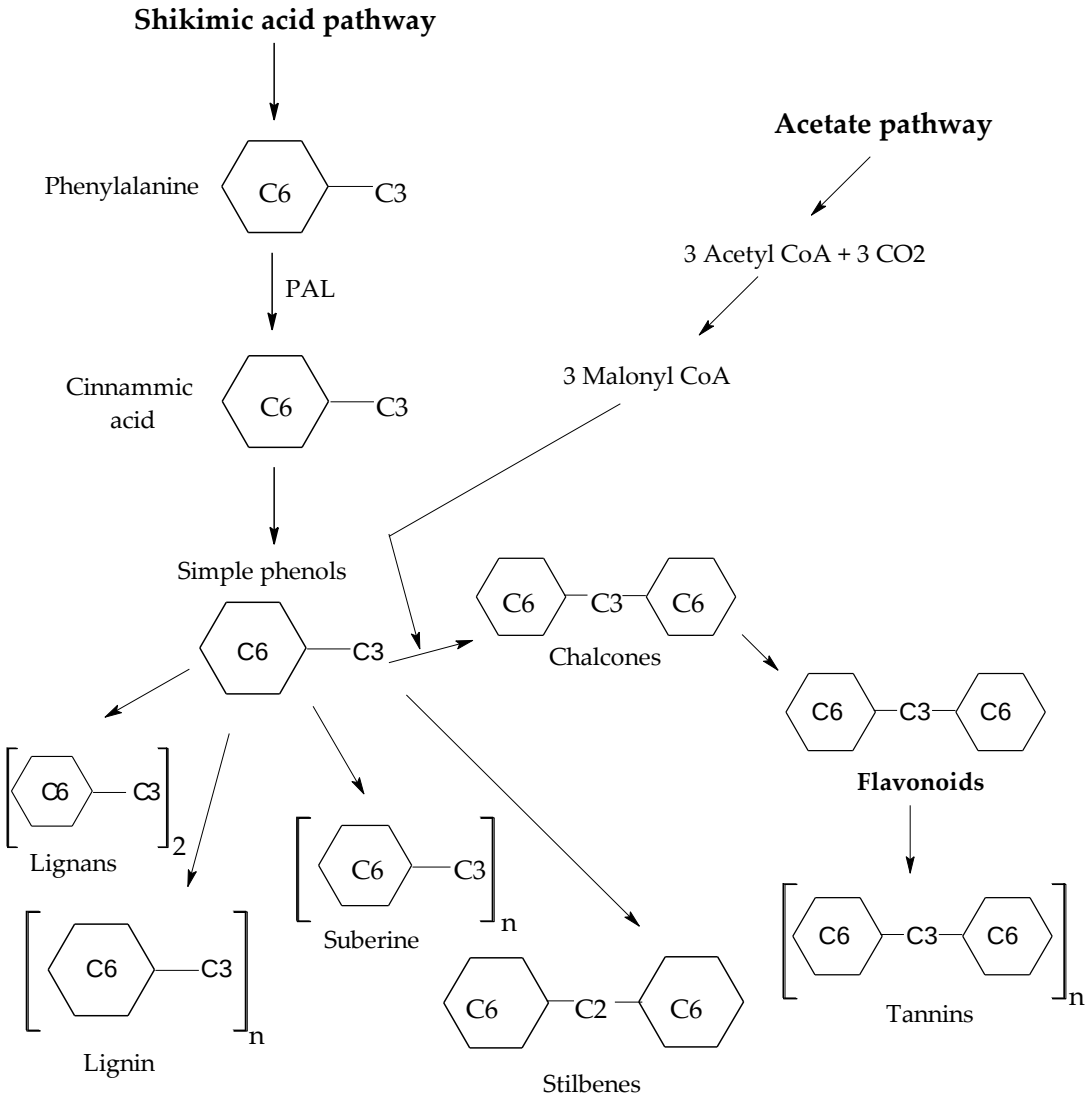


Figure 1.2: End products of shikimate and acetate pathways.

Phenolic acids are organic acids containing two distinguishing constitutive carbon frameworks: the hydroxycinnamic and hydroxybenzoic structures. Even if the basic skeleton remains the same, the numbers and the positions of the hydroxyl groups on the aromatic ring create the variety [17].

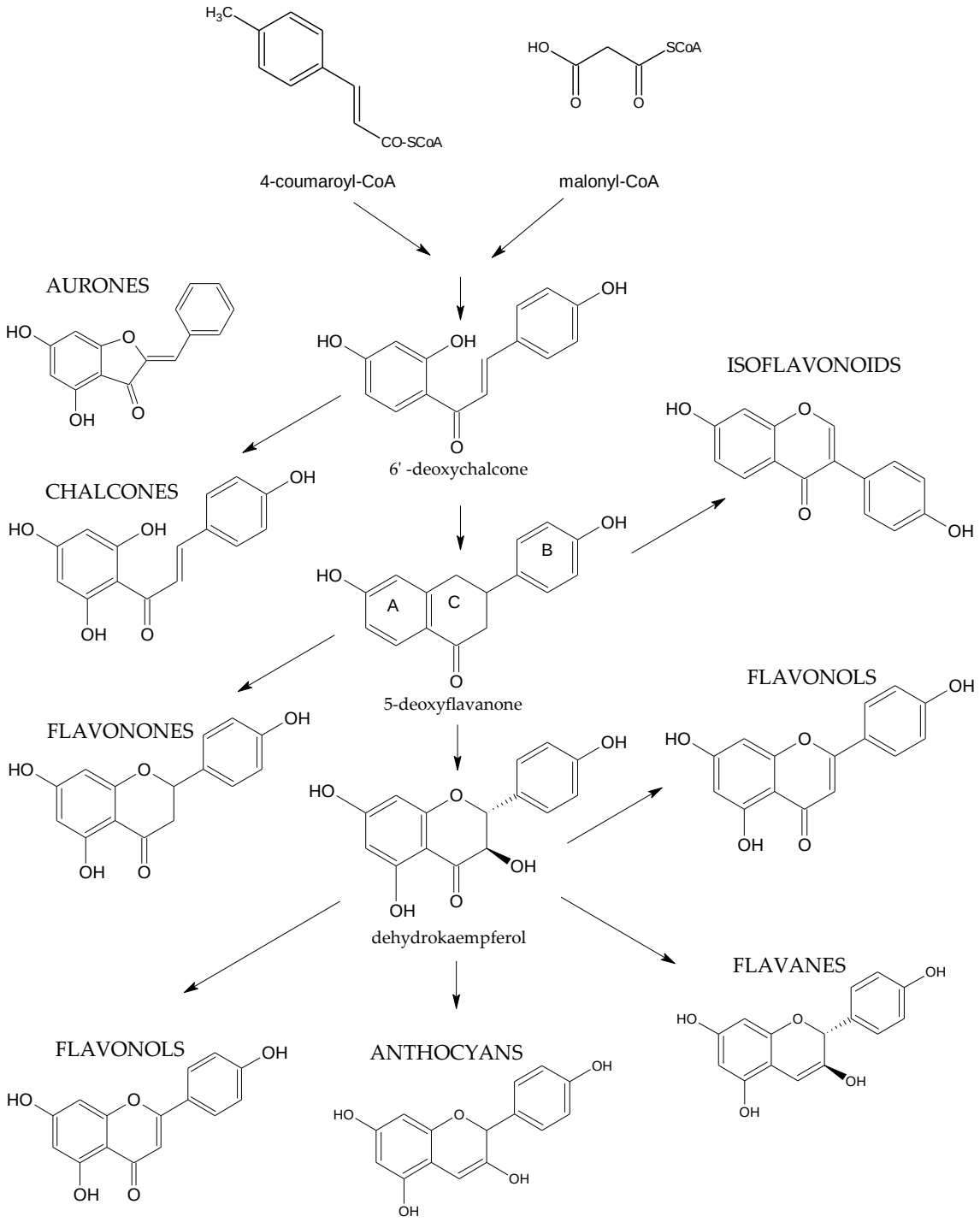


Figure 1.3: Flavonoid biosynthesis

Hydroxycinnamic acids are ubiquitously found in fruits, vegetables, and cereals [19], and the major are *p*-coumaric, caffeic, ferulic, and sinapic acids.The benzoic acid derivatives, such as gallic, syringic, vanillic, and salicylic acids, although not strictly phenylpropanoids themselves because they lack the three carbon side chain, originate from the phenylpropanoids cinnamate and *p*-coumarate [20]

Flavonoids are polyphenolic compounds comprising fifteen carbons, with two aromatic rings (rings A and B in Figure 1.3) connected by a three-carbon chain that forms a closed pyran ring (heterocyclic ring containing oxygen, the C ring) with ring A. Therefore, their structure is also referred to as C6-C3-C6. They can be subdivided into different subgroups depending on the carbon of the C ring on which B ring is attached, and the degree of unsaturation and oxidation of the C ring. Flavonoids in which B ring is linked in position 3 of the ring C are called isoflavones; those in which B ring is linked in position 4, neoflavonoids, while those in which the B ring is linked in position 2 can be further subdivided into several subgroups on the basis of the structural features of the C ring. These subgroup are: flavones**,** flavonols, flavanones, flavanonols, flavanols or catechins and anthocyanins. Finally, flavonoids with open C ring are called chalcones. In each subgroup flavonoids can be differentiated based on the substituents on rings A and B [21]. Hydroxyl groups are usually present at the 4, 5 and 7 positions. Sugars are very common with the majority of flavonoids existing naturally as glycosides. Whereas both sugars and hydroxyl groups increase the water solubility of flavonoids, other substituents, such as methyl groups, make flavonoids lipophilic. Their biological activities depend on the structural features of the ring B and the patterns of glycosylation and hydroxylation of the three rings, making the flavonoids one of the most diversified groups of phytochemicals [22]. Although flavonoids are ubiquitous in higher plants, their substitution with specific groups may be peculiar to certain species.

Tannins are a group of polyphenols possessing 12±16 phenolic groups and 5±7 aromatic rings per 1000 units of relative molecular mass [23]. This feature, together with their high molecular weight, clearly makes the tannins and similar phenolic polymers different both in structure and properties from the low-molecular-weight phenolic acids and monomeric flavonoids [24]. The observation that many tannins can be fractionated hydrolytically into their components led to the classification of such tannins as ‘hydrolyzable tannins’, whereas nonhydrolyzable tannins are classified as condensed tannins. Ellagitannins (ETs), which belong to the hydrolyzable tannin class of polyphenols, are complex derivatives of ellagic

acid. Hydrolysis of ETs with acids or bases yields hexahydroxydiphenic acid (HHDP), which spontaneously lactonizes to ellagic acid (Figure 1.4). ETs constitute a complex class of polyphenols characterized by one or more HHDP moieties esterified to a sugar, usually glucose and usually linked to galloyl residues. ET compounds have an enormous structural variability due to the different possibilities for the linkage of HHDP residues with the glucose moiety, and to their tendency to form oligomeric derivatives. The antioxidant efficiency of ETs and ellagic acid is directly correlated with their degree of hydroxylation and decreases with the presence of a sugar moiety. They are abundant in some fruits, nuts and seeds such as pomegranates, raspberries, strawberries and almonds [25].

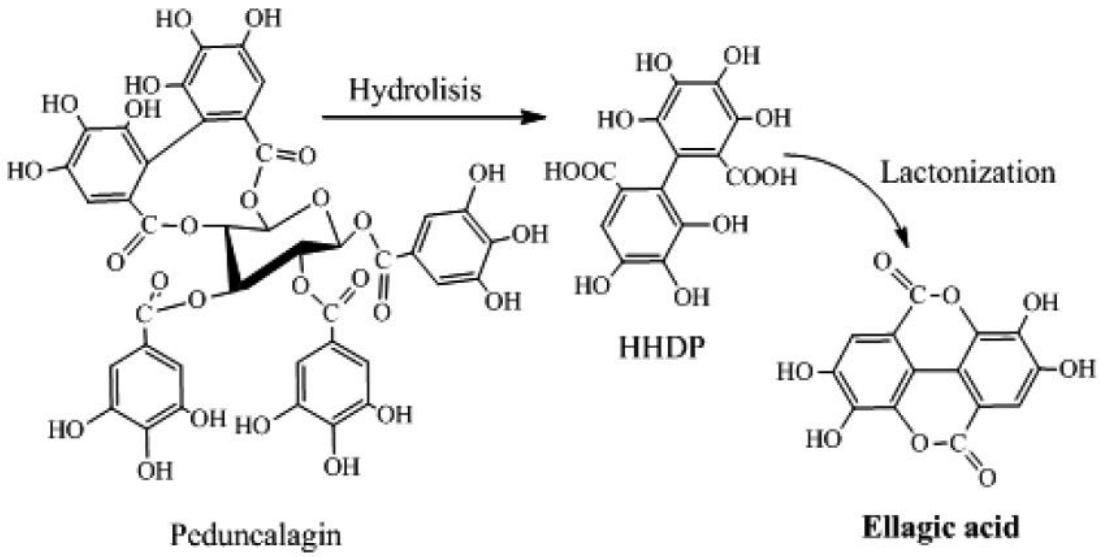


Figure 1.4: Hydrolysis and spontaneous lactonization of HHDP in ETs [26].

Proanthocyanidins (PAs) belong to the class of condensed tannins. They can be divided into several classes based on the hydroxylation of their constitutive units (flavan-3-ols) and the linkages between them. The most common constitutive units are (epi)catechins, (epi)gallocatechins, and (epi)afzelechins, leading to procyanidin, prodelphinidin, and propelargonidin structures, respectively (Figure 1.5). Flavan-3-ol units are most frequently linked via B-type bonds, that is, C4-C8 or C4-C6 linkages. Occasionally, an additional C2-O7 or C2-O5 linkage may exist, leading to doubly bonded A-type PAs. The physical, chemical, and biological features of PAs depend largely on their structure and particularly on their degree of polymerization. PAs are powerful antioxidants, but they have also been reported to demonstrate antibacterial, antiviral, anticarcinogenic, anti-inflammatory, and

vasodilatory activities. They are widely present in the plant kingdom, for example, in fruits, berries, nuts, seeds, and bark of pine trees [27].

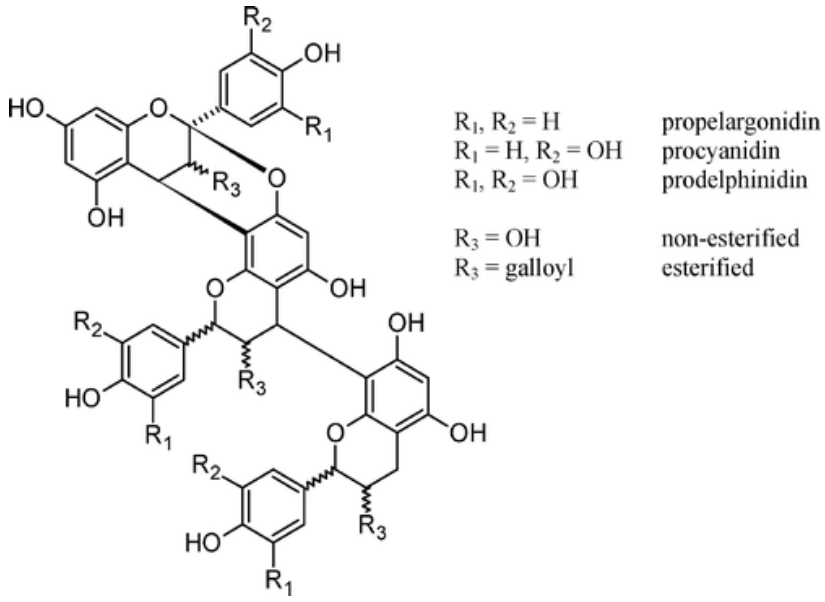


Figure 1.5: Typical structure of an oligomeric PA [27].

3.1.2 Glucosinolates

Glucosinolates (GLSs) have been found mainly in the order Capparales, which includes broccoli, rape and most notably salad and herb species, such rocket salad or rucola, of the Brassica genus [28, 29]. They are nitrogen- and sulfur-containing secondary metabolites playing crucial roles in plant development as well as in the interaction of a plant with its biotic and abiotic environment. GLSs, also known as (Z)-N-hydroximinosulfate esters, consist of a common glycone group and a variable aglycone side chain (R) derived from amino acids, in particular methionine, phenylalanine, tyrosine, and tryptophan [30]. The glycone is characterized by a thioglucose and a sulfonated oxime group, both attached to the C-2 carbon of the parental amino acid. The thioglucose can be conjugated to some phenolic acids such as cinnamic acid, coumaric acid, etc. All GLSs are permanently negatively charged as a common sulfate moiety is present (Figure 1.6). To date, more than 120 different GLSs have been identified and classified as aliphatic, aromatic, v-methylthioalkyl and heterocyclic (e.g., indole) according to the type of side chain (R) [29]. GLSs are hydrolyzed

by endogenous thioglucosidases, called myrosinases, to produce a wide range of degradation products (isothiocyanates, nitriles, epithiocyanates, oxazolidine- 2-thiones, and thiocyanates) with diverse biological activities. Frequent consumption of high-GLS-content vegetables is associated with a lowered risk of cancer and cardiovascular disease [31].

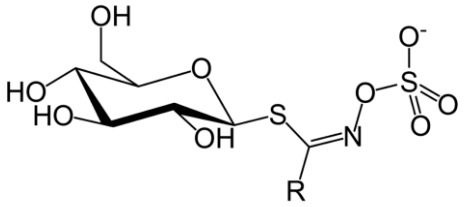


Figure 1.6: Typical structure of a generic GLSs, where the –R moiety can be substituted by a wide class of groups.

4 Food and nutrition

Nowadays, food is investigated not only as a source of energy but also as a potential health promoter. Interest in nutrition and health has grown considerably, as evidenced by a rapid proliferation of studies examining nutritional exposures in relation to several diseases [32].

The first studies on diet and health evaluated correlations between country-specific diseases rates and food consumption patterns. These ecologic studies utilized aggregate data to identify possible relationships and generate hypotheses. However, ecologic studies are highly susceptible to confounding, where the observed relationship between two factors is actually caused by a third (or more) factor(s). Case-control studies, in which diseases cases are compared, for a limited period, with healthy controls derived from the same source population from which the cases arose, improve on ecologic studies. However, even this approach is problematic because dietary assessment occurs after diagnosis. A key development in nutritional epidemiology over the past two decades has been the increase in results from large prospective cohort studies. In these studies, a cohort of healthy individuals is assembled and exposures are assessed at baseline, then the cohort is followed over time and diseases cases are identified as they develop. A disadvantage of cohort studies is that they must be very large and have long follow-up time to accrue sufficient numbers of disease cases. In addition, many dietary factors are consumed together, which can make it difficult

or impossible to isolate the effect of one specific factor. In contrast to observational studies, in experimental studies the investigator controls the assignment of subjects to the treatment or control group. Randomized clinical trials are often considered the "gold standard" when testing hypotheses because proper randomization theoretically results in treatment and control groups that are similar in all ways except the intervention, and observed associations can be directly attributed to the intervention. However, a randomized clinical trials can only test a specific intervention, in a specific population, over a relatively short period of time. A common approach in nutritional epidemiology has been to use results of observational studies to identify nutrients or other dietary factors that modify the risk of diseases, and then test these factors in a clinical trial. Observational and experimental studies both contribute to our knowledge of nutrition and health, and no single study can provide a definitive answer.

Finally, measurement error is another considerable problem for both case-control and cohort studies of diet. There are three main methods of assessing diet in epidemiologic studies: the food diary (participants record what they consume over a period of time); the dietary recall (participants report what they consumed during a preceding period); and the food frequency questionnaire (listing of foods for which participants report the usual frequency and amount of consumption)n. Dietary assessment relies on the ability of individuals to recall a complex collection of exposures and is known to contain measurement error. The use of food frequency questionnaires in studies of diet and diseases and the possible impact of measurement error, needs to be considered when interpreting results of observational studies. Due to this complexity it is probably too early to conclude on the relationships between many substances and health. Several of the health benefits assigned to many dietary constituents are still under controversy as can be deduced from the large number of applications rejected by the European Food Safety Authority (EFSA) [33] about health claims of foods and ingredients [34]. More scientific evidences are needed to demonstrate the claimed beneficial effects of the constituents of food.

Chapter 2:

Liquid chromatography-mass spectrometry for food safety, quality and nutrition.

1 Liquid chromatography-mass spectrometry developments

Over the past 20 years there have been a number of key developments in mass spectrometry (MS) which have led to step-changes in the way food analysis has been conducted [35].

The development of electrospray (ESI) and atmospheric pressure chemical ionization (APCI) interfaces, together with progressive improvements in MS designs, opened up liquid chromatography-mass spectrometry (LC-MS) to the wider analytical community, because of the possibility of ionize and detect the widest range of molecules. Latterly, further improvements have seen movement from LC-MS to LC tandem MS (MS/MS), which has now become the standard approach to characterize the molecular structure of the analytes. Another step change has been the development of high-resolution mass spectrometry (HRMS) through time-of-flight (TOF) instruments in the form of LC-TOF MS in various configurations [36], together with the development of Orbitrap technology [37], that allowed the detection of masses until 1-2 ppm of error, thus giving the possibility of identify unknown compounds. The same trend with HRMS has been shown, in the last few years, in terms of software being progressively improved. HRMS together with dedicated software for data analysis allowed a more comprehensive analysis and characterization of complex matrices. Finally, the coupling of MS with ultra-high performance liquid chromatography (UHPLC) also allowed to obtain new important results for the analysis of complex biological matrices, such as food products. Several advantages and limitations of the different LC-MS techniques in their application to food safety and quality have been assessed [38].

In the next paragraphs, LC and MS techniques will be discussed in detail, together with their application in food safety, quality assessment and in nutrition.

2 Liquid chromatography

When analysing complex mixtures by means of MS, a good upstream separation process is critical for detection of as many compounds as possible. Chromatographic separation is needed when using ESI, to avoid the ionization suppression observed when too many

substances are present in the ionization source simultaneously (see the next paragraphs). In this context, chromatographic techniques are the most suitable for small molecules separation because several differences among molecules can be exploited, such as differences in hydrophobicity and charge. The LC is probably the most widely used separation technique in both analytical and research laboratories. It is used for the separation and determination of compounds in very different areas of science such as food analysis.

The fundamental chromatographic principle involves the different affinity of a compound to two phases, one stationary and one mobile phase located within a separation column. Analytes move through the chromatographic column by means of a certain flow of mobile phase. The column contains a stationary phase, sometimes in the form of polymerized monoliths or, more often, small particles which carry functional groups. Each compound undergoes a different deal between the stationary phase and the mobile phase. If the analyte-stationary phase interaction is stronger than the analyte-mobile phase interaction, the analyte will be attracted to the stationary phase and will thus be retained, otherwise it will be eluted by the mobile phase. In practical terms, the stationary phase is kept constant, whereas the mobile phase composition is varied in a gradient over time. Hence, physicochemical distinct molecules can be separated in timely manner by sequential elution off the column.

Depending on the interaction phenomenon that occurs with the stationary phase, it is possible to differentiate various types of LC: distribution, ion exchange, adsorption, molecular exclusion, affinity or chiral. Given the high resolving power and volatile solvent components compatible with online coupling to the mass spectrometer, the stationary phase generally used in LC-MS setups is reversed-phase. In the typical reversed-phase chromatography setup, columns are packed with silica beads which are functionalized with hydrophobic C18 alkyl chains. In reversed-phase the elution is based on hydrophobicity, and it is achieved by increasing the concentrations of the organic modifier (e.g. acetonitrile ACN) in the mobile phase.

Among the major drawbacks of conventional LC, despite its high popularity, there are the high consumption of solvents and sample needed to perform the chromatographic separation and the low sensitivity obtained with some detection systems due to dilution of the sample in the mobile phase. These drawbacks can be largely resolved by chromatographic columns of reduced internal diameter and packed with smaller sized particle. This LC mode, known as UHPLC, allowed to obtain several advantages such as: lower consumption of sample, mobile phase, and stationary phase; increased thermostatting

capacity, which allows maintaining a certain constant temperature throughout the column; increased sensitivity due to the fact that the lower volumes of mobile phase used (due to the smaller internal diameter of columns) make a smaller dilution of the sample and thus a more concentrated sample reaches the detector; increased efficiency because of a greater number of theoretical plates and thus decreased peak width; higher permeability achieved with the use of smaller particles. Moreover, UHPLC chromatography also makes easier to interface the chromatographic system to the mass spectrometer. UHPLC development has required several modifications of the classical HPLC instrumentation. The most obvious is the construction of systems capable of handling the increased pressure. The new UHPLC instruments are capable of operating at pressures up to 1500 bar, compared to about 200 bar for traditional HPLC. It is also important to reduce the dead-volume sites in the system, to limit peak broadening. Electronics have also been upgraded to accommodate faster data collection as the peak widths can be as much as 10 times smaller than that seen with traditional HPLC systems, thus requiring a much more rapid data collection interface[39, 40].

3 Mass spectrometry

MS is a highly sensitive instrumental technique and its use as a detector offers clear advantages over detectors traditionally used in LC. Firstly, the high sensitivity allows to determine compounds usually present at very low concentrations. Secondly, the possibility of determine the mass of the analytes, together with the fragmentation pattern, provides the unambiguous identification of the compounds of interest.

MS fundamental task is the separation of previously ionized molecular or atomic particles according to their mass to charge ratio *(m/z).* The process occurring in a mass spectrometer comprises three stages, as shown in Figure 2.1: sample ionization to charge the analytes, that takes place in the ion source; acceleration of the ions by an electric field and separation according to their *m/z* ratio, that takes place in the analyzer; detection of ions and production of an electrical signal taking place in the detector [41].

In the next paragraphs, a brief overview of the instruments that will be mentioned in the next chapters will be presented.

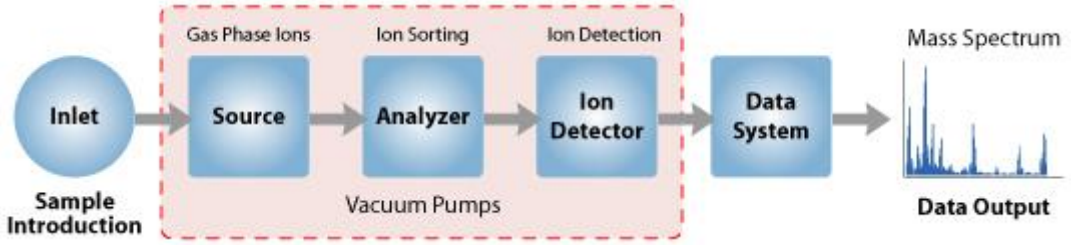


Figure 2.1. Basic scheme of a mass spectrometer[42].

3.1 Ionization sources

The main ionization sources available for the analysis of polar-medium polar compounds are matrix assisted laser desorption ionization (MALDI), APCI, and ESI. In the present work, only ESI source has been employed.

3.1.1 ESI

Despite numerous sources currently available for ionization, ESI ionization is the technique that allows the ionization of the widest range of molecules, therefore it is the most commonly used. In Figure 2.2 the ESI ionization process is schematically shown.

In the initial step of the ionization process, the analyte solution passes through a thin conductive capillary to which a voltage of a few kV is applied. The influence of the electric field on the liquid on the tip of the capillary, in conjunction with the surface tension, causes the formation of a Taylor cone. The potential difference between the capillary and the counter electrode pulls the positive ions to the surface of the solvent. Once the electric field overcomes the surface tension, the Taylor cone emits a solvent jet. The tip of the jet is inherently unstable and collapses into charged droplets. A number of parameters, including the applied potential, the solution flow rate and solvent properties, influences the diameter of the formed droplets.

Solvent evaporates from these droplets until the charge density on the surface of the droplets reaches the Rayleigh limit. Once this limit is passed, the Coulombic repulsion force is greater than the surface tension causing the droplet to undergo fission into smaller droplets, the so called Coulombic explosion. These small, highly charged droplets will continue to lose solvent, and when the electric field on their surface becomes large enough, desorption of ions from the surface occurs. As ESI is a soft ionization source, gaseous ions formed are generally protonated or deprotonated molecules depending on selected ionization polarity mode [43]. Once the compounds are ionized, they can be separated according to their *m/z* ratios in the mass analyzer.

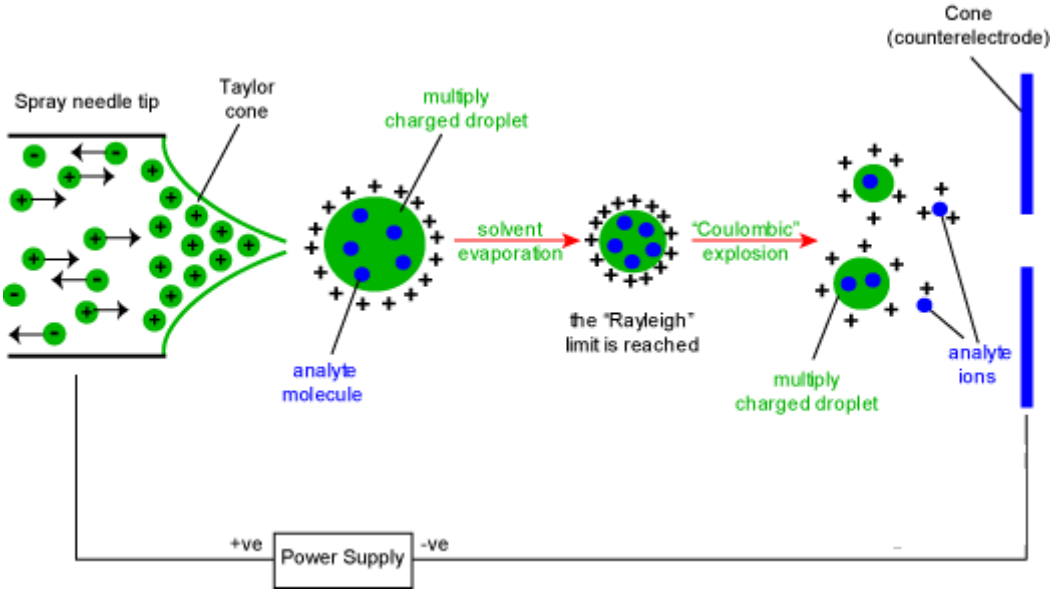


Figure 2.2. ESI ionization process [44].

ESI ionization is a liquid phase ionization because the ionization takes place in the solution itself. This is one of the main advantages of this ionization technique, since it enable the on-line coupling of LC to a mass spectrometer. As mentioned before, the upstream chromatographic separation is required to be efficient. Indeed, when too many analytes are ionized simultaneously, whether it is due to more abundant molecules or contaminants, only the most hydrophobic molecules are ionized [45]. Thus, a good separation is needed to avoid the ionization suppression and allow an efficient ionization of the analytes of interest.

3.2 Mass analyzers

Several mass analyzers have been proposed during history such as ion trap (IT), quadrupole, Fourier transform (FT), TOF and Orbitrap. Among the different analyzers, Triple quadrupole (QqQ), Orbitrap and quadrupole-TOF (qTOF) were used in the PhD project, therefore a detailed description of these instruments will be given in the following paragraphs.

3.2.1 Triple quadrupole

The quadrupole analyzer is a device that uses the stability of the trajectories of ions in oscillating electric fields to separate them according to their *m/z* ratios.

Quadrupole analyzers are made up of four perfectly parallel cylindrical rods serving as electrodes as shown in Figure 2.3. A direct current is applied to the four rods, whilst a variable radiofrequency is applied to each pair of bars. Therefore, ions are subjected to the influence of a total electric field made up of a quadrupolar alternative field and a constant field. When a positive ion enters the space between the rods it will be drawn towards a negative rod. If the potential changes sign before it discharges itself on this rod, the ion will change direction. Adjusting the potentials of the direct current and the variable radiofrequency, only ions with a specific m/z will be able to pass the quadrupole without touching the rods.

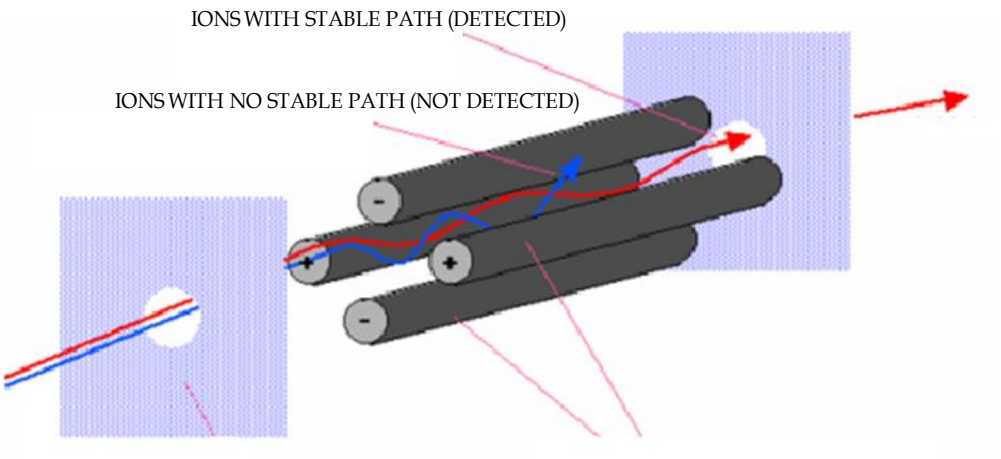


Figure 2.3. Diagram of a quadrupole.

Therefore, when the ions coming from the ionization source reach the quadrupole, only those ions having a certain m/z, and then a stable trajectory between the four bars, will reach the detector[41].

The quadrupole is the most simple, economical, compact and robust analyzer. However, it is a low-resolution instrument, meaning it is only able to distinguish ions differing by one unit mass. Moreover, quadrupole sensitivity is quite low when monitoring a large number of ions. However, if only one ion, or a limited number of ions, is analyzed, the sensitivity of the detector can increase. In this context, the use of MS/MS can largely increase the selectivity, thus increasing signal to noise ratio.

When three quadrupoles are combined together, in the configuration QqQ, MS/MS experiments can be performed. Figure 2.4 shows the general scheme of a QqQ system that consists of two quadrupoles as analyzers and a quadrupole in the centre that acts as collision cell for the fragmentation of the ions. In the first quadrupole a specific ion which is known as "precursor ion" is selected; the precursor ion passes into the second quadrupole which, by means of collision of molecules of an inert gas (He, N_2 or Ar) with the precursor ion, fragments it in "product ions"; product ions are then separated in the third quadrupole and detected generating an MS/MS spectrum.

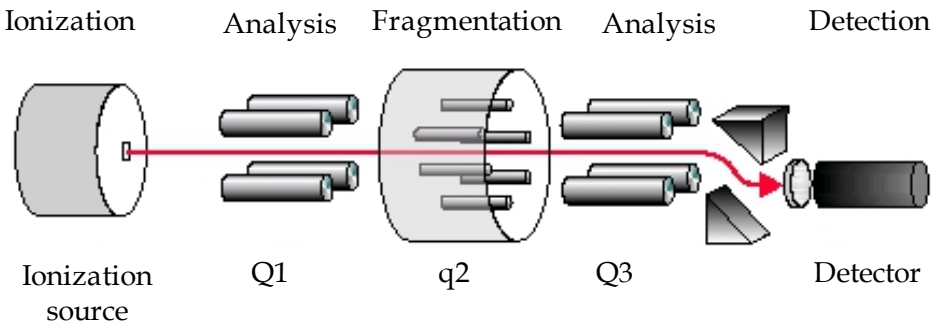


Figure 2.4. QqQ mass spectrometer.

The main advantage of MS/MS experiments is the improvement in selectivity and sensitivity, thus making the QqQ the most versatile analyzer for quantitative analysis. This analyzer has the inherent qualities of a quadrupole, such as easy handling and control, small size and relatively low cost. The QqQ analyzer offers a great versatility of acquisition modes as shown in Figure 2.5:

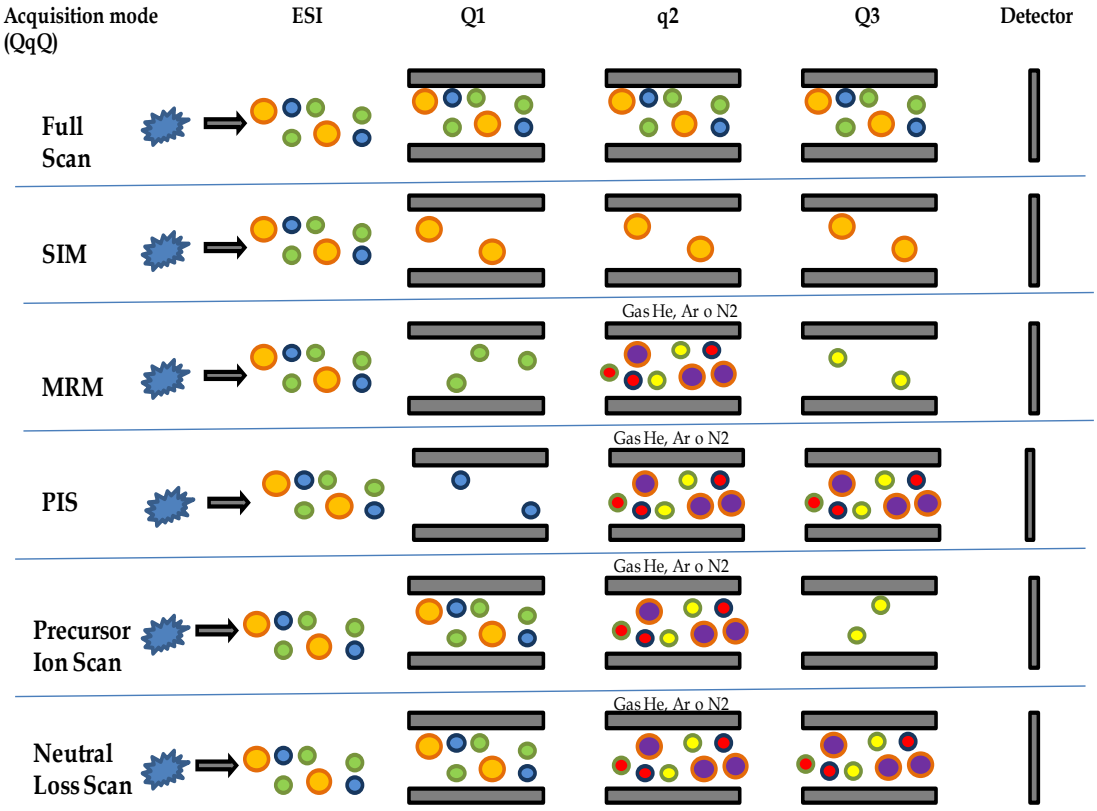


Figure 2.5. Acquisition modes of a QqQ instrument.

-Full Scan acquisition mode consists in the scanning over a whole range of *m/z* values, detecting all ions that pass through the spectrometer.

-Single Ion Monitoring (SIM) consists in the selection of a specific m/z ratio ion in the first analyzer. Only the selected ion will be detected. The absence of scanning allows to focus on the precursor over longer times, increasing the sensitivity compared to the Full Scan mode.

-Selected Reaction Monitoring (SRM) or Multiple Reaction Monitoring (MRM) consists of selecting a fragmentation reaction. An ion with a specific *m/z* ratio is selected in the first analyzer, fragmented in the collision cell, and only a specific ion produced by the fragmentation of the precursor will be selected and detected. The ions selected by the first mass analyzer are only detected if they produce a given fragment. The absence of scanning allows to focus on the precursor and fragment ions over longer times, increasing the sensitivity as for selected ion monitoring, but this sensitivity is now associated with a high increase in selectivity.

-The Product Ion Scan (PIS) consists of selecting a precursor ion of a chosen *m/z* ratio and determining all the product ions resulting from fragmentation.

-The Precursor Ion Scan consists of choosing a product ion and determining the precursor ions. This scan requires the focusing of the second spectrometer on a selected ion while scanning the masses using the first spectrometer. All the precursor ions that produce ions with the selected mass through fragmentations thus are detected.

-The Neutral loss scan consists of selecting a neutral fragment and detecting all the fragmentations leading to the loss of that neutral. This scan requires that both mass spectrometers are scanned together, but with a constant mass offset between the two. Thus, for a mass difference *a*, when an ion of mass *m* goes through the first mass spectrometer, detection occurs if this ion has produced a fragment ion of mass (*m*−*a*) when it leaves the collision cell.

Each of these modes of acquisition has characteristics that make it suitable for a particular purpose. However, the acquisition in MRM mode is the most used in quantitative analysis, to determine a large number of compounds while minimizing the presence of interfering substances.

3.2.2 Orbitrap

The Orbitrap mass analyzer consists of an inner spindle-shaped central electrode surrounded by a pair of bell-shaped outer electrodes (Figure 2.6). The electric field created by both electrodes traps the ions inside the analyzer. The ions rotate around the central electrode and oscillate along its axis. The movement is harmonic and the frequency ω depends only on the *m*/z ratio and an instrumental constant *k*:

$$\omega = \sqrt{\frac{ze}{m}} \text{ x } k$$

The oscillation induces an image current in the outer halves of the Orbitrap. The image current is then transformed in a mass spectrum by FT[41].

Orbitrap instruments are always hybrid mass spectrometers. Several hybrid instruments are available in the market in which the combination of different analyzer allows several possibilities. In the works described in this thesis a Q Exactive Orbitrap mass spectrometer has been employed.

The Q Exactive mass spectrometer is usually equipped with an ESI source and includes, a stacked-ring ion guide (S-lens) in the source region, a quadrupole mass filter, a C-trap, an

higher-energy collisional dissociation (HCD) cell, and an Orbitrap mass analyzer as shown in Figure 2.7. Ions are formed in the ion source, pass through the ion transfer tube where the pressure is gradually reduced from the atmospheric values of the source to void of the mass spectrometer. Ions pass from the transfer tube to an S-lens described and then via an injection multipole into a bent flatapole. The bent flatapole is oriented in such a way that the line of sight from the S-lens is open for clusters and droplets to fly unimpeded out of the flatapole. Ion optics use a combination of direct current, radio frequency voltage and a vacuum gradient to guide the ions towards a hyperbolic quadrupole, capable of isolating ions down to an isolation width of 0.4 Th at *m/z* 400. A short octapole then brings ions into the C-trap. The C-trap is the key of the instruments, because allows accumulation of ions, before injection into the Orbitrap and permits to interface continuous ion sources (like ESI) to a discontinuous analyzer. The c-trap is also interfaced to an HCD cell. In the gas-filled HCD cell, fragmentation of ions is achieved by adjusting the voltage offset of the rods and the axial field to provide the required collision energy. As long as this offset remains negative relative to the C-trap, all fragments remain trapped inside the HCD cell. The fragments can be then transferred back into the C-trap, ejected into the Orbitrap analyzer and analyzed in a single Orbitrap detection cycle.

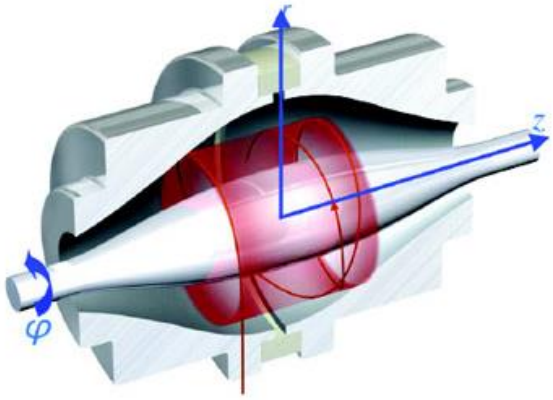


Figure 2.6: Orbitrap analyzer[46].

The mass range covered by the instrument is *m/z* 50–4000. Acquisition speed ranges from 12 Hz for resolving power 17,500 at *m/z* 200 to 1.5 Hz for resolving power 140,000 at *m/z* 200[47].

The quadrupole and the Orbitrap are two independent MS detectors. The quadrupole has the function of isolating the ions and sending them to the Orbitrap to be analyzed. However,

after the isolation process, the ions can be either ejected and analyzed from the Orbitrap, or subjected to fragmentation to yield product ions and then ejected and analyzed from Orbitrap again.

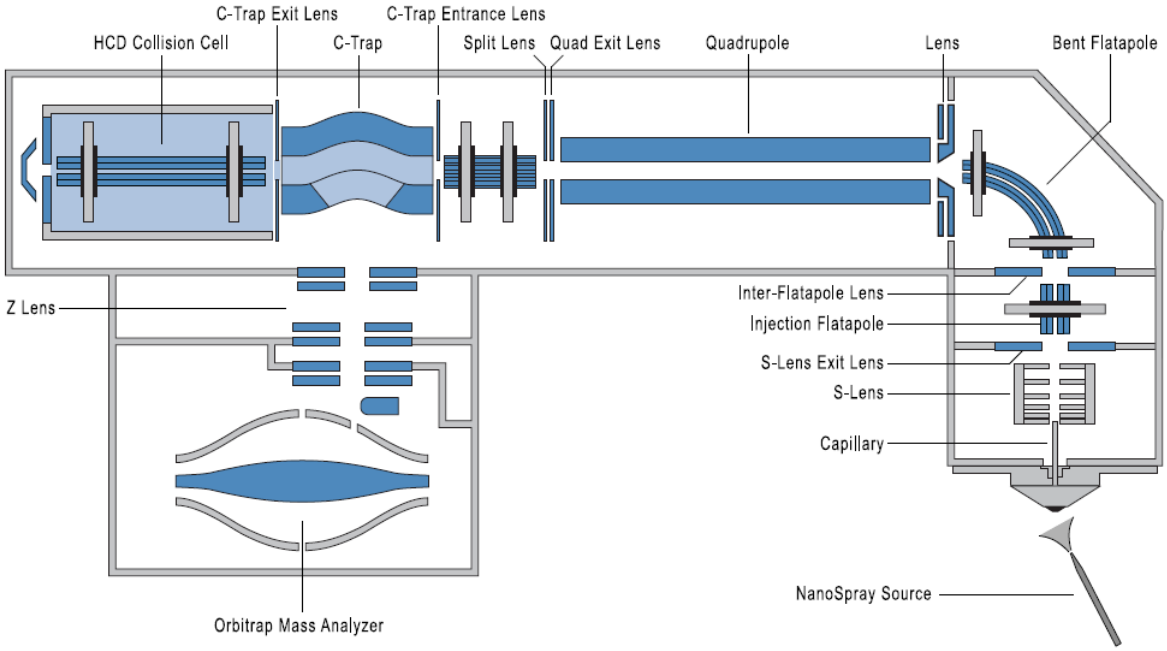


Figure 2.7: Diagram of a Q-Exactive mass spectrometer [47].

Due to the characteristic of this hybrid mass spectrometer several acquisition mode are available such as Full Scan, SIM, Product ion, Neutral Loss etc. However, two acquisition modes in particular make the difference compared to other instruments:

- data dependent acquisition (DDA) is the most commonly used approach. In a first scan the instrument is operated in Full Scan mode to detect and analyze all the ions coming from the source. In the next scan, the most abundant ions (it is possible to choose from 5 to 20 ions usually referred to as TOPN fragmentation) are then selected from the quadrupole, and ejected to be fragmented in the collision cell. After the fragmentation the ions are analyzed by the Orbitrap. In order to avoid resampling of earlier selected precursors, a dynamic exclusion list is used. Additionally, it is possible to use an inclusion list, which forces the instrument to select a desired peak for fragmentation, regardless of its abundance, to ensure the identification.

-data independent acquisition (DIA) or all ion fragmentation (AIF) has been introduced in the last years. In this case no precursor selection is performed, rather all present precursors of a distinct mass window are co-fragmented.

In the case of targeted approaches, in which a specific class of compounds have to be analyzed, it is also possible to work in MRM mode. The major limitation of MRM is that it is important to validate targeted masses and its transitions, prior to the experiment; while DDA has the limit of a stochastic and irreproducible selection of precursors for fragmentation, compromising the reproducibility (mainly of less abundant ions). This is the reason why the use of DIA is continuously increasing.

However, in both DDA and DIA, the ability to fill the HCD cell or the C-trap with ions while a previous Orbitrap detection cycle is still ongoing is an important innovation that allows to significantly increase the acquisition speed and quality of spectra. The advantage of this configuration is the possibility of working in parallel with precursor master scans in the Orbitrap and fast acquisition of fragmentation spectra on selected ions in the HCD cell. Moreover, fragment ions are analyzed in the Orbitrap, allowing their detection with a good mass accuracy, a quite important feature because errors in the determination of fragments *m*/*z* values affects the final identifications.

Finally, two fragmentation techniques can be used in Q-Exactive: Collision-induced dissociation (CID) and HCD. CID and HCD use an inert gas for fragmentation: the energy acquired during the collisions of the intact molecule with the gas molecules (for example He) leads to the breakage of chemical bonds. Depending on the fragmentation technique, different fragments are created.

3.2.3 Time-of-flight (TOF)

The TOF analyzer separates ions, after their initial acceleration by an electric field, according to their velocities when they drift in a free-field region that is called a flight tube. Ions are expelled from the source and then accelerated towards the flight tube by a difference of potential applied between an electrode and the extraction grid. As all the ions acquire the same kinetic energy, ions characterized by a distribution of their masses present a distribution of their velocities. When leaving the acceleration region, they enter into a field-free region where they are separated according to their velocities, before reaching the detector positioned at the other extremity of the flight tube. The *m/z* ratios are determined by measuring the time that ions take to move through a field-free region between the source and

the detector. Considering *V* is the potential applied to accelerate the ion with *m/z*, the time *t* needed to cover the distance *L* before reaching the detector is given by

$$t^2 = \frac{m}{z}\left(\frac{L^2}{2eV}\right)$$

In principle, the upper mass range of a TOF instrument has no limit, which makes it especially suitable for soft ionization techniques. Another advantage of these instruments is their high transmission efficiency that leads to very high sensitivity. However, the most important drawback of the first TOF analyzers was their poor mass resolution. Mass resolution is affected by factors that create a distribution in flight times among ions with the same *m/z* ratio, such as the variation of the initial kinetic energy of the ions (kinetic energy distribution). A way to improve mass resolution is to use an electrostatic reflector also called reflectron. The reflectron corrects the kinetic energy dispersion of the ions with same *m/z* leaving the source. Ions with more kinetic energy and hence more velocity will penetrate the reflectron more deeply than ions with lower kinetic energy. Consequently, the faster ions will spend more time in the reflectron and will reach the detector at the same time than slower ions with the same *m/z*. The reflectron performance may be improved by using a two-stage reflectron that enables the beam to traverse the flight tube twice (Figure 2.8). This doubling of the flight path enables the increase of the resolution [41].

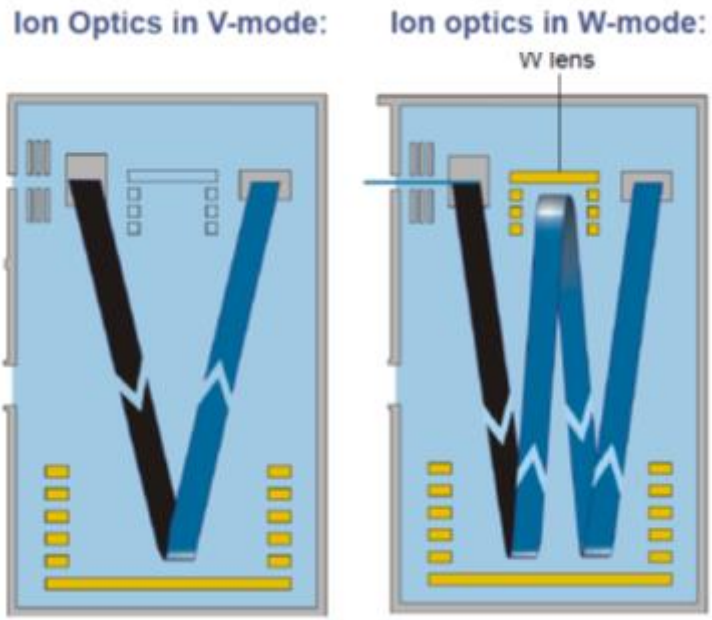


Figure 2.8: Schematic diagram of V-mode and W-mode ion optics.

To allow the coupling of TOF analyzer with continuous ionization techniques, such as ESI or API, the orthogonal acceleration (oa) technique has been developed. The sample is continuously ionized in the ion source. Ion optics focus the resulting ions into a parallel beam and direct it to the orthogonal accelerator. Applying a potential to the accelerator the bean will be directed toward the flight tube and then set again to 0. During the time that the ions continue their flight, the orthogonal accelerator is refilled with new ion beam. Another flight cycle will begin by reapplying a pulse voltage to the plate[41].

TOF analyzers are nowadays always hybrid mass spectrometers. The combination of different analyzers allows several possibilities. In the work described in this thesis a qTOF premier orthogonal mass spectrometer has been employed. The qTOF Premier is a hybrid quadrupole-orthogonal acceleration TOF mass spectrometer (Figure 2.9). It is provided with a reflectron to reflect ions back again towards the detector (V-optics) allowing a resolution of 10 000 FWHM (full width at half maximum), but is also equipped with a second reflectron that enables the beam to traverse the flight tube twice (W-Optics) enabling a resolution of 17,500 FWHM.

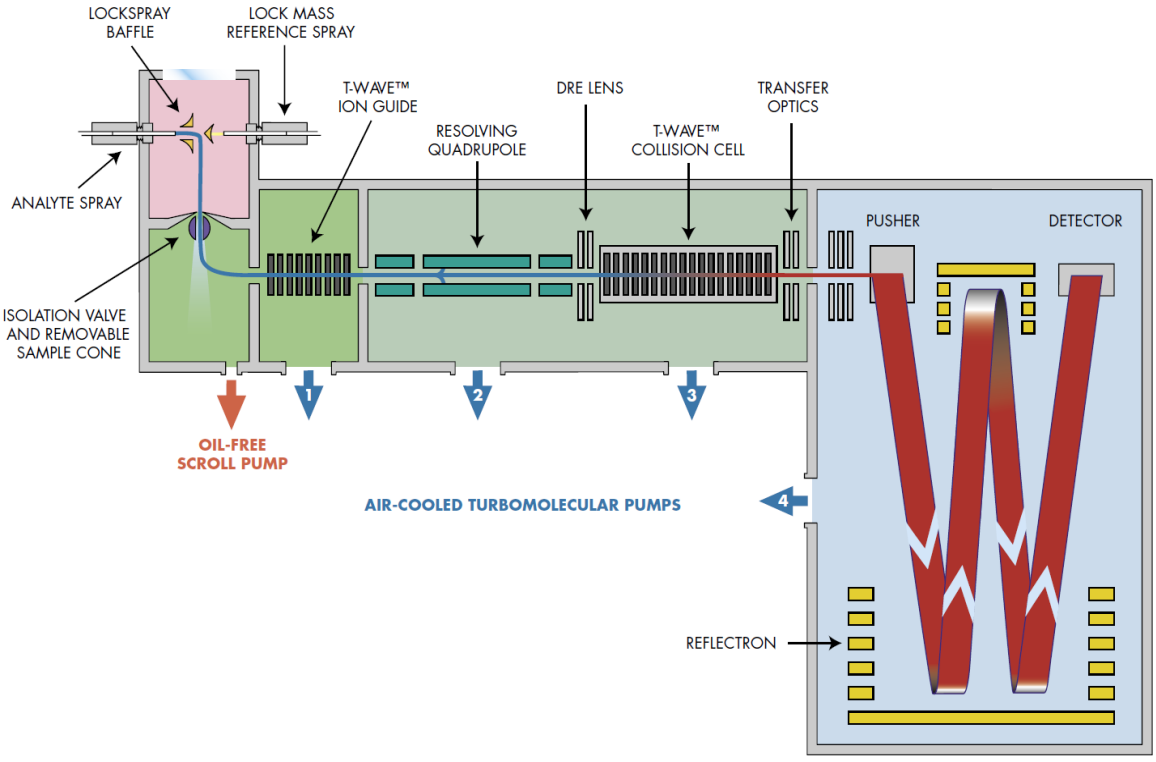


Figure 2.9: Schematic diagram of a Premier orthogonal QTOF

The hybrid configuration of qTOF premier enables automated exact mass measurement of both precursor and fragment ions. Indeed, qTOF technology provides both quadrupole and TOF mass analyzers, with an intermediate collision cell for fragmentation, if required. This powerful combination allows ions to be selected from the quadrupole, individually fragmented in the collision cell, and then measured to a high degree of mass accuracy by the TOF. The quadrupole can be operated in several modes:

-Full scan mode, in which the ions in a wide range of *m/z* pass through the quadrupole and are accurately measured by the TOF.

-The quadrupole can either be parked on one specific mass (TOF MS/MS) or can be made to scan through a wide mass range in search of candidate ions for fragmentation (parent ion scanning).

-DDA mode when the instrument automatically switching between full scan and fragmentation of the most abundant ions detected during the full scan acquisition.

3.2.4 Resolving power

Resolution or resolving power is the ability of a mass analyzer to yield distinct signals for two ions with a small *m/z* difference. The exact definition of these terms is one of the more confusing subjects of MS terminology that continues to be debated. Two peaks are considered to be resolved if the valley between them is equal to 50% of the weaker peak intensity when using quadrupoles, IT, TOF, and so on. If Δm is the smallest mass difference for which two peaks with masses m and $m+\Delta m$ are resolved, the definition of the resolving power R is $R=m/\Delta m$. Therefore, a greater resolving power corresponds to the increased ability to distinguish ions with a smaller mass difference (Figure 2.10 a). The resolving power can also be determined with an isolated peak. Based on the definition of Marshall [41], indeed, the resolving power is determined using the peak width Δm at 50% of the peak height, FWHM. Therefore, the narrower the peak width, the higher the resolution[41] (Figure 2.10 b).

Mass accuracy indicates the accuracy of the *m/z* provided by the mass analyzer. It is the difference that is observed between the theoretical *m/z* (*m* theoretical) and the measured *m/z* (*m* measured). It can be expressed in millimass units (mmu) but is often expressed in parts per million (ppm). Mass accuracy is largely linked to the stability and the resolution of the

analyzer. A low-resolution instrument cannot provide high accuracy. Indeed, the precision obtained on the mass of the analyzed sample depends also on the determination of the centroid of the peak[41].

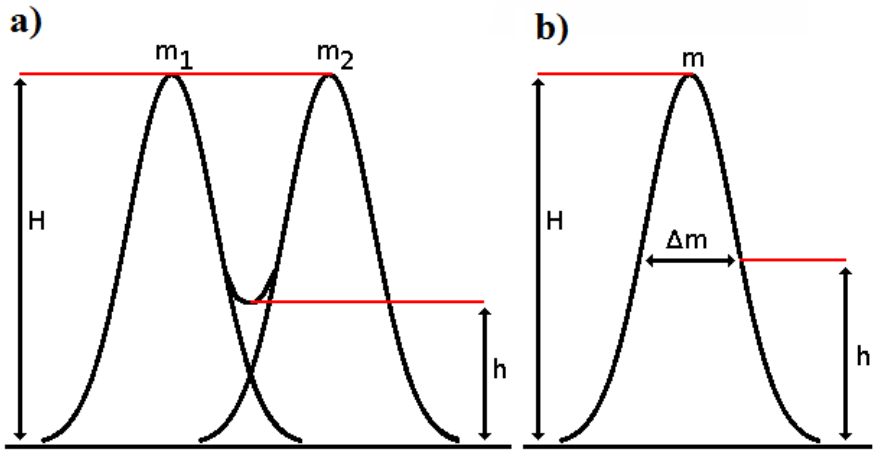


Figure 2.10: Resolution calculation based on the Marshall definition (a) and on the FWHM definition (b).

Low resolution (LR) or high resolution (HR) is usually used to describe analyzers with a resolving power that is less or greater than about 10 000 (FWHM), respectively. However, there is no exact definition of the boundary between these two terms. LR mass spectrometers, such as QqQs and quadrupole, operate at unit resolution that is a resolution that is sufficient to separate two peaks one mass unit apart. TOF instruments are now generally capable of 20,000 resolving power. Orbitrap instruments are capable of HR up to 140,000.

In TOF analyzers the resolution depends on the capacity of the instrument of eliminating the initial spread of kinetic energy of the injected ions, this makes resolution independent from detection time (T_d) or *m*/*z*. In Orbitrap analyzers resolution depends on the ratio of the detection time (T_d) to the period of main oscillations (T_ω). Because the ion motion is determined by the electrostatic field, T_ω is proportional (and resolution inversely proportional) to the square root of *m*/*z*. Finally, TOF resolution is largely independent from *m*/*z* also in the MS/MS mode, whereas in the Orbitrap analyzer the resolution in MS/MS mode is often sacrificed to increase scan speed[2]. However, the highest resolution available in commercial TOF devices is several times lower than the highest resolution in Orbitrap. The FTICR and Orbitrap analyzers outweigh any other mass analyzer with respect to the achievable maximum mass resolution and accuracy.

4 Low-resolution versus high-resolution mass spectrometry

The development of chromatography coupled to MS allowed several progresses for the identification and quantification of both endogenous compounds and contaminants in complex matrices, such as food. Several methods have been developed for the analysis of several classes of compounds. Several advantages and disadvantages has been shown depending on the use of HR or LR mass spectrometers for the identification and quantification of analytes.

4.1 Identification

The methods used for the analytes identification change considerably when operating with LR mass spectrometers compared to HR mass spectrometers.

LR mass spectrometers allow to determine the nominal mass of compounds. Therefore, to discriminate between compounds with isobaric masses is necessary to perform MS/MS. For a large range of compounds, SRM of precursor-product ion transitions by using QqQ allows to identify compounds at low concentrations. A range of studies have shown, however, that monitoring only one transition might result in false positive identifications for individual compounds and thus at least two transitions are required[48]. However, several limitations become evident: (i) Under the constraints of at least two transitions for the identification, SRM methods are typically limited to about 100–150 target analytes depending on chromatographic separation, as otherwise accuracy or sensitivity deteriorate due to an insufficient temporal peak resolution or too short acquisition times for the individual MS/MS transitions, respectively. (ii) For some analytes, only non-specific transitions such as the neutral loss of H_2O or CO_2 might occur, which are common also for matrix interferences. (iii) Some analyte ions, particularly those of low molecular weight, show only one transition. These limitations, together with the impossibility of LR mass spectrometers to determine the accurate mass of analytes, do not allow the certain identification of unknown compounds. For instance, two isomeric compounds, showing same nominal mass and same transitions, cannot be distinguished and identified. In these cases, it is necessary to compare the retention

time (RT) and MS/MS spectra of the analytes with those of a reference standard for the unambiguous identification[49].

HRMS analysis offers promising solutions to these limitations. All compounds present in a sample can be determined simultaneously with HRMS instruments operating in full-scan mode, and fragmented by means of acquisition modes such as DDA, making no preselection of compounds and associated SRM transitions necessary. Additionally, HR mass spectrometers allow the determination of accurate mass, and consequently, ions with slightly different *m/z* ratios can be distinguished from one another. The most important feature of HRMS is the capacity to determine analytes molecular formula from accurate mass measurements. The accurate mass together with the fragmentation obtained by means of MS/MS experiments allow the certain identification of unknown compounds. In theory, the presence of an unlimited number of compounds can be investigated at proper sensitivity, without requiring the preselection of analytes or even without having reference standards available. Moreover, the possibility to store a high volume of full-scan and MS/MS data of high mass accuracy, allows to carry out also retrospective analysis without the necessity to re-running samples.

Due to the high quality information by combining sensitive full-spectrum data and high mass accuracy, HRMS is a promising technique that has opened new horizons in screening and rapid identification of a wide range of compounds and unknowns identification. Nowadays, three main approaches (see Figure 2.11) for the screening and identification of substances can be pointed out[49]:

-Target screening involves the use of a reference standard for identification and quantification of a compound or a class of compounds. RT, MS and MS/MS spectra of the analyte have to be compared with those of the reference standard, measured under the same analytical conditions. This approach imply the previous knowledge of the compounds possibly present in the matrix of interest, and it requires the availability of the relative standard that, in the case of some contaminants or natural compounds, can be difficult.

-Suspect screening is performed when prior information from various sources indicates that a given structure may be present in the sample. But, in contrast to target analysis, the suspects screening approach does not rely on reference standards for identification. Indeed, the exact mass and isotopic pattern can be used to screen for the investigated substances in the sample.

-Non-target screening involves the study of all compounds detected in a sample where no prior information on the composition is available. As no structural information is available

in advance, a full non-target identification starting from the exact mass, isotopic pattern, adducts and fragmentation information needs to be performed.

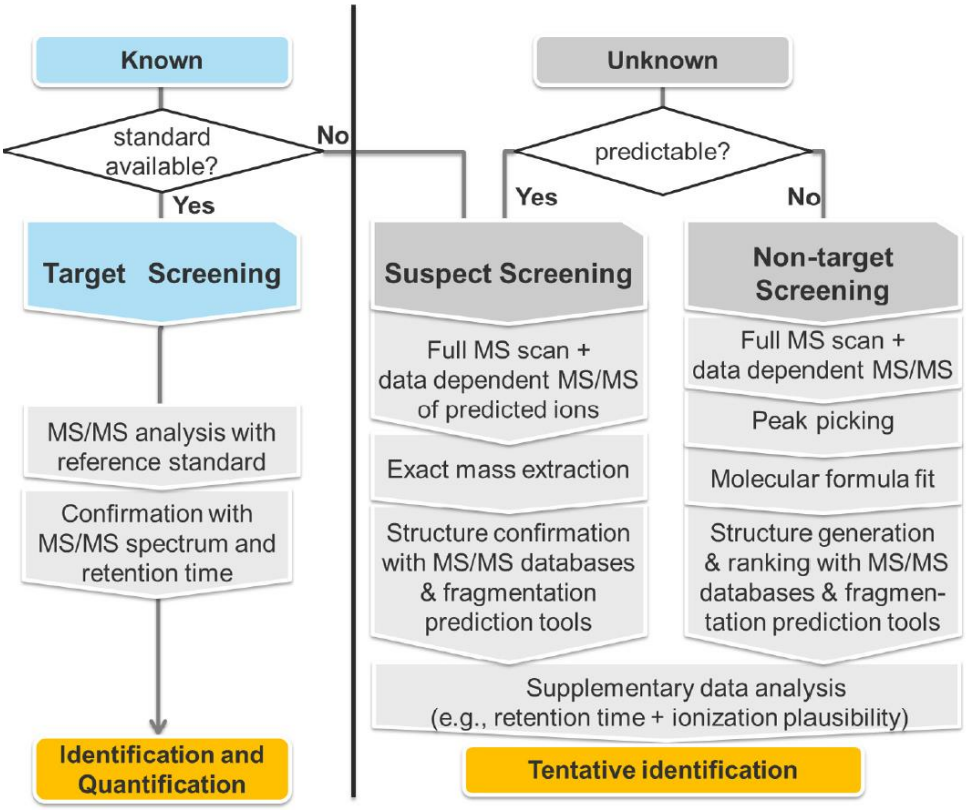


Figure 2.11. Target, suspect and non-target analysis approaches

4.2 Quantification

As previously mentioned, the identification and quantification of compounds at low concentrations requires both a high sensitivity and selectivity against complex matrix backgrounds. For a large range of compounds, SRM of precursor-product ion transitions by using QqQ fulfils these prerequisites. However, as explained before, under the constraints of at least two transitions for the quantification, SRM methods are limited to about 100–150 target analytes, as otherwise sensitivity deteriorate. When the number of target analytes is limited the problem can be overcome. In this case, indeed, QqQ analyzer in its MRM mode results the most adequate one for quantitative methods. However, although LR mass

spectrometers such as QqQ result to be very robust and allow a reliable quantification of known compounds, they do not allow the certain identification of unknown compounds.

HRMS target analysis offers promising solutions to the limitations of SRM analysis. All compounds present in a sample can be determined simultaneously operating in full-scan mode, making no preselection of compounds and associated SRM transitions necessary. However, on qTOF instruments, these capabilities are impaired by the limited sensitivity, which is about 1–2 orders of magnitude lower than those of QqQ instruments in SRM mode, and the limited dynamic range which is about 10-fold below that of QqQs. Thus, qTOF instruments have been used only occasionally for quantification and one established strategy is to use QqQ for quantification and a separate qTOF analytical run for confirmation. The Orbitrap instrument offers a better dynamic range and a sensitivity close to that of many QqQ instruments, thus allowing for quantification and confirmation in a single analytical run[49].

However, several publications describing the new generation of HRMS systems and highlighting differences between LR MS and HR capabilities[50, 51] in target quantification are available. The biggest difference between low and HR mass spectrometers, although, is the relative quantification of thousands of compounds in untargeted analysis. The untargeted analysis deals with detection and identification of the greatest possible number of compounds present in the sample, usually with the purpose of a (semi)quantitative comparison of different groups of samples[52].

The current trend in untargeted workflows is the analysis of large sets of samples to obtain representative information from the biological system under investigation. Thus, most studies require long periods to analyze samples in which the quantitative response can fluctuate by alteration of the instrument performance owing to accumulation of matrix components in different instrumental zones or simply by periodic practices such as instrument calibration or cleaning protocols. These sources of instrumental variability are generally corrected in targeted analysis by using isotopically labelled internal standards. However, this is not viable in untargeted analysis owing to the wide chemical heterogeneity of metabolites. The most common practice to monitor experimental variability (including instrumental variability) in untargeted analysis is the implementation of quality control samples (QCs), periodically inserted in the sequence of analyzes programmed for each batch of samples. A cut-off value in terms of variability is set to filter molecular entities or metabolites in the final data set. Other tools classify the metabolites according to their

behaviour along the sequence of analyzes and use correction functions for each group to correct instrumental variability.

4.3 Data analysis

In contrast to targeted approach and suspected approach, in which a small range of known molecules have to be analyzed, in a typical large-scale untargeted experiment a large number of unknowns have to be analyzed at the same time. The main difference between these approaches, which is basically dealing with known or unknown molecules, leads to differences in the way data analysis is accomplished. Whereas in target or suspected analysis the presence or absence of each substance is determined individually using the extracted ion chromatogram (XIC), in non-target methods a screening of detected compounds can be performed by means of dedicated software for the data pre-processing and statistical analysis. Several approaches are discussed in the literature [53, 54] reporting the following key steps: (i) an automated peak detection by exact mass filtering from the chromatographic run by means of several pre-processing tools (MZmine, XCMS etc); (ii) an assignment of an elemental formula to the exact mass of interest by means of pre-processing tools or online available tools (PredRet etc.); and (iii) a database search of plausible structures for the determined elemental formula (available databases such as Chemspider, Metlin, human metabolome database etc.); (iiii) the investigation of MS/MS spectra to elucidate the structure and comparison with available databases (such as Mass Bank, or MZCloud). Sometimes, also *in silico* fragmentation prediction tools can be used, such as Met Frag or Mass Frontier[55]. These algorithms match the experimental precursor masses and their MS/MS spectra to those generated by *in-silico fragmentation;* the statistical analysis of the comparison between the experimental spectrum and the theoretical one provides an ion score for the match, which is calculated on the basis of the percentage of fragmentation matching.

However, the identification of a compound requires several information to be investigated including the isotopic pattern, presence of additional adducts, RT, fragmentation information. Several papers discuss this in greater detail [56, 57] and introduce the concept of “identification levels” in HRMS analysis. The concepts of identification strategy and confidence are merged in Figure 2.12, showing that target, suspect and non-target compounds start by definition at Levels 1 (reference standard available), 3 (tentative

candidate(s)) and 5 (no information), respectively. If sufficient MS (exact mass, isotope, adduct), MS/MS (i.e., fragmentation) and experimental information (e.g. retention behaviour, presence of related substances) is available, suspect and non-target components can gain in confidence through to Level 2 (library match and/or diagnostic fragments) and even Level 1 following purchase of the corresponding standard for identifications (green arrows in Figure 2.12).

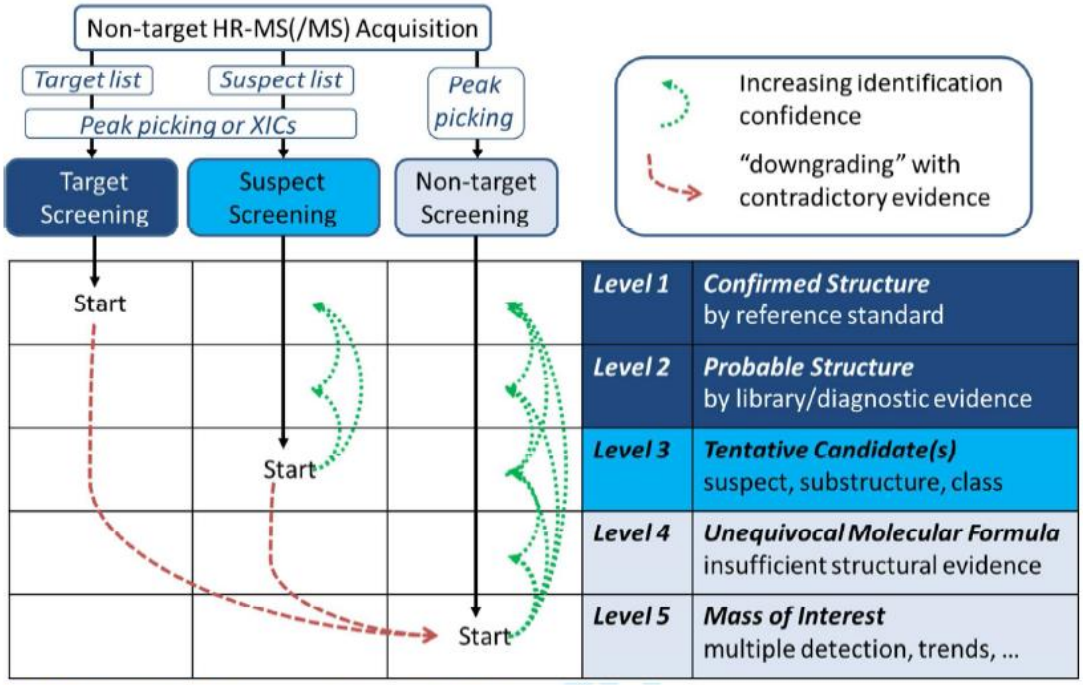


Figure 2.12: Scheme of the 5 identification level related to target, suspect and non—target screening[54].

It is important to note that this elaborate matrix of confidence is necessary with soft ionization HRMS/MS analysis at this stage for one main reason: the lack of comprehensive spectral libraries for soft ionization techniques due to the lack of reproducibility between soft ionization instruments and various settings and the relative newness of the technique [58]. As a consequence, the identification of "unknown" substances measured with soft ionization techniques are quite complicated. This is the reason why, before starting with the identification of unknowns, it is really important the selection of the candidates of interest for the topic under investigation.

A number of strategies for the selection of candidates for suspected and non-target screening using HR-MS/MS information have been developed. Suspect screening has been

performed using predicted transformation products [59] or metabolites of interest. For non-target identification, a differential analysis is usually performed between several groups of samples in order to select the most significantly different compound to identify[60]. Different chemometric tools can be used to statistically assess possible differences among samples. Multivariate analysis is often used, although the particular statistical approach to be used will largely depend on the objectives of the study.

5 Metabolomics for food safety, quality and nutrition

The always more-demanding fields of food safety, quality and nutrition are continuously fostering the development of robust, efficient, sensitive and cost-effective analytical methodologies. MS-based metabolomics is a key tool nowadays with great potential in many analytical fields and has been demonstrated to be capable of facing some important challenges related to the food science domain[61].

Metabolomics, together with genomics, transcriptomics and proteomics, is involved in the study of the food and nutrition domains through Foodomics approaches. As per definition, metabolomics includes the exhaustive study of the whole small metabolite composition (molecular weight below 1500 Da) of a particular system. In practice, this aim is difficult to achieve, because it implies the development of a universal approach to analyze metabolites belonging to very different chemical classes and present in a very wide dynamic range. In this regard, the food metabolome is not an exception as quite diverse compounds, such as carbohydrates, lipids, proteins, amino acids, amines, steroids, phenolic compounds, carotenoids, alkaloids, among others are frequently present.

There are three basic approaches that can be used in metabolomics: target analysis, metabolic profiling, and metabolic fingerprinting. Target analysis aims the quantitative measurement of selected analytes, such as specific contaminants, reaction products, bioactive compounds or biomarkers by using authentic standards. The resulting data can be used to distinguish different sample classes, however, most of the metabolite information of complex samples is lost. Metabolic profiling is a suspected screening that focuses on the study of a group of metabolites related to a specific metabolic pathway. Metabolic profiling refers mainly to the analysis of metabolites often belonging to the same chemical class, which are frequently identified and quantified [62, 63]. Finally, metabolic fingerprinting is

referred to the analysis of as many compounds as possible within a system, including their detection and the subsequent statistical treatment of the obtained results. Under this approach, the identification and quantification of the detected metabolites may not be a necessity. Fingerprinting aim is not to identify all metabolites, but to compare the patterns of metabolites that change in response to external factors.

In any case, as the complexity of the set of metabolites to be analyzed is quite high, suitable analytical techniques are needed. Firstly, a proper sample treatment methodology is required. Food are usually quite complex matrices full of potentially disturbing components for the analysis of metabolites. It is important to extract the analytes of interest but, as a universal sample treatment directed to the extraction of the full metabolome does not exist, some components may be lost during this phase. Secondly, a proper separation technique is required prior to the detection. The technique of choice in metabolomics is mostly LC that can be operated in several separation modes, increasing its versatility towards different metabolites. Particularly, in the last years, methods based on the use of UHPLC have gained considerable popularity thanks to the advantages that this technique can provide. Finally, a good detection method is essential for the coverage of the whole metabolome. In the last years MS has gradually substituted the use of NMR because of the possibility for coupling with a separation technique, as well as the development and improved affordability of HRMS instruments. HR instruments allowed to significantly enhance the capabilities for the identification of unknown metabolites and thus developing the metabolomics field. As a direct consequence of the improvement of the available analytical tools, samples with higher complexity can be analyzed. The datasets generated after sample analyzes in a typical metabolomics study is of extremely great complexity, requiring the development of dedicated bioinformatics tools. However, the particular statistical analyzes made are usually different depending also on the topic of the study, i.e., food and health relationships, biomarker discovery, food quality, food safety or traceability, among others.

5.1 Metabolomics for food safety

Food safety is one of the most-important topics within food analysis. In this context, the use of MS within metabolomics-based approaches has allowed significantly raising the level of the analytical determinations possible nowadays in this field[61].

The first part of any MS-based metabolomics study for the detection of food contaminants is sample preparation. As foods may be considered as very complex matrices, suitable sample preparation steps are needed in order to allow a proper detection of contaminants which will surely be present in very low amounts. Different methods such as solid-liquid extraction (SLE) or liquid-liquid extraction (LLE), solid-phase extraction (SPE) have been widely used to extract and/or concentrate the analytes of interest. However, following the latest trends regarding the application of "Green Chemistry" principles, other miniaturized protocols limiting the volumes of solvents employed have been also proposed and employed in the last years such as solid-phase microextraction (SPME) [64] and QuEChERS (Quick, Easy, Cheap, Effective, Rugged, and Safe) methods[65].

Methods directed to quantification of chemical contaminants in food are strongly influenced by current international legislation, which is generally directed to the establishment of maximum residue levels (MRLs) on certain substances, and to specify the banned compounds that cannot be present at any concentration. MRLs for pesticides, veterinary drugs and contaminants are available[9]. The most frequent analytical approach to determine contaminants in foods relies on the use of QqQ analyzers under SRM. Two transition are monitored: the most-intense one is used for quantification whereas the second is employed for qualification purposes. This detection procedure allows complying with European legislation on banned and controlled substances in foods. This regulation establishes the requirements that an analytical method must meet for an unequivocal identification and quantification of a controlled substance in a food sample, which means at least four identification points: one identification point is gained by RT confirmation with a commercial standard, whereas additional 1.5 identification points are gained for each ion transition successfully confirmed.

In the last years, the use of QqQs in MRM mode is still the most-extended approach, even though the use of HRMS in the field of food safety is increasing. In fact, the use of the above described targeted approach has important limitations, which are mainly related to the determination of unknown compounds as well as the need of reference commercial standards. For this reason, the use of non-targeted analysis of contaminants is increasingly proposed [66]. HRMS also help to discover and identify new or unexpected chemical contaminants. This technique is clearly more capable of discovering new safety hazards beyond the use of the regulated compounds and contaminants. Every year new compounds are individuated as potential chemical pollutants and in conclusion the number of illegal

compounds is in continuous growth. However, the use of these diverse non-targeted methodologies is still somewhat limited compared to the targeted approach. This is the case, for example, of natural toxins. For the detection and quantification of toxins in foods, similar approaches to those already described for chemical contaminants are widely employed. However, in this case, the natural toxin variability potentially present in a particular food product mean that less compounds have to be analyzed, and thus, advanced metabolomics-based approaches are not required.

5.2 Metabolomics for food quality

Nowadays, food quality is one of the major concerns of the food industry. Its evaluation is a complex task due to the multiple aspects that may be considered to achieve an appropriate food quality. Food composition, aroma, flavour, or nutritional properties are among the most important aspects that may be evaluated in food quality assessments. MS based metabolomics approaches are gaining attention due to their demonstrated capability to establish links between the chemical composition and food quality, to control food authentication and adulteration, or to differentiate food samples according to their variety[61]. The ultimate goal of these researches is to determine relevant compounds that may be selected as quality markers. However, one of the main strategies to evaluate food quality is the investigation of the relationship between composition and bioactivity. Diet is a source of bioactive compounds able to contribute to the improvement of human health. Consequently, the identification of such compounds is crucial to provide customers with a healthy, balanced diet. Bioactive compounds are being studied in the prevention of cancer, heart disease, and other diseases. In this context, the vegetable kingdom is a primary source of such phytochemicals with potential health benefits. At present, researchers have identified hundreds of compounds with health-promoting, disease-preventing, or curative properties, and new discoveries concerning the complex interactions between secondary metabolites in plants and health are continually made [67].

The advent of metabolomics, which is a powerful approach enabling the comprehensive, qualitative and quantitative analysis of all metabolites [68], prompted research in this field; HRMS coupled to UHPLC is increasingly used in metabolomics. The recent innovations in instrumentation and informatics currently allow performing a comprehensive analysis of

metabolites. However, rigorous methodologies are required. The development of powerful analytical techniques is essential for the advancement in bioactive compounds research. For more details on the modern analytical approaches, based on HPLC and UHPLC coupled with HRMS and MS/MS, employed to characterize and/or identify bioactives in food a review, specifically on phytochemicals analysis by means of HRMS, has been reported in this thesis (Appendix). In the review all the relevant steps to assess food quality by MS-based metabolomics, such as proper sample preparation procedures separation techniques and mass spectrometric approaches, will be discussed.

5.3 Metabolomics for nutrition

In the last years, the interest in establishing clear relationships between diet and health is continuously growing. However, it is interesting to remark that several of the health benefits assigned to many dietary constituents are still under controversy as can be deduced from the large number of applications rejected by the EFSA about health claims of foods and or new ingredients[33, 34]. More sound scientific evidences are needed to demonstrate the claimed beneficial effects of these foods and constituents.

Understanding the roles of nutritional compounds at molecular level (i.e., their interaction with genes, proteins and metabolites) is the main challenge that food scientists and nutritionists have to face to adequately answer the new emerging questions. Nutrients can be considered as signaling molecules that are recognized by specific cellular-sensing mechanisms. However, unlike pharmaceuticals, the simultaneous presence of a variety of nutrients with diverse chemical structures and concentrations and having numerous targets with different affinities and specificities increases enormously the complexity of the problem. Therefore, it is necessary to look at hundreds of test compounds simultaneously and observe the diverse temporal and spatial responses.

In this sense, new strategies seem to be essential to understand how the bioactive compounds from diet interact at molecular and cellular levels, and their crucial influence on the transcriptome, proteome and the metabolome. The combination of the information from the three expression levels (gene, protein, and metabolite) can be crucial to adequately understand and scientifically sustain the health benefits from food ingredients[1].

Metabolomics is nowadays considered an adequate strategy to investigate the complex issues related to prevention of future diseases and health promotion through food intake. It can be a major tool for detecting small changes induced by food ingredient(s) at the metabolic level, making possible new investigations on food bioactivity and its effect on human health at molecular level. The analysis of metabolic patterns and the changes in the metabolism in the nutrition field can be, therefore, very interesting to locate; for example, variations in different metabolic pathways due to the consumption of different compounds in the diet. The recent years have therefore seen an increased application of metabolomics to monitor specific food intake and dietary patterns in humans and to investigate food-related diseases via the discovery of dietary biomarkers[69].

5.3.1 Dietary intake of foods

Generally, food frequency questionnaires or multiple days of diet records have been used as the routine methods for collecting dietary intake data to evaluate positive or negative effects of certain foods or diets on large population cohorts [70]. However, these methods have many limitations, such as high cost and time consumption, recall bias and measurement error, the misreporting or underreporting of dietary intake, difficulty in determining accurate portion sizes and inappropriateness for some populations with cognitive impairment [71]. To overcome such limitations of these conventional data-collecting methods, metabolomics has been used as an alternative for evaluating the positive or negative consequences of certain foods or diets and for analyzing the dietary patterns by identifying new biomarkers of food intake. Many studies have identified novel biomarkers associated with the intake of specific foods such as juice, fruits, vegetables, grain, fish, wine and coffee and complex intakes following dietary patterns [72–74]. These results have shown that specific foods or dietary patterns significantly affect various metabolisms such as amino acid and lipid metabolism connected with human health. These studies have demonstrated that metabolomics is an effective and useful tool to reveal metabolic changes after specific food consumption and to analyze dietary patterns for nutritional epidemiological studies.

5.3.2 Food and health relationship

Epidemiological studies have reported that diet is related to certain diseases. To obtain more persuasive and reliable results about diet-related diseases, metabolomics has been used to reveal metabolic alterations associated with diet-related diseases and the results of diet intervention. The relationship between green tea consumption and obesity prevention has been elucidated by metabolomic analysis[75]. Additionally, metabolomics analyzes of the serum metabolites produced after coffee consumption have revealed that caffeine-related metabolites are inversely associated with colorectal cancer[76]. The metabolomic study of dietary curcumin, found that it partially recovered metabolic disorders of the glycolysis and fatty acid metabolism in hyperlipidemia[77]. Furthermore, metabolomics has been used to identify biomarkers of human diseases and to understand metabolisms of human diseases[78, 79]. Recently, the human gut microbiota has received increased attention, since microorganisms in the human intestinal along with the metabolites they generated were found to play an important role in human health and diseases. The composition of the human gut microbiota is affected by food intake and its manipulation by food intake can be a potential therapy of certain diseases. Therefore, it is necessary to understand the metabolism and functional status of the human gut microbiota and its interaction with the host, factors that metabolomics can be a powerful tool for characterizing[80].

6 Aim of the thesis:

In the previous paragraphs we focused the attention on several topics concerning metabolomics for food analysis. Firstly, the importance of the assessment of food safety, food quality and nutrition. Secondly, the importance of MS and the overall approaches available to face the food analysis issue such as targeted, suspected and untargeted approaches. All the methods have been studied in details to understand the difficulties, the advantages and disadvantages of the mentioned approaches. Several steps have been proved to be necessary for the development of a reliable analysis of food such as sample preparation, chromatographic and mass spectrometric method, data analysis and identification.

In this thesis, the potentialities of MS have been exploited from several points of view in order to give an overview of its application on the field of food metabolomics. All the aspects concerning the development of a reliable method for food analysis have been treated and critically discussed.

The thesis is divided in three main sections: targeted approach for food safety, suspected approach for food quality, untargeted approach for nutrition. In each section, one of the three main approaches (targeted, suspected, untargeted) used in food metabolomics has been studied for its application on one of the three different food analysis fields (safety, quality, nutrition). Moreover, each topic has been treated focusing the attention on a specific step of the method development (sample preparation, LC-MS, data analysis) (Figure 2.13).

The topic of metabolomics for food safety has been studied in a targeted approach by developing a method for the analysis of secondary metabolites of fungi, namely mycotoxins, in food. In Paper I, the development of a particular material for the clean-up of mycotoxins from milk extracts is presented. In Paper II, an application of the material developed in paper I has been shown for the analysis of mycotoxins on a different matrix, i.e. cereals. In the two works, a typical workflow for the targeted identification and quantification of analytes in food by means of MS is presented. However, particular emphasis has been given to the importance of a proper sample preparation for the development of an efficient method. The topic of metabolomics for food quality has been treated in a suspected approach by developing a method for the analysis of several classes of secondary metabolites of plants such as phenolic acids, flavonoids, PAs, ETs and GLSs in food. In Paper III, is presented the development of a chromatographic and mass spectrometric method for the profiling of strawberry. In Paper IV, is presented a columns evaluation for the profiling of strawberry.

Paper V is focused on a chromatographic method evaluation for the profiling of GLSs in cauliflower. The three works present the typical workflow of a suspected analysis for the metabolic profiling of specific classes of compounds in food. However, the main theme of the papers is showing how a chromatographic and mass spectrometric method development is essential for the comprehensive profiling of a complex matrix.

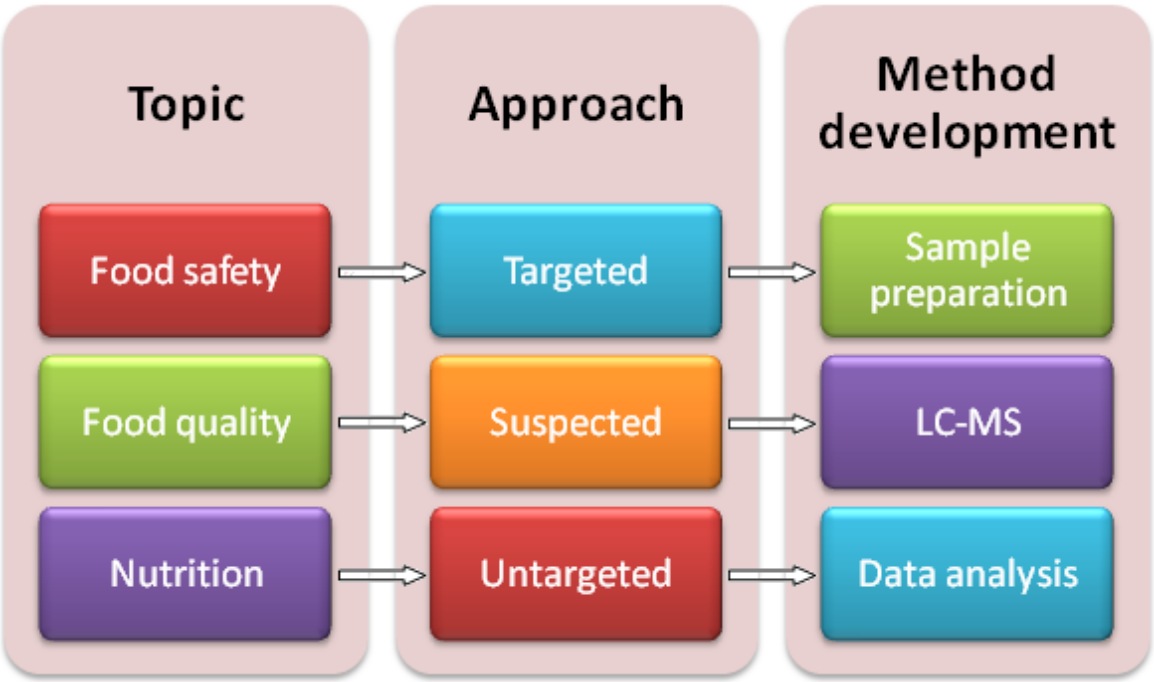


Figure 2.13: Diagram showing the thesis structure

The topic of metabolomics for nutrition has been treated in an untargeted approach by developing a method for the analysis of the human urinary metabolome after the consumption of meat and dairy products. “Paper VI”, not published yet, is focused on the identification of urinary biomarkers after the consumption of meat either dairy products. In this section, a typical untargeted workflow is presented, focusing in particular on the importance of the data analysis and identification step on the investigation of the whole metabolome.

Finally, a review on bioactive compounds in vegetables and fruits by means of HRMS is reported in the Appendix to have more detailed information on MS and the approaches used in food metabolomics.

Chapter 3:

Targeted approach for food safety

1 Mycotoxins targeted analysis in food

The first section of the thesis concerns the application of a typical targeted approach for the analysis of mycotoxins in food. Two papers, strictly related to each others, will be presented in the next paragraphs, developing a method for the analysis of different classes of mycotoxins in two different matrices.

1.1 Background

As previously reported (Chapter 1, paragraph 2.2), mycotoxins can be found on different crops, in particular cereal grains. However, human exposure might occur not only directly through the intake of contaminated agricultural products, but also indirectly through the consumption of products of animal origin obtained from animals that were fed with contaminated feed[8]. In Paper I and Paper II different matrices, i.e., milk in Paper I and corn meal and durum wheat flour in Paper II have been analyzed for the detection of several classes of mycotoxins. In particular, Zearalenone (ZEN) and its derivatives, namely α-zearalenol (α-ZEL), α-zearalenol (α-ZEL), α-zearalanol (α-ZAL), β-zearalanol (β -ZAL), and zearalanone (ZAN), that are classified as mycoestrogens[81] because of their ability to bind the estrogen receptor [82], have been analyzed in Paper I. Whereas the main four aflatoxins (AFB1, AFB2, AFG1, AFG2), ochratoxin A (OTA), and ZEN have been analyzed in Paper II.

Milk, as well as cereals, is a complex food matrix, therefore, sample pre-treatment and/or clean-up and enrichment steps are generally required to reduce matrix effects (MEs), i.e., ion suppression, and/or to preserve LC columns. Although analysis time and cost are increased in this case, the analytical method benefits from increased sensitivity and robustness (e.g., reducing column blockage and contamination). After a preliminary pre-treatment of the matrix, the mycotoxin extract is generally cleaned up/enriched by SPE. Recently, dispersive SPE in a magnetic mode using nanoparticles is attracting increasing scientific interest [83, 84]. The mechanisms occurring in magnetic SPE (mSPE) are analogous to those observed in classical on-column SPE, where the interactions between target molecules and adsorbent functional groups determine the efficiency of the system. In mSPE, the dispersion of the magnetic material into the solution containing the target

molecules assures a continuous and dynamic contact with the adsorbent surface, leading to a more efficient analyte retention. The separation of the magnetic material with the adsorbed analytes from the solution is then realized by applying a magnet outside the vessel, avoiding centrifugation or filtration steps. Finally, after eventual washings, analytes are eluted from the magnetic material by a proper solvent mixture. Magnetic SPE presents various operational advantages over classical SPE and, moreover, magnetic nanoparticles (MNPs) are easy to prepare, and various materials can be used in their synthesis. In both Paper I and Paper II, a magnetic graphitized carbon black (mGCB) was chosen as the adsorbent. Indeed, as graphitized carbon black (GCB) as adsorbent in SPE had already shown selectivity towards estrogenic compounds and mycoestrogens in milk [85], it was reasonable testing it in the magnetized form for mycoestrogens. Moreover, due to the promising results obtained in Paper I, in Paper II the same material was also tested for mycotoxins in cereals.

1.2 Magnetized graphitized carbon black

GCB may establish different types of interaction with the organic molecules, acting both as reversed-phase and anion-exchanger sorbent. Indeed, the planar graphitic surface allows the hydrophobic interaction with analytes by polarizable bonds and π interactions; furthermore, some polar and positively charged chemical heterogeneities (e.g., oxonium groups) on its surface allow the GCB to establish anion–exchange interactions and hydrogen bridges between protonated functional groups of the analytes and carbonylic groups of GCB or vice versa. GCB is a non-porous material whose surface area varies among the diverse commercially available products.

Because of the lack of proper functional groups on their surface, some carbonaceous materials require an oxidation step to improve the wettability and dispersion in water. GCB, even possessing some polar functionalities in its structure, was previously treated with HNO_3 in mild conditions, which led to an only 1% oxidation. This is because in preliminary experiments with non-oxidized GCB (data not shown), elution of the analytes from MNPs required a high volume of eluent, which could be reduced only after oxidation.

The graphitized carbon black, previously treated with HNO_3, was magnetized with Fe_3O_4, which was confirmed by the characterization analyzes. The GCB material was fully characterized before and after oxidation with nitric acid, as well as after magnetization, to

observe changes in morphology or in structure. The starting material was a GCB with particle size of 120–400 mesh and a surface area of 100 m^2/g. After magnetization, the available surface area and the pore volume were half as much. Termogravimetric analysis showed that the starting GCB, as well as the material after oxidation, was free of organic impurities but, after magnetization, the inorganic components showed a significant increase to 55%. The FTIR analysis showed signals belonging to OH, C=O and C–O after the oxidation and to typical Fe–O bond after magnetization. Finally, the morphological structure of the materials was monitored by TEM analysis, which showed that after magnetization spherical nanoparticles (100–300 nm) were grown on the GCB surface (Figure 3.1).

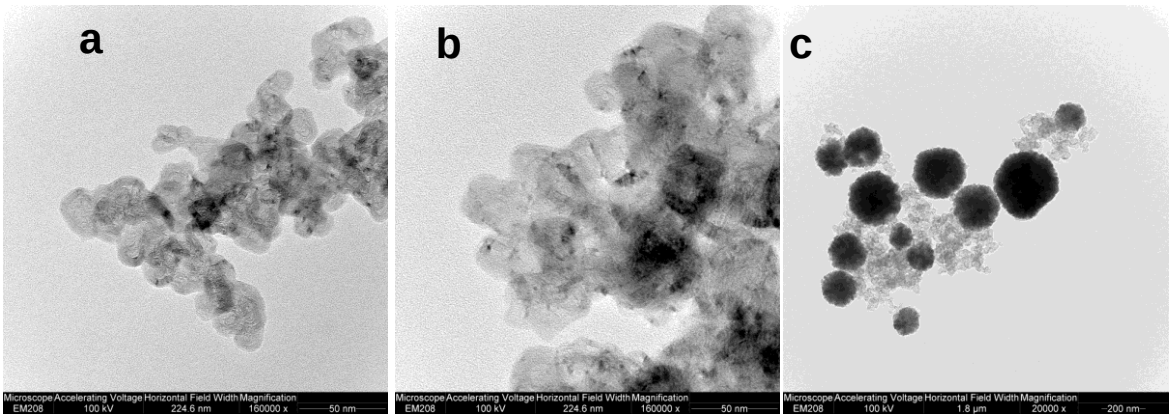


Figure 3.1: TEM images to characterize GCB during the various preparation steps: a) GCB starting material; b) GCB after treatment with HNO_3; C) GCB after magnetization.

1.3 Chromatographic and mass spectrometric method

Once characterized the material, a pre-treatment step and the clean-up step for the extraction of mycoestrogens from milk and mycotoxins from cereals had to be performed. The evaluation of the pre-treatment and clean-up procedures was performed, in both Paper I and Paper II, by means of a targeted analysis of the standards of the above mentioned mycotoxins by means of UHPLC-MS/MS.

An UHPLC coupled by means of an ESI source to a QqQ TSQ Vantage was used. The electric (e.g. S-lens) and MS and MS/MS acquisition parameters (precursor ions, SRM transitions, and collision energy) were optimized by directly infusing each analyte standard solution. Gas pressures, temperatures, and spray voltage parameters of ESI source were

optimized by introducing simultaneously, through a tee-junction, the analyte standard solution at 300 µL/min. The chromatographic conditions were slightly different in Paper I and Paper II, whereas in both papers a 10 cm long column namely C18 Cortecs was chosen.

Spectra were acquired in SRM mode, monitoring two transitions of the deprotonated molecule $[M–H]^-$ for each compound. For each analyte, the unambiguous identification was based on comparison with the authentic standard (RTs, relative intensity ratios of MRM transition pairs) following the criteria reported in the Decision 2002/657/EC[86]. Recovery (RE), ME, and Process efficiency (PE) were determined as follows: blank matrix spiked with standards before (set 1) and after (set 2) extraction procedure, and neat solution standards (set 3) were prepared; Analyte response was measured as absolute peak area. RE was assessed by the ratio between the peak area of sample set 1 and sample set 2 according to the formula:

$$RE(\%) = (Areaset1/Areaset2)\times 100$$

ME was estimated by the ratio between the peak area of sample set 2 and sample set 3 according to the formula

$$ME(\%) = (Areaset2/Areaset3)\times 100$$

PE was estimated by the ratio between the peak area of sample set 1 and sample set 3, (i.e., the product between PE and RE), according to the formula:

$$PE(\%) = (Areaset1/Areaset3)\times 100$$

1.4 Sample pre-treatment and clean-up

Once chosen the best chromatographic and mass spectrometric condition, the pre-treatment and clean-up method was evaluated. Certainly, the matrix composition affects both the pre-treatment and the selection of the best parameters for the clean-up steps [84]. Therefore, two completely different protocols were used in Paper I and Paper II starting from two different matrices and analysing two slightly different classes of compounds. Several different pre-treatment techniques were tested and several mSPE clean-up protocols were tested to evaluate the efficiency of the whole methods.

As explained in Paper I, to reduce MEs, usually milk deproteinization step is performed before extraction by adding an organic solvent, such as methanol (MeOH) or ACN [87, 88]. Three different solvents (MeOH, ACN, acidic MeOH) were tested, but measurements of the PEs showed very low values for all the solvents. Also a QuEChERS- like procedure was evaluated for sample pre-treatment with NH_4SO_4, ACN and hexane, but the RE resulted to be 4-16%. Finally, the milk sample was diluted in phosphate buffer at different ratios without any protein removal, showing the best results in term of PE in particular with 1:5 sample to buffer ratio. After chosen the best pre-treatment step, the optimization of the clean-up conditions was carried out.

The diluted skimmed milk were added to GCB MNPs, previously conditioned. After the adsorption of the analytes onto the MNPs, supernatant was removed by means of magnetic decantation. Resorcylic acid lactones were eluted from the sorbent material. The amount of MNPs (100 mg) required to achieve an efficient MSPE procedure was chosen on the basis of previous experience[89]. Magnetic decantation time depended on the medium solvent (15 min in aqueous medium, 5 min in organic solvent). Initially, ACN was used as elution solvent; however, REs were not satisfactory for α-ZEL, ZAN, and ZEN. Therefore, the mycoestrogens were eluted with the mixture CH_2Cl_2/MeOH 80:20 v/v neutral or containing 0.1% and 1% v/v HCOOH. Considering the PE, the neutral one gave the best results and was chosen as elution mixture. Therefore, in the final optimized conditions, aliquots of 8 mL of milk diluted with phosphate buffer in a sample-to-buffer ratio 1:5 were adsorbed onto the GCB MNPs and eluted with the mixture CH_2Cl_2/MeOH 80:20 v/v, then dried and reconstituted with H_2O/MeOH 50:50 v/v.

Of course, the optimized pre-treatment steps and clean-up steps were different for the analysis of mycotoxins in cereals in Paper II. Most steps of the analytical method were developed using corn meal samples, and only successively validated for wheat flour samples. Generally, ACN/water mixtures are used to extract mycotoxins from cereal samples; in some cases, acetic or formic acid up to 1% (v/v) is added to the mixture[8]. In Paper II, corn meal and durum wheat flour samples were extracted with ACN/water/formic acid 80:19.8:0.2 (v/v/v). Once the corn meal and durum wheat for samples were extracted the clean-up procedure, with the same material and procedure as in Paper I was carried out. However, the optimization of the parameters gave different results. Different sample to mGCB ratios were tested keeping the magnetic material amount to 50 mg, and using 1000 mg, 500 mg, and 250 mg of maize meal. Results, showed that the smallest sample amount allowed obtaining the

best PE values for most mycotoxins. Five mL of dichloromethane/MeOH 80:20 (v/v) containing 0.2% formic acid was used to elute the retained analytes from mGCB. A mixture dichloromethane/MeOH, either neutral or containing up to 0.2% of a weak acid, was generally used to elute the analytes from GCB in classical SPE mode. Because in the analysis of maize samples, ZEN and its derivatives were eluted a neutral mixture[90], while AFs with and acid mixture the elution conditions were chosen to enhance the REs of the AFs, which are the natural least abundant and most toxic mycotoxins. The eluted sample was dried and reconstituted with 250 µL of MeOH/water 80:20 (v/v) containing 5 $mmolL^{-1}$ ammonium formate. The choice of this reconstitution mixture was preferred to a mixture of ACN/water because of better REs very likely due to solubility issues.

1.5 Method validation

Once optimized the extraction and clean-up methods for mycoestrogens in milk and mycotoxins in cereals, the methods validation has been performed. To assess the performance of the developed analytical method, overall PE, trueness and precision (both intra-day and inter-day precision), limits of detection (LODs), and limits of quantification (LOQs) were considered.

For each analyte two calibration graphs, named "standard" and "matrix matched," respectively, were constructed. To construct the standard calibration graphs of the analytes, standard solutions were prepared at five and six concentration levels for Paper I and II, respectively, and adding the same amount of internal standard (IS). To construct the matrix-matched calibration graph for each analyte, matrix-matched solutions were prepared at five and six concentration levels for paper I and paper II, respectively, by spiking analyte free sample before the extraction to obtain in the final extracts the same nominal concentrations used for standard calibration, adding the same amount of IS in the final extract, and following the optimized extraction protocol. For each analyte, the analyte to IS peak area ratio versus the analyte concentration was plotted, considering the sum of both SRM transitions to measure the areas. Unweighted regression lines for calibration graphs were calculated.

Method trueness was estimated by analyte apparent recoveries, calculated by comparing the analyte to IS peak area ratios in free-analyte samples spiked before and after the extraction procedure.

$$\text{Apparent RE (\%)} = \frac{\text{Area analyte set 1/Area IS set 1}}{\text{Area analyte set 2/Area IS set 2}}$$

Within laboratory precision of the method was estimated as both intraday (repeatability) and interday (reproducibility). The intraday precision was expressed as relative standard deviation r (RSD) of the RE values of six spiked samples, for each concentration level, analyzed in the same day. The interday precision was expressed as RSD of the RE values of six spiked samples, for each concentration level, analyzed in six consecutive days.

The estimation of LOD and LOQ was conducted as already reported [13, 28, 31, 32]. Briefly, a first LOD and LOQ estimation was performed considering the SD of the intercept (σ) and the slope of the calibration graphs (*S*), according to the formulas

$$LOD = 3\frac{\sigma}{S} \qquad LOQ = 10\frac{\sigma}{S}$$

The standard calibration curve was used for the instrumental LODs and instrumental LOQs, whereas the matrix-matched calibration curve was used for the method (M)LODs and MLOQs. Only data generated from regression statistics obtained from calibration points in the range of the lower concentrations were employed to assess these values. The peak area of the second most intense SRM transition and the sum of the two SRM transitions were used to determine LOD an LOQ values, respectively. To verify MLODs and MLOQs, milk samples in Paper I and corn meal and wheat flour samples in Paper II were fortified with the analytes at the extrapolated values, and the samples were handled as reported above. For MLODs and MLOQ confirmation, S/N >3 and >10 were set, respectively.

Method validation results are shown in detailed in the published papers and supplementary materials. However, the results are summarized below.

In Paper I REs were >69% and MEs ranged between 67 and 115%, giving overall PEs >52%. Both standard and matrix-matched calibration graphs were linear over the explored range (R^2 >0.9989 and for R^2 =0.9956–0.9996, respectively). Apparent recoveries were > 75%. Within laboratory precision was 2–16 and 8–22% for intra-day and inter-day experiments, respectively. Calculated MLODs and MLOQs in Paper I resulted to be lower compared to other papers published in the literature. In a previous work of our group[85], using classical GCB-SPE, only for ZAN a lower MLOQ was obtained.

Table 3.1: RE, ME, PE obtained analyzing 8 mL milk samples spiked at 31.2, 156.2, and 312.5 ng/L

Analyte	Spiking level								
	31.2 ng/L			156.2 ng/L			312.5 ng/L		
	RE	ME	PE	RE	ME	PE	RE	ME	PE
β-ZAL	70	115	80	79	106	84	89	115	102
β-ZEL	79	88	70	77	112	86	84	111	93
α-ZAL	76	79	60	71	88	62	72	112	81
α-ZEL	77	67	52	82	87	71	87	68	59
ZAN	91	91	83	79	90	71	74	103	76
ZEN	88	72	63	69	96	66	73	106	77

Results are obtained by comparing analyte area without normalization to that of IS. The experiments were replicated six times. For RE and ME calculation, see text.

In Paper II REs were >67% at the lowest fortification level. The MEs of signal suppression affected in particular AFB1 and OTA. The only signal enhancement ME was observed for ZEN, however it cannot be excluded that this signal enhancement was due to a natural contamination below method detection limit (MLOD). The apparent REs were higher for OTA and ZEN compared to the others analytes. Compared to the work by Hashemi et al. [91], the MLOQs obtained in the present work for AFB2 are slightly higher. Compared to the work by Mashhadizadeh et al. [92], the limits for OTA are three times higher. However, both these works used HPLC-FD for determination, thus the criteria for MLOD and MLOQ estimation are different, and they analyzed from one up to two mycotoxins. In other papers using coated nanoparticles [93] instead, the MLODs values were comparable to those obtained in the present work, whereas the MLOQs were higher.

Table 3.2: RE, ME for AFs, OTA and ZEN in corn meal and durum wheat flour samples. Fortification levels were maximum limit (ML), 0.5 X ML and 2 X ML

Analyte	0.5 × ML				ML				2 × ML			
	Corn		Durum Wheat		Corn		Durum Wheat		Corn		Durum Wheat	
	RE	ME	RE	ME	RE	ME	RE	ME	RE	ME	RE	ME
AFG2	78	76	74	86	67	79	69	79	71	84	67	88
AFG1	71	75	74	79	66	73	74	75	68	78	73	84
AFB2	69	76	68	78	63	85	76	86	74	89	71	76
AFB1	73	68	71	70	74	69	73	68	72	69	69	73
OTA	67	72	73	67	83	68	76	71	79	69	81	68
ZEN	78	87	79	102	89	94	82	116	84	104	88	108

1.6 Method application

One of the problems arising during method development in Paper II was to find samples free of ZEN, which, indeed, contaminates most of the corn meal samples[94]. OTA was also detected in some samples at concentrations below MLOQ. In a short survey carried out on 10 corn meal samples, OTA was detected in one sample at 1.3 µg kg^{-1} level, whereas it was detected at values >MLOD but <MLOQ in the other two samples. In eight out of 10 samples, ZEN was detected at values >MLOD but <MLOQ, and in one sample at 72.9 1.3 µg kg^{-1} level. Quantification was made by matrix-matched calibration. None of the investigated mycotoxins were detected in the five durum wheat flour samples. Indeed, the maize plant is very susceptible to contamination by Fusarium species [94], so it is generally more affected by certain mycotoxin contamination than wheat (in particular ZEN contamination).

1.7 Conclusions

In Paper I is shown that the overall PE ranged between 52 and 102%, with ME below 33%. The obtained method limits of quantification were below those of other published methods that employ classical SPE protocols. The MLOQs ranging between 8 and 15 ng/L make the present method able to determine the compounds of interest in milk samples at reasonable contamination levels. Moreover, the employment of MSPE instead of SPE allowed to extract a turbid sample as milk avoiding possible clogging of SPE cartridge frit, as well as decreased RE due to colloid–analyte interaction. Although the equilibration time (NPs-analytes interaction time 30 min) could appear long, it should be considered that, on the contrary of SPE, a large number of samples can be processed simultaneously, and the time needed for analyte retrieval is about the same as in the classical SPE.

Due to the positive results obtained for this method development applied to mycoestrogens in milk, we decided to use the same material on a different class of mycotoxins in cereals. In Paper II is shown that REs were >60% in both cereals analyzed, even if the MEs were not negligible. The LODs were comparable to or higher than those of other mSPE methods. Firstly, the different detection techniques used in the other works make MLOQ calculation modes not really comparable Secondly, this method, on the contrary of

the previous works, allows the simultaneous investigation of a larger number of mycotoxins belonging to three different chemical classes.

These results show that several aspects, related to sample preparation, have to be considered when developing a method for a targeted analysis: recoveries, time, costs, and applicability to different matrices or to a wide range of diverse compounds. The overall process efficiency of the developed method has to be a compromise between performance and easiness and rapidity of application. The trend in contaminants analysis in food is moving toward the development of multi-residual analysis, consisting in the analysis of a wide range of compounds (which is difficult because of the different chemical properties) in an effective and low time consuming process.

2 Paper I

Mycoestrogen determination in cow milk: Magnetic solid-phase extraction followed by liquid chromatography and tandem mass spectrometry analysis


ISSN 1615-9306 · JSSCCJ 39 (24) 4667–4876 (2016) · Vol. 39 · No. 24 · December 2016 · D 10609
JOURNAL OF
SEPARATION SCIENCE
24|16
Methods
Chromatography · Electroseparation
Applications
Biomedicine · Foods · Environment
www.jss-journal.com
WILEY-VCH

Anna Laura Capriotti
Chiara Cavaliere
Patrizia Foglia
Giorgia La Barbera
Roberto Samperi
Salvatore Ventura
Aldo Laganà

Department of Chemistry, University of Rome "La Sapienza", Rome, Italy

Received August 5, 2016
Revised October 11, 2016
Accepted October 13, 2016

Research Article

Mycoestrogen determination in cow milk: Magnetic solid-phase extraction followed by liquid chromatography and tandem mass spectrometry analysis

Recently, magnetic solid-phase extraction has gained interest because it presents various operational advantages over classical solid-phase extraction. Furthermore, magnetic nanoparticles are easy to prepare, and various materials can be used in their synthesis. In the literature, there are only few studies on the determination of mycoestrogens in milk, although their carryover in milk has occurred. In this work, we wanted to develop the first (to the best of our knowledge) magnetic solid-phase extraction protocol for six mycoestrogens from milk, followed by liquid chromatography and tandem mass spectrometry analysis. Magnetic graphitized carbon black was chosen as the adsorbent, as this carbonaceous material, which is very different from the most diffuse graphene and carbon nanotubes, had already shown selectivity towards estrogenic compounds in milk. The graphitized carbon black was decorated with Fe_3O_4, which was confirmed by the characterization analyses. A milk deproteinization step was avoided, using only a suitable dilution in phosphate buffer as sample pretreatment. The overall process efficiency ranged between 52 and 102%, whereas the matrix effect considered as signal suppression was below 33% for all the analytes even at the lowest spiking level. The obtained method limits of quantification were below those of other published methods that employ classical solid-phase extraction protocols.

Keywords: Dispersive solid-phase extraction / Graphitized carbon black / Magnetic nanoparticles / Mycoestrogens / Resorcylic acid lactones
DOI 10.1002/jssc.201600879

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Zearalenone (ZEN), a macrocyclic β-resorcylic acid lactone, is a mycotoxin produced by several filamentous fungi of *Fusarium* species. ZEN and its metabolites and derivatives, namely α-zearalenol (α-ZEL), β-zearalenol (β-ZEL), α-zearalanol (α-ZAL), β-zearalanol (β-ZAL), and zearalanone (ZAN), are classified as mycoestrogens [1] because they are able to bind the estrogen receptor and consequently lead to problems in the mammalian reproductive system [2]. In particular, α-ZAL is used as a growth promotor and to reduce stress in cattle in the USA and several other countries, whereas it has been banned in the European Union (EU) since 1985 [3].

Mycotoxins can be found on different crops, in particular cereal grains. However, human exposure might occur not only directly through the intake of contaminated agricultural products, but also indirectly through the consumption of products of animal origin (e.g. meat, milk, eggs, etc.) obtained from animals that were fed with contaminated feed [4]. Indeed, carryover of ZEN and/or its metabolites into cow milk has been reported [2], since in mammals, ZEN is mainly biotransformed in β-ZEL and α-ZEL [5].

Due to their toxic effects, Council Directive 96/23/EC [6] lists resorcylic acid lactones in group A, substances having anabolic effect and unauthorized substances, which have to be monitored in live animals, and animal products. Moreover, in 2000, the Joint FAO/WHO Expert Committee on Food Additives has recommended a total provisional

Correspondence: Professor Chiara Cavaliere, Dipartimento di Chimica, Università di Roma "La Sapienza", Piazzale Aldo Moro 5, 00185 Rome, Italy
E-mail: chiara.cavaliere@uniroma1.it
Fax: +39-06-490631

Abbreviations: CNT, carbon nanotube; **GCB**, graphitized carbon black; **IS**, internal standard; **MLOD**, method limit of detection; **MLOQ**, method limit of quantification; **MNP**, magnetic nanoparticle; **MSPE**, magnetic solid-phase extraction; **ME**, matrix effect; **pDA**, polydopamine; **PE**, process efficiency; **RE**, recovery; **SRM**, selected reaction monitoring; **TGA**, thermogravimetric analysis; **ZAN**, zearalanone; **ZEN**, zearalenone; **α-ZEL**, α-zearalenol; **β-ZEL**, β-zearalenol; **α-ZAL**, α-zearalanol; **β-ZAL**, β-zearalanol; **ZEN-d6**, deuterated zearalenone

maximum tolerable daily intake of 0.5 μg/kg b.w. for ZEN and its metabolites. Resorcylic acids are not listed among the substances to be detected in milk [6]. Although the limited data available seem to indicate that in milk this mycotoxin class does not pose a serious danger to humans [2], nevertheless it has to be considered that the effect of estrogenic substances might be additive or synergic, and the knowledge of every estrogenic compound source is important for human health protection. Therefore, the availability of rapid and reliable analytical methods for the determination of resorcylic acid lactones in animal-derived matrices is highly important for food safety assurance [7].

SPE followed by LC coupled to MS are the techniques of choice for the analysis of mycoestrogens in milk [5, 8], that are analyzed alone [9] or more frequently together with other mycotoxins [10, 11] or estrogenic compounds [3, 12]. Nevertheless, even using a such selective technique, an adequate sample pretreatment is required for the elimination of matrix interferences and for the enrichment of the selected analytes [8]. To reduce matrix effects (MEs), i.e., ion suppression or enhancement typically observed when an ESI source is used, and/or to preserve SPE or LC columns, usually milk deproteinization step is performed before extraction by adding an organic solvent, such as MeOH or ACN [8, 14]. However, in this case, the resulting solution possesses a too high organic content to be retained, for example, on a conventional RP SPE column and requires a suitable aqueous dilution [15].

Recently, magnetic nanoparticles (MNPs) have been proposed as sorbents for dispersive SPE; their advantages over classical SPE sorbents have been already reported in several reviews [16, 17]. The magnetic core of MNPs, usually constituted by an iron oxide such as magnetite (Fe_3O_4), can be grafted or coated with organic or inorganic layers, to enhance MNP stability and selectivity [18]. Silica is the material most used for coating, but also carbonaceous materials are gaining interest. In other cases, Fe_3O_4 is deposited onto the structure of a carbon-based material [19], such as graphene [20] and graphene oxide [21], multi-walled carbon nanotubes (CNTs) [22], (nano)porous carbon [23,24], and carbon nanofibers [25].

As in classical SPE, the separation mechanism occurring in magnetic (M)SPE depends on the type of sorbent, and is connected with the interaction of analyte molecules with the functional groups of the MNPs. The choice of the right sorbent also depends on the nature of the sample being tested [17].

MSPE has never been used for mycoestrogens determination in milk, whereas MNPs with various coating have been employed for extracting estrogens from milk: polypyrrole [14], CNTs [26], and cetyltrimethyl ammonium bromide [27].

In a recent work [28], we exploited the ability of MNPs coated with polydopamine (pDA) to selectively retain estrogens, mycoestrogens and phytoestrogens from water samples. The Fe_3O_4@pDA particles gave high recoveries for the analytes when used in water samples, nevertheless they failed when applied to milk samples, because ME dramatically increased.

In this work, we wanted to test the capability of magnetic graphitized carbon black (GCB) in mycoestrogen extraction from milk samples. For this purpose, GCB, previously treated with HNO_3, was magnetized and used as MSPE sorbent before UHPLC–MS/MS analysis. To author's best knowledge, this is the first application of MSPE for mycoestrogens in milk, as well as the employment of magnetic GCB in a food matrix.

2 Materials and methods

2.1 Chemical and reagents

Formic acid, nitric acid, ethylene glycol, iron(III) chloride hexahydrate, trisodium citrate, sodium acetate, and organic solvents of analytical grade were obtained from Sigma–Aldrich (St. Louis, MI, USA). For LC–MS analysis, only ultrapure water and ACN Optima LC/MS grade from Fisher Scientific (Illkirch, France) were used.

Standards of the analytes (purity >99%) ZEN, α-ZEL, β-ZEL, ZAN, α-ZAL, β-ZAL were purchased from Sigma–Aldrich (St. Louis, MI, USA). The pure standard of the commercially available deuterated zearalenone-d6 (ZEN-d6) was acquired from Wellington Laboratories (Toronto, Ontario, Canada) and was used as internal standard (IS).

For each analyte and the IS, an individual stock standard solution at 200 μg/mL concentration was prepared by dissolving the suitable amount of standard in MeOH. A composite working standard solution was prepared by combining aliquots of the individual solutions of the six analytes and diluting with MeOH to obtain the required concentration. This mixture was renewed weekly. All the solutions were stored in the dark at −20°C and brought to room temperature before use.

2.2 Preparation of graphitized carbon black magnetic nanoparticles

The GCB MNPs were prepared following a literature preparation of magnetic carbon nanotubes, with some modifications [29]. Four-hundred mg of the commercial GCB Supelclean ENVI-Carb (Sigma–Aldrich, St. Louis, MO) was initially activated by stirring the material in 50 mL of concentrated nitric acid at room temperature for 7 h. At the end of this step, and after extensive washings with water until neutral pH, the material was dried overnight at 50°C. Finally, 150 mg of the activated GCB were dispersed into 40 mL of ethylene glycol and added with $FeCl_3 \cdot 6H_2O$ (810 mg), trisodium citrate (150 mg), sodium acetate (3.6 g), and PEG 10k (1.0 g). The mixture thus obtained was sonicated for 3 h and then sealed in a 125 mL autoclave for 10 h at 200°C.

After cooling to room temperature, the obtained GCB MNPs were washed with water and ethanol and dried; finally, they were stored sealed in a glass flask at 4°C in the dark.

2.3 Characterization of graphitized carbon black material

The GCB material was characterized before and after treatment with nitric acid, and after magnetization. TEM images were obtained with a Philips EM 208 transmission electron microscope with a standard loop filament, Olimpus Morada 2K × 2K CCD camera. Porosimetry and specific surface area were assessed by an 3Flex Surface Characterization Analyzer (Micromeritics Instrument Corporation, Norcross, GA, USA). The FTIR spectra were recorded on Perkin Elmer 2000 spectrophotometer. All the spectra were collected in the range 4000–400 cm^{-1}. Thermogravimetric analysis (TGA) was performed on TA Instruments TGA Q500 under an air flow of 90 mL/min with a ramp of 10°C/min from 100 to 800°C. Raman spectra were recorded on a Renishaw instrument, model Invia reflex equipped with 532 laser operating at 22.5 W.

All the analyses were duplicated.

2.4 Samples and sample preparation

Pasteurized whole milk packages of the most widespread brands of the area were acquired in local markets in Rome. As reported on the label, composition was 3.6 g of fat and 3.3 g of protein per 100 mL; original pH ranged between 6.6 and 6.8.

Whole milk was previously skimmed by centrifugation at 12 700 × g at 4°C for 15 min; afterward, cream was removed from the surface and the underlying milk was used for experiments.

A total of 8 mL of skimmed milk was placed in a 50 mL polypropylene centrifuge tube and added with 40 mL of 60 mmol/L phosphate buffer (pH 7). After that, 100 mg of GCB MNPs, previously conditioned with 10 mL MeOH followed by 5 mL H_2O, were added to the mixture. After 1 min vortexing, the solution was kept under agitation for 30 min to allow the adsorption of the analytes onto the MNPs. Before removing the supernatant, a permanent magnetic disk composed Nd-Fe-B of 25 mm × 5 mm from Supermagnete (Gottmadingen, Germany) was placed on the bottom of the polypropylene tube for 15 min to sediment the GCB MNPs. To remove residual water, 500 µL MeOH were added in the tube and then the supernatant was discarded.

Resorcylic acid lactones were eluted from the sorbent material by 10 mL CH_2Cl_2/MeOH 80:20 v/v under slight agitation for 5 min, and then by 5 min magnetic decantation. The extract was collected in a 20 mL round bottom glass vial and evaporated under a gentle nitrogen stream in a water bath at 37°C, and the residue was reconstituted with 250 µL of H_2O/MeOH 50:50 v/v. Before injection, the extract was filtered with 13 mm GHP membrane syringe filters (0.2 µm, Pall, MI, USA).

2.5 LC–MS/MS analysis

The mass spectrometer was a TSQ Vantage EMR triple quadrupole (Thermo Fisher Scientific, Bremen, Germany), provided with a heated ESI source. It was interfaced to an UHPLC system Ultimate 3000 binary pump (Thermo Fisher Scientific) equipped with thermostatted column compartment and microwell-plate autosampler.

Analytes were separated onto a Cortecs UHPLC C18+ column (100 mm × 2.1 mm id, 1.6 µm particle size), provided with a Cortecs UHPLC C18+ VanGuard pre-column (5 mm × 2.1 mm id, 1.6 µm particle size), both from Waters (Milford, MA, USA). The column was maintained at 40°C in the thermostatted oven of the LC apparatus, and it was operated at 300 µL/min flow rate; sample aliquots of 10 µL were injected. Using water (A) and ACN (B) as mobile phase, the elution gradient consisted in an isocratic step at 15% B for 2 min, followed by a linear increase of B to 60% within 14 min; then B was brought to 99% in 4 min, and held constant for 10 min to rinse the column. Equilibration time at 15% B lasted 10 min.

The ESI source was operated in negative ion mode: spray voltage was set at −2.8 kV; vaporizer and capillary temperatures were set at 250 and 220°C, respectively; gas pressures were 40 and 20 (arbitrary units) for sheath and auxiliary gas, respectively. Spectra were acquired in selected reaction monitoring (SRM) mode, monitoring two transitions of the deprotonated molecule $[M–H]^-$ for each compound (see Table 1).

Table 1. Full names, acronym, chromatographic retention time, precursor ion, product ions with the relative collision energy (CE), and S-lens of the target analytes and the deuterated internal standard (IS)

Analyte	Abbreviation	Retention time (min)	Precursor ion $[M-H]^-$ (*m/z*)	Product ion (*m/z*) (CE, eV)	S-lens
β-zearalanol	β-ZAL	10.7	321.0	161.1 (32); 277.2 (23)	125
β-zearalenol	β-ZEL	10.9	319.1	188.0 (30); 275.1 (23)	150
α-zearalanol	α-ZAL	11.9	321.0	132.9 (38); 161.2 (32)	125
α-zearalenol	α-ZEL	12.1	319.1	160.0 (36); 275.1 (23)	150
zearalanone	ZAN	13.5	319.1	275.2 (22); 161.1 (29)	125
zearalenone-d6	ZEN-d6 (IS)	13.6	323.0	161.1 (34); 175.1 (26)	125
zearalenone	ZEN	13.6	317.0	131.1 (33); 175.1 (25)	150

Once per month, the manufacturer solution (mass range 69 to 2800 m/z) was used for mass calibrations and resolution adjustments on the resolving lens and quadrupoles.

The entire LC–MS system was managed by software Xcalibur™, v.2.2 (Thermo Fisher Scientific).

2.6 Evaluation of analytical method performance

Performance of the optimized analytical method was assessed through overall process efficiency (PE), trueness and precision (both intraday and interday precision), LODs, and LOQs.

2.6.1 Standard and matrix-matched calibration graphs

Two calibration graphs, named "standard" and "matrix-matched," respectively, were constructed. Standard solutions were prepared at five concentration levels (0.5, 1.0, 2.0, 5.0, and 10 pg/μL) by appropriate dilution of the composite working standard solution in H_2O/MeOH (50:50, v/v), and adding the same amount of IS (2 pg/μL). These solutions were used to construct the standard calibration graphs of the analytes.

Matrix-matched solutions were prepared at five concentration levels (15.6, 31.2, 62.5, 156.2, 312.5 ng/L) by spiking analyte-free milk samples, to obtain in the final extracts the same nominal concentrations used for standard calibration, adding the same amount of IS (2 pg/μL in the final extract, corresponding to 62.5 ng/L in milk) and following the optimized extraction protocol. These solutions were used to construct the matrix-matched calibration graph for each analyte.

Each solution was prepared in duplicate and injected twice, starting from the lowest up to the highest concentration level; finally, the results were averaged to give rise to a single calibration graph.

For each analyte, the combined ion current profile for the selected transitions was extracted from the LC–SRM dataset; the resulting traces were smoothed (Gaussian type, 7 points) by applying the automatic processing smoothing of Xcalibur™ software.

The analyte to IS peak area ratio versus the analyte concentration was plotted, considering the sum of both SRM transitions to measure the areas. Unweighted regression lines for standard and matrix-matched calibration graphs were calculated using Xcalibur™ QuanBrowser (Thermo Fisher Scientific).

2.6.2 Recovery, matrix effect, and overall process efficiency

Recovery (RE), ME, and PE were determined following Matuszewski approach [30], as also reported in a previous work [28]. For this purpose, blank matrix spiked with standards before (set 1) and after (set 2) extraction procedure, and neat solution standards (set 3) were prepared. Analyte response was measured as absolute peak area.

RE was assessed by the ratio between the peak area of sample set 1 and sample set 2 according to the formula: $RE(\%) = (Area_{set1}/Area_{set2}) \times 100$. ME was estimated by the ratio between the peak area of sample set 2 and sample set 3 according to the formula $ME(\%) = (Area_{set2}/Area_{set3}) \times 100$. PE was estimated by the ratio between the peak area of sample set 1 and sample set 3, according to the formula: $PE(\%) = (Area_{set1}/Area_{set3}) \times 100$. (i.e., the product between PE and RE).

Three different fortification levels were prepared by adding the suitable volume of the composite working standard solution of the mycoestrogens to the sample sets. The levels were 31.2, 156.2, and 312.5 ng/L in milk samples, corresponding to 1, 5, and 10 pg/μL in the neat standard solutions. For each spiking level, six replicates were performed.

2.6.3 Trueness and precision

Method trueness was estimated by analyte apparent recoveries, calculated by comparing the analyte to IS peak area ratios in free-analyte milk samples spiked before and after the extraction procedure. Resorcylic acid lactones fortification levels were 31.2, 156.2, and 312.5 ng/L, whereas the concentration level of IS was 2 pg/μL in the final extract (i.e., 62.5 ng/L if referred to milk) in all samples. Six replicates were carried on for each spiking level.

Within laboratory precision of the method was estimated as both intraday (repeatability) and interday (reproducibility). The intraday precision was expressed as RSD_r of the recovery values of six spiked samples, for each concentration level, analyzed in the same day. The interday precision was expressed as RSD_R of the recovery values of six spiked samples, for each concentration level, analyzed in six consecutive days.

2.6.4 Limits of detection and quantification

The estimation of LOD and LOQ was conducted as already reported [13,28,31,32]. Briefly, a first LOD and LOQ estimation was performed considering the SD of the intercept (σ) and the slope of the calibration graphs (S), according to the formulas $LOD = 3\,\sigma/S$ and $LOQ = 10\,\sigma/S$. The standard calibration curve was used for the instrumental LODs and instrumental LOQs, whereas the matrix-matched calibration curve was used for the method (M)LODs and MLOQs.

Only data generated from regression statistics obtained from standard and matrix-matched calibration points in the range of the lower concentrations were employed to assess these values. The peak area of the second most intense SRM transition and the sum of the two SRM transitions were used to determine LOD an LOQ values, respectively.

To verify MLODs and MLOQs, milk samples were fortified with the analytes at the extrapolated values, and the samples were handled as reported above. For MLODs and MLOQ confirmation, S/N >3 and >10 were set, respectively.

2.7 QC

For each processed sample batch, a blank sample was randomly injected. However, neither contamination nor carryover was detected during method development and validation.

3 Results and discussion

3.1 LC–MS/MS conditions

The electric (e.g. S-lens) and MS and MS/MS acquisition parameters (precursor ions, SRM transitions, and collision energy) were optimized by directly infusing each analyte standard solution at 1 ng/μL in H_2O/ACN (50:50, v/v). Gas pressures, temperatures, and spray voltage parameters of ESI source were optimized by introducing simultaneously, through a tee-junction, the analyte standard solution at 1 ng/μL and H_2O/ACN (50:50, v/v) at 300 μL/min. The main MS and MS/MS parameters are reported in Table 1.

For chromatographic mobile phase, ACN was used as organic modifier on the basis of previous experience [13], where ACN resulted better than MeOH in terms of S/N. No additives were added to the mobile phase. Initially, a 5 cm long column was used, however, to better separate the analytes from some coeluting interferences, a 10 cm long column (see Subsection 2.5) was preferred. Indeed, injecting the same blank extract spiked with the analytes onto the two chromatographic columns, with the longer one MEs were less pronounced (from 35–61% to 55–84%).

3.2 Choice of magnetic graphitized carbon black as adsorbent material

In MSPE, the choice of the suitable sorbent depends not only on the analyte nature, but also on the composition of the matrix, because interfering substances could affect the strength of the interaction of the analyte with the sorbent [17]. Moreover, the nonselective adsorption on the MNPs of matrix components could lead to a pronounced ME. Another important parameter is the elution solvent that should have a suitable elution strength to quantitatively remove the analytes from the surface of the sorbent.

Several carbonaceous nanomaterials have been proposed as adsorbents in the analytical chemistry field, including for MSPE applications [19]. The most successful ones are graphene that has shown excellent properties, especially in oxide form [33], and CNTs [16, 34].

Starting from this point, we decided to exploit the performance of GCB in MSPE. The peculiar characteristic of GCB material, making it a suitable adsorbent for both polar and nonpolar molecules, have been deeply described in previous works [35, 36]. Briefly, GCB may establish different types of interaction with the organic molecules, acting both as RP and anion-exchanger sorbent. Indeed, the planar graphitic surface allows the hydrophobic interaction with analytes by polarizable bonds and π interactions; furthermore, some polar and positively charged chemical heterogeneities (e.g., oxonium groups) on its surface allow the GCB to establish anion–exchange interactions and hydrogen bridges between protonated functional groups of the analytes and carbonylic groups of GCB or vice versa. GCB is a non-porous material whose surface area varies among the diverse commercially available products.

Taking into account that our group has already successfully employed GCB as adsorbent in SPE of free and conjugated estrogens [37] and estrogens and mycoestrogens from milk [13], we decided to test GCB MNPs for the extraction of mycoestrogens from milk.

Because of the lack of proper functional groups on their surface, some carbonaceous materials require an oxidation step to improve the wettability and dispersion in water, as in the case of graphene, very often used in its oxide form [38], and multi-walled CNTs [22]. GCB, even possessing some polar functionalities in its structure, was previously treated with HNO_3 in mild conditions, which led to an only 1% oxidation. This is because in preliminary experiments with non-oxidized GCB (data not shown), elution of the analytes from MNPs required a high volume of eluent, which could be reduced only after oxidation.

3.3 Characterization of graphitized carbon black material

The GCB material was fully characterized before and after oxidation with nitric acid, as well as after magnetization, to observe changes in morphology or in structure.

The starting material was a GCB with particle size of 120–400 mesh and a surface area of 100 m^2/g, as declared by the manufacturer. The porosimetry analysis confirmed these data, providing a surface area of 98.0 ± 1 m^2/g. Additionally, in the starting material a mesopore structure was observed, with a total pore volume of 0.60 cm^3/g diameters in the range 20–35 and 80–800 Å (largest population, maximum at 460 Å). After magnetization, the available surface area and the pore volume were half as much, 55.0 ± 0.5 m^2/g and 0.30 cm^3/g, respectively. The mesoporous structure showed no significant changes in the pore diameters (20–40 and 80–800 Å), but the distribution was affected and the first range was representative of half the population.

The increase of inorganic species in the material was monitored by TGA in air (See Supporting Information Fig. S1). The starting GCB was found free of organic impurities (no weight loss was observed at low temperatures) or inorganic impurities (no loss above 720°C), whereas the sharp weight loss in the 650–720°C range confirmed high homogeneity of the graphitic structure. After oxidation, such profile was maintained, with only a 2% weight loss at 600°C as a result of the chemical oxidation. After magnetization, however, the inorganic components showed a significant increase to 55%.

Table 2. Recovery (RE), matrix effect (ME), and overall process efficiency (PE, product of RE and ME) obtained by the analysis of milk sample fortified at 100 ng/L with the analytes by varying the sample pretreatment: (i) protein precipitation with cold MeOH; (ii) protein precipitation with ACN; (iii) protein precipitation with acidic MeOH; (iv) phase partition with ACN and hexane

Analyte	MeOH			ACN			Acidic MeOH			ACN+hexane		
	RE	ME	PE	RE	ME	PE	RE	ME	PE	RE	ME	PE
β-ZAL	112	18	20	68	40	27	59	42	25	7	112	8
β-ZEL	51	15	8	65	41	27	45	44	20	10	97	10
α-ZAL	91	16	15	55	35	19	70	34	24	4	105	4
α-ZEL	49	17	8	46	44	20	54	43	23	13	107	14
ZAN	96	13	12	76	38	29	62	38	24	16	103	16
ZEN	83	12	10	85	37	31	62	37	23	14	109	15

Results are reported as average from duplicate experiments. For RE and ME calculation, see text.

The FTIR analysis of the starting GCB was consistent with a graphitic structure and showed the typical C–H stretching (2880 and 2865 cm^{-1}), rocking (1350 cm^{-1}), and scissoring (1430 cm^{-1}). After oxidation, additional signals were monitored, namely a broad band for the OH stretch (3000–3600 cm^{-1}), and signals indicating C=O and C–O bonds (1400–1700 and 1000–1300 cm^{-1}). After magnetization, signal from magnetite nanoparticles were observed below 600 cm^{-1}, including the typical Fe–O bond signal at 590 cm^{-1}.

The Raman analysis was consistent with a pure graphitic structure, with typical D, G, and 2D bands at ca. 1350, 1580, and 2700 cm^{-1} respectively (ID/IG of 0.72), which remained unaffected after oxidation. Presence of magnetite was next confirmed by the signal at ca. 670 cm^{-1} (See Supporting Information Fig. S2).

The morphological structure of the materials was monitored by TEM analysis, which showed that after magnetization spherical nanoparticles (100–300 nm) were grown on the GCB surface (See Supporting Information Fig. S3).

3.4 Sample pretreatment

3.4.1 Milk deproteinization by organic solvents

For milk deproteinization, three different solvents were tested: (i) MeOH; (ii) ACN; (iii) acidic MeOH (by adding an HCOOH amount able to lower pH value to 4.6 in the same volume of H_2O).

For protocol (i), 1 mL of skimmed milk was fortified at 100 ng/L with the analytes, added with 5 mL of cold MeOH and kept at –20°C for 30 min to allow both protein and all lipid precipitation. After centrifugation at 7200 × g at 4°C for 15 min, the supernatant was collected and diluted to 25 mL with ultrapure H_2O. Then, the MNPs were added following the extraction procedure reported in the Experimental section, with the only difference that elution from MNPs was achieved by 5 mL ACN. However, a pronounced ME (signal suppression) was observed.

The procedure was analog by using (ii) ACN and (iii) acidic MeOH for milk deproteinization. By adding the three different solvents, similar results were obtained (see Table 2). With the use of ACN, lipids were not removed from the sample, however the ME was less pronounced than that observed by using MeOH, and very similar to that observed by adding acidic MeOH. Considering the PEs, ACN performed better than MeOH and acidic MeOH; however, the very low values obtained even with this solvent (19>PEs<31), led us to look for a sample treatment different from deproteinization.

3.4.2 Phase partition with acetonitrile and hexane

For the extraction of resorcylic acid lactones, also a QuEChERS-like procedure was evaluated for sample pretreatment. In this case, 5 mL of skimmed milk was spiked with the analytes at 100 ng/L and was added with 3.5 g of NH_4SO_4, 5 mL ACN, and 5 mL hexane. After 2 min vortexing, the solution was centrifuged at 7200 × g at 4°C for 15 min, obtaining the separation in three phases, on the top hexane, in the middle ACN and on the bottom aqueous phase. The ACN layer was collected and diluted to 25 mL with H_2O. After that, 100 mg of magnetic GCB was added; after 30 min shaking, elution was carried on with 10 mL ACN. The recovery ranged between 4–16%, even if ME was negligible (see Table 2).

3.4.3 Dilution in phosphate buffer

Finally, the milk sample was diluted in phosphate buffer (as described in Section 2) without any protein removal, as reported by other authors [14] for the MSPE of estrogens from milk. Following this protocol, Gao et al. [14] obtained performance comparable with those of other published SPE-based methods involving protein removal step.

Therefore, 2 mL skimmed milk samples fortified with the analytes at 100 ng/L were added with 60 mmol/L phosphate buffer, in the sample to phosphate buffer ratios 1:2, 1:4, 1:5,

Table 3. Recovery (RE), matrix effect (ME), and overall process efficiency (PE, product of RE and ME) obtained by the analysis of 2 mL milk samples fortified at 100 ng/L with the analytes by varying the sample to phosphate buffer ratio

Analyte	Milk sample:phosphate buffer											
	1:2			1:4			1:5			1:10		
	RE	ME	PE	RE	ME	PE	RE	ME	PE	RE	ME	PE
β-ZAL	82	144	118	82	105	86	84	141	118	89	142	126
β-ZEL	87	117	102	46	110	50	93	79	73	75	82	36
α-ZAL	43	83	36	80	51	41	89	69	61	56	99	56
α-ZEL	79	71	56	58	65	37	65	76	49	51	96	49
ZAN	32	105	34	61	64	39	59	97	57	67	97	64
ZEN	50	74	37	48	68	33	47	80	38	62	93	58

Phospahte buffer concentration was 60 mmol/L and the sample to buffer ratios were 1:2, 1:4, 1:5, and 1:10. Elution was obtained by ACN. Results are reported as average from duplicate experiments. For RE and ME calculation, see text.

Table 4. Recovery (RE), matrix effect (ME), and overall process efficiency (PE, product of RE and ME) obtained by the analysis of milk samples fortified at 50 ng/L with the analytes by varying the acidity of the elution mixture CH_2Cl_2/MeOH 80:20 (*v/v*): (i) neutral; containing 0.1% (*v/v*) HCOOH; (ii) containing 1% (*v/v*) HCOOH

Analyte	Elution mixture								
	CH_2Cl_2/MeOH 80:20			CH_2Cl_2/MeOH 80:20 with 0.1% HCOOH			CH_2Cl_2/MeOH 80:20 with 1% HCOOH		
	RE	ME	PE	RE	ME	PE	RE	ME	PE
β-ZAL	89	112	100	91	60	55	85	58	49
β-ZEL	77	117	90	98	54	53	105	55	58
α-ZAL	69	114	79	81	69	56	68	76	52
α-ZEL	77	74	57	90	77	69	84	75	63
ZAN	70	103	72	89	84	75	85	79	67
ZEN	70	111	78	94	81	76	92	77	71

Results are reported as average from duplicate experiments. For RE and ME calculation, see text.

Table 5. Recovery (RE), matrix effect (ME), and overall process efficiency (PE, product of RE and ME) obtained analyzing 8 mL milk samples spiked at 31.2, 156.2, and 312.5 ng/L

Analyte	Spiking level								
	31.2 ng/L			156.2 ng/L			312.5 ng/L		
	RE	ME	PE	RE	ME	PE	RE	ME	PE
β-ZAL	70	115	80	79	106	84	89	115	102
β-ZEL	79	88	70	77	112	86	84	111	93
α-ZAL	76	79	60	71	88	62	72	112	81
α-ZEL	77	67	52	82	87	71	87	68	59
ZAN	91	91	83	79	90	71	74	103	76
ZEN	88	72	63	69	96	66	73	106	77

Results are obtained by comparing analyte area without normalization to that of IS. The experiments were replicated six times. For RE and ME calculation, see text.

and 1:10. The different ratios were tested to reduce the sample volume before adding MNPs and thus facilitate the analyte–adsorbent interactions. Then, 100 mg MNPs were added and after 30-min shaking, the elution was achieved with 10 mL ACN.

Among sample pretreatments, the phosphate buffer dilution gave the best results in terms of PE (see Table 3), in particular with the 1:5 sample to buffer ratio.

We used a buffer concentration of 60 mmol/L instead of 10 mmol/L in a sample-to-buffer ratio 1:9 used by Gao

Table 6. Trueness, measured by apparent recoveries obtained at three spiking levels: 31.2, 156.2, and 312.5 ng/L, maintaining the amount of the IS at 2 pg/μL in the final extract, corresponding to 62.5 ng/L in milk; within laboratory precision, reported as intraday and interday RSD_r and RSD_R, respectively ($n = 6$); method limit of detection (MLOD) and method limit of quantitation (MLOQ)

Analyte	Trueness						
	31.2 ng/L	156.2 ng/L	312.5 ng/L	Intraday RSD_r	Interday RSD_R	MLOD (ng/L)	MLOQ (ng/L)
β-ZAL	97	95	102	9	21	6	13
β-ZEL	98	106	97	2	8	5	13
α-ZAL	93	96	94	5	12	3	9
α-ZEL	103	107	105	9	10	9	15
ZAN	98	102	99	7	17	6	14
ZEN	101	99	104	16	22	3	8

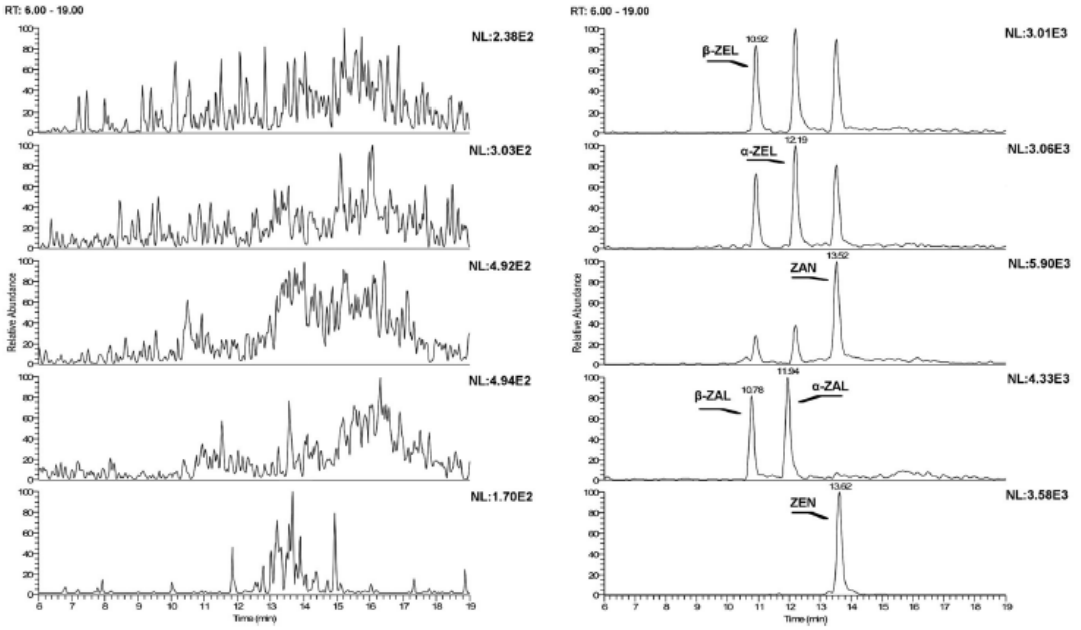


Figure 1. LC–SRM mass chromatograms of the six analytes obtained by the analysis of a blank milk sample (left panel) and a milk sample spiked at 62.5 ng/L before extraction (right panel).

et al. [14], to reduce the added volume of buffer. However, the role of phosphate buffer is not clear; the simple dilution in water was not effective in reducing ME.

3.5 Optimization of the extraction conditions

The amount of MNPs (100 mg) required to achieve an efficient MSPE procedure was chosen on the basis of previous experience [28] and from preliminary experiments (data not shown). Magnetic decantation time depended on the medium solvent: at least 15 min were necessary for a complete MNP deposition in aqueous medium, whereas 5 min was enough when the medium was an organic solvent, as in the elution step.

3.5.1 Elution solvent

Initially, ACN was used as elution solvent; however, even after sample pretreatment optimization, REs were not satisfactory for α-ZEL, ZAN, and ZEN (see Table 3), very likely because 10 mL ACN was not able to completely elute the analytes from magnetic GCB. This hypothesis was indirectly verified by analyzing the loading solution in which analytes were below LOD (data not shown).

Therefore, 2 mL skimmed milk samples were spiked at 50 ng/L with the analytes, diluted in phosphate buffer (1:5 sample to buffer ratio) and, after adding MNPs and 30-min shaking, the mycoestrogens were eluted with the mixture CH_2Cl_2/MeOH 80:20 v/v neutral or containing 0.1% and 1% v/v HCOOH. Considering the PE, the neutral one gave the best results (see Table 4) and was chosen as elution mixture.

3.5.2 Sample amount

In the final optimized conditions, aliquots of 8 mL of milk samples diluted with phosphate buffer in a sample-to-buffer ratio 1:5 were handled. Indeed, the higher ME observed handling higher sample volume was counterbalanced by improved MLOD and MLOQ values.

3.6 Method performance

Table 5 shows REs, MEs, and PEs obtained analyzing in the optimized conditions 8 mL milk sample fortified at three different levels. REs were >69% and MEs ranged between 67 and 115%, giving overall PEs >52%

Both standard and matrix-matched calibration graphs were linear over the explored range. For all the six analytes, coefficients of determination R^2 of unweighted regression lines were >0.9989 for the standard calibration curve and between 0.9956–0.9996 for the matrix-matched calibration curves (see Supporting Information Fig. S4).

Trueness and precision were estimated by analyzing analyte-free samples spiked before and after MSPE protocol at three different concentration levels; in this case, normalization to IS was used for quantitation. As can be seen in Table 6, apparent recoveries were > 75%. Within laboratory precision was 2–16 and 8–22% for intra-day and inter-day experiments, respectively. Table 6 also reports the MLODs and MLOQs calculated after verification of the extrapolated values. Figure 1 reports the LC–SRM mass chromatograms of the six analytes obtained by the analysis of a blank milk sample and a spiked milk sample.

Compared to other papers published in the literature for the determination of mycoestrogens in milk, the MLOQ values obtained in the present work are lower. In a previous work of our group [13], using classical GCB-SPE, only for ZAN a lower MLOQ was obtained, whereas using an OASIS MAX cartridge, Xia et al. [9] obtained similar values only for ZAN and ZEN. Socas-Rodríguez et al. [39] used multiwalled CNTs for dispersive SPE from milk infant formula, obtaining MLOQs ranging 8–137 μg/L, i.e., three orders of magnitude higher than those obtained in the present work.

4 Concluding remarks

The MSPE technique possesses several advantages if compared to classical SPE. Nevertheless, the problem consists in the choice of the right sorbent for the selected analytes in a certain matrix.

In the present work, an effective MSPE protocol for the extraction of resorcylic acid lactones from milk is presented. To the best of our knowledge, this is the first application of MSPE for mycoestrogens in this matrix. Moreover, also the use of magnetic GCB is still unexplored in this field. Differently from most of the MNPs used in MSPE, possessing a magnetic core grafted with a suitable organic or inorganic layer, in this case GCB was decorated by Fe_3O_4, as already reported in the literature for other carbonaceous materials. In the optimized conditions, overall process efficiency ranged between 52 and 102%, whereas apparent recovery was >93% for all the investigated analytes. The MLOQs ranging between 8 and 15 ng/L are lower than those reported in the few published works on the LC–MS determination of mycoestrogens in milk; moreover, these MLOQ values make the present method able to determine the compounds of interest in milk samples at reasonable contamination levels. Finally, the employment of MSPE instead of SPE allowed to extract a turbid sample as milk avoiding possible clogging of SPE cartridge frit, as well as decreased recoveries due to colloid–analyte interaction. Although the equilibration time (30 min) could appear long, it should be considered that a large number of samples can be processed simultaneously, and the time needed for analyte retrieval is about the same as in the classical SPE.

We would like to thank Dr. Ida Pettiti of Chemistry Department of La Sapienza for porosimetry analysis. We would like to thank Dr. Silvia Marchesan of Trieste University for all the other GCB characterization.

The authors have declared no conflict of interest.

5 References

[1] Metzler, M., Pfeiffer, E., Hildebrand, A., Zearalenone and its metabolites as endocrine disrupting chemicals. *World Mycotoxin J.* 2010, *3*, 385–401.

[2] Flores-Flores, M. E., Lizarraga, E., López de Cerain, A., González-Peñas, E., Presence of mycotoxins in animal milk: a review. *Food Control* 2015, *53*, 163–176.

[3] European Commission, Council Directive 85/649/EEC of 31 December 1985 prohibiting the use in livestock farming of certain substances having a hormonal action. *Off. J.* 1985, *L 382*, 228–231.

[4] Capriotti, A. L., Caruso, G., Cavaliere, C., Foglia, P., Samperi, R., Laganà, A., Multiclass mycotoxin analysis in food, environmental and biological matrices with chromatography/mass spectrometry. *Mass Spectrom. Rev.* 2012, *31*, 466–503.

[5] Capriotti, A. L., Cavaliere, C., Colapicchioni, V., Piovesana, S., Samperi, R., Laganà, A., Analytical strategies based on chromatography-mass spectrometry for the determination of estrogen-mimicking compounds in food. *J. Chromatogr. A* 2013, *1313*, 62–77.

[6] European Commission, Council Directive 96/23/EC of 29 April 1996 on measures to monitor certain substances and residues thereof in live animals and animal products and repealing Directives 85/358/EEC and 86/469/EEC and Decisions 89/187/EEC and 91/664/EEC. *Off. J.* 1996, *L 125*, 10–32.

[7] Malik, A. K., Blasco, C., Picó, Y., Liquid chromatography-mass spectrometry in food safety. *J. Chromatogr. A* 2010, *1217*, 4018–4040.

[8] Socas-Rodríguez, B., Asensio-Ramos, M., Hernández-Borges, J., Herrera-Herrera, A. V., Rodríguez-Delgado, M. Á., Chromatographic analysis of natural and synthetic estrogens in milk and dairy products. *TrAC Trends Anal. Chem.* 2013, *44*, 58–77.

[9] Xia, X., Li, X., Ding, S., Zhang, S., Jiang, H., Li, J., Shen, J., Ultra-high-pressure liquid chromatography–tandem mass spectrometry for the analysis of six resorcylic acid lactones in bovine milk. *J. Chromatogr. A* 2009, *1216*, 2587–2591.

[10] Huang, L. C., Zheng, N., Zheng, B. Q., Wen, F., Cheng, J. B., Han, R. W., Xu, X. M., Li, S. L., Wang, J. Q., Simultaneous determination of aflatoxin M1, ochratoxin A, zearalenone and α-zearalenol in milk by UHPLC–MS/MS *Food Chem.* 2014, *146*, 242–249.

[11] Sørensen, L., Elbak, T., Determination of mycotoxins in bovine milk by liquid chromatography tandem mass spectrometry. *J. Chromatogr. B* 2005, *820*, 183–196.

[12] Kaklamanos, G., Theodoridis, G., Rapid multi-method for the determination of growth promoters in bovine milk by liquid chromatography–tandem mass spectrometry. *J. Chromatogr. B* 2013, *930*, 22–29.

[13] Capriotti, A. L., Cavaliere, C., Piovesana, S., Stampachiacchiere, S., Samperi, R., Ventura, S., Laganà, A., Simultaneous determination of naturally occurring estrogens and mycoestrogens in milk by ultrahigh-performance liquid chromatography-tandem mass spectrometry analysis. *J. Agric. Food Chem.* 2015, *63*, 8940–8946.

[14] Gao, Q., Luo, D., Bai, M., Chen, Z.-W., Feng, Y.-Q., Rapid determination of estrogens in milk samples based on magnetite nanoparticles/polypyrrole magnetic solid-phase extraction coupled with liquid chromatography–tandem mass spectrometry. *J. Agric. Food Chem.* 2011, *59*, 8543–8549.

[15] Yan, W., Li, Y., Zhao, L., Lin, J.-M., Determination of estrogens and bisphenol A in bovine milk by automated on-line C30 solid-phase extraction coupled with high-performance liquid chromatography–mass spectrometry. *J. Chromatogr. A* 2009, *1216*, 7539–7545.

[16] Ríos, A., Zougagh, M., Bouri, M., Magnetic (nano)materials as an useful tool for sample preparation in analytical methods. A review. *Anal. Methods* 2013, *5*, 4558–4573.

[17] Wierucka, M., Biziuk, M., Application of magnetic nanoparticles for magnetic solid-phase extraction in preparing biological, environmental and food samples. *TrAC Trends Anal. Chem.* 2014, *59*, 50–58.

[18] Pérez, R. A., Albero, B., Tadeo, J. L., Molero, E., Sánchez-Brunete, C., Application of magnetic iron oxide nanoparticles for the analysis of PCBs in water and soil leachates by gas chromatography–tandem mass spectrometry. *Anal. Bioanal. Chem.* 2015, *407*, 1913–1924.

[19] Speltini, A., Sturini, M., Maraschi, F., Profumo, A., Recent trends in the application of the newest carbonaceous materials for magnetic solid-phase extraction of environmental pollutants. *Trends Environ. Anal. Chem.* 2016, *10*, 11–23.

[20] Cao, X., Chen, J., Ye, X., Zhang, F., Shen, L., Mo, W., Ultrasound-assisted magnetic SPE based on Fe_3O_4 - grafted graphene for the determination of polychlorinated biphenyls in water samples: sample preparation. *J. Sep. Sci.* 2013, *36*, 3579–3585.

[21] Zhang, Y., Cheng, Y., Chen, N., Zhou, Y., Li, B., Gu, W., Shi, X., Xian, Y., Recyclable removal of bisphenol A from aqueous solution by reduced graphene oxide–magnetic nanoparticles: adsorption and desorption. *J. Colloid Interface Sci.* 2014, *421*, 85–92.

[22] Amoli-Diva, M., Pourghazi, K., Hajjaran, S., Dispersive micro-solid phase extraction using magnetic nanoparticle modified multi-walled carbon nanotubes coupled with surfactant-enhanced spectrofluorimetry for sensitive determination of lomefloxacin and ofloxacin from biological samples. *Mater. Sci. Eng. C* 2016, *60*, 30–36.

[23] Liu, L., Hao, Y., Ren, Y., Wang, C., Wu, Q., Wang, Z., Magnetic nanoporous carbon as an adsorbent for the extraction of phthalate esters in environmental water and aloe juice samples: sample preparation. *J. Sep. Sci.* 2015, *38*, 1411–1418.

[24] Diao, C., Yang, X., Sun, A., Liu, R., A combined technique for the pretreatment of ultra trace bisphenol A in environmental water based on magnetic matrix solid phase extraction assisted dispersive liquid–liquid microextraction. *Anal. Methods* 2015, *7*, 10170–10176.

[25] Sarafraz-Yazdi, A., Rokhian, T., Amiri, A., Ghaemi, F., Carbon nanofibers decorated with magnetic nanoparticles as a new sorbent for the magnetic solid phase extraction of selected polycyclic aromatic hydrocarbons from water samples. *New J. Chem.* 2015, *39*, 5621–5627.

[26] Ding, J., Gao, Q., Li, X.-S., Huang, W., Shi, Z.-G., Feng, Y.-Q., Magnetic solid-phase extraction based on magnetic carbon nanotube for the determination of estrogens in milk. *J. Sep. Sci.* 2011, *34*, 2498–2504.

[27] Wang, J., Cheng, C., Yang, Y., Determination of estrogens in milk samples by magnetic-solid-phase extraction technique coupled with high-performance liquid chromatography. *J. Food Sci.* 2015, *80*, C2655–C2661.

[28] Capriotti, A. L., Cavaliere, C., La Barbera, G., Piovesana, S., Samperi, R., Zenezini Chiozzi, R., Laganà, A., Polydopamine-coated magnetic nanoparticles for isolation and enrichment of estrogenic compounds from surface water samples followed by liquid chromatography-tandem mass spectrometry determination. *Anal. Bioanal. Chem.* 2016, *408*, 4011–4020.

[29] Yan, Y., Zheng, Z., Deng, C., Zhang, X., Yang, P., Selective enrichment of phosphopeptides by titania nanoparticles coated magnetic carbon nanotubes. *Talanta* 2014, *118*, 14–20.

[30] Matuszewski, B. K., Constanzer, M. L., Chavez-Eng, C. M., Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC–MS/MS *Anal. Chem.* 2003, *75*, 3019–3030.

[31] Belen Serrano, A., Capriotti, A. L., Cavaliere, C., Piovesana, S., Samperi, R., Ventura, S., Laganà, A., Development of a rapid LC-MS/MS method for the determination of emerging *Fusarium* mycotoxins enniatins and beauvericin in human biological fluids. *Toxins* 2015, *7*, 3554–3571.

[32] Cavaliere, C., Capriotti, A. L., Foglia, P., Piovesana, S., Samperi, R., Ventura, S., Laganà, A., Natural estrogens in dairy products: determination of free and conjugated forms by ultra high performance liquid chromatography with tandem mass spectrometry. *J. Sep. Sci.* 2015, *38*, 3599–3606.

[33] Wang, X., Liu, B., Lu, Q., Qu, Q., Graphene-based materials: fabrication and application for adsorption in analytical chemistry. *J. Chromatogr. A* 2014, *1362*, 1–15.

[34] Herrero-Latorre, C., Barciela-García, J., García-Martín, S., Peña-Crecente, R. M., Otárola-Jiménez, J., Magnetic solid-phase extraction using carbon nanotubes as sorbents: a review. *Anal. Chim. Acta* 2015, *892*, 10–26.

[35] Crescenzi, C., Di Corcia, A., Passariello, G., Samperi, R., Carou, M. I. T., Evaluation of two new examples of graphitized carbon blacks for use in solid-phase extraction cartridges. *J. Chromatogr. A* 1996, *733*, 41–55.

[36] Altenbach, B., Giger, W., Determination of benzene- and naphthalenesulfonates in wastewater by solid-phase extraction with graphitized carbon black and ion-pair liquid chromatography with UV detection. *Anal. Chem.* 1995, *67*, 2325–2333.

[37] Capriotti, A. L., Cavaliere, C., Foglia, P., Samperi, R., Stampachiacchiere, S., Ventura, S., Laganà, A., Ultra-high-performance liquid chromatography-tandem mass spectrometry for the analysis of free and conjugated natural estrogens in cow milk without deconjugation. *Anal. Bioanal. Chem.* 2015, *407*, 1705–1719.

[38] Chen, X.-H., Pan, S.-D., Ye, M.-J., Li, X.-P., Zhao, Y.-G., Jin, M.-C., Magnetic solid-phase extraction based on a triethylenetetramine-functionalized magnetic graphene oxide composite for the detection of ten trace phenolic environmental estrogens in environmental water: Sample Preparation. *J. Sep. Sci.* 2016, *39*, 762–768.

[39] Socas-Rodríguez, B., González-Sálamo, J., Hernández-Borges, J., Rodríguez Delgado, M. Á., Application of multiwalled carbon nanotubes as sorbents for the extraction of mycotoxins in water samples and infant milk formula prior to high performance liquid chromatography mass spectrometry analysis: liquid phase separations. *Electrophoresis* 2016, *37*, 1359–1366.

3 Paper II:

A rapid magnetic solid phase extraction method followed by liquid Chromatography-Tandem mass spectrometry analysis for the determination of mycotoxins in cereals

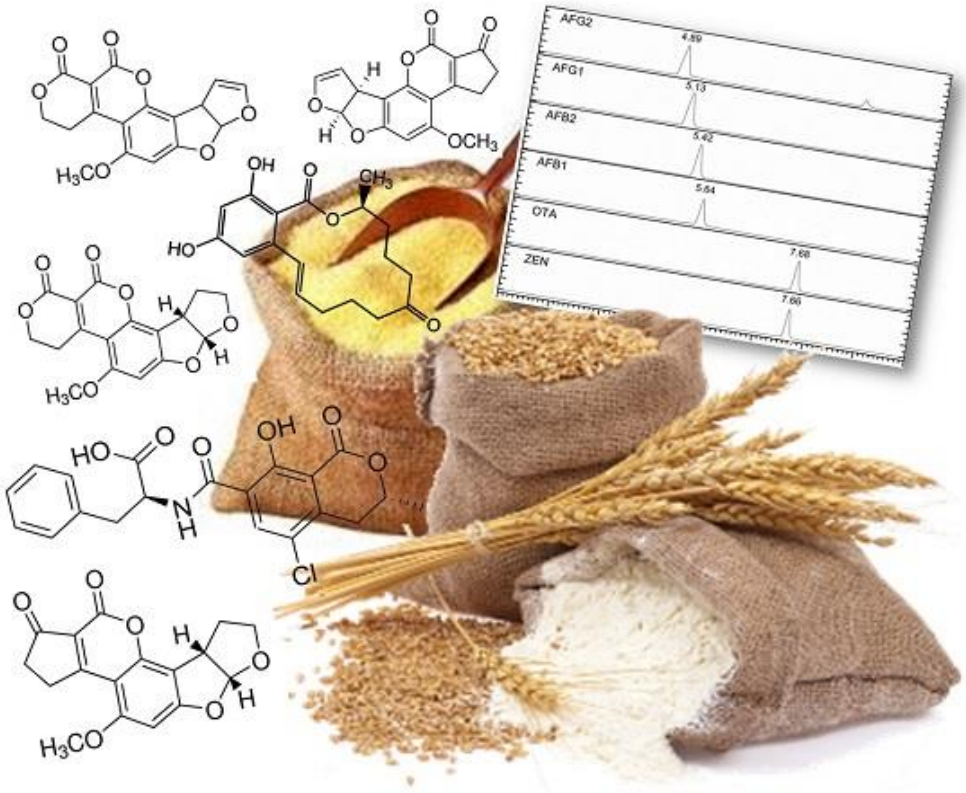


Figure 3.2: Graphical abstract of Paper II

toxins

MDPI

Article

A Rapid Magnetic Solid Phase Extraction Method Followed by Liquid Chromatography-Tandem Mass Spectrometry Analysis for the Determination of Mycotoxins in Cereals

Giorgia La Barbera, Anna Laura Capriotti, Chiara Cavaliere, Patrizia Foglia, Carmela Maria Montone, Riccardo Zenezini Chiozzi * and Aldo Laganà

Department of Chemistry, University of Rome "La Sapienza", 00185 Roma, Italy; giorgia.labarbera@uniroma1.it (G.L.B.); annalaura.capriotti@uniroma1.it (A.L.C.); chiara.cavaliere@uniroma1.it (C.C.); patrizia.foglia@uniroma1.it (P.F.); carmelamaria.montone@uniroma1.it (C.M.M.); aldo.lagana@uniroma1.it (A.L.)

* Correspondence: riccardo.zenezini@uniroma1.it

Academic Editor: Laura Anfossi

Received: 28 February 2017; Accepted: 19 April 2017; Published: 21 April 2017

Abstract: Mycotoxins can contaminate various food commodities, including cereals. Moreover, mycotoxins of different classes can co-contaminate food, increasing human health risk. Several analytical methods have been published in the literature dealing with mycotoxins determination in cereals. Nevertheless, in the present work, the aim was to propose an easy and effective system for the extraction of six of the main mycotoxins from corn meal and durum wheat flour, i.e., the main four aflatoxins, ochratoxin A, and the mycoestrogen zearalenone. The developed method exploited magnetic solid phase extraction (SPE), a technique that is attracting an increasing interest as an alternative to classical SPE. Therefore, the use of magnetic graphitized carbon black as a suitable extracting material was tested. The same magnetic material proved to be effective in the extraction of mycoestrogens from milk, but has never been applied to complex matrices as cereals. Ultra high–performance liquid chromatography tandem mass spectrometry was used for detection. Recoveries were >60% in both cereals, even if the matrix effects were not negligible. The limits of quantification of the method results were comparable to those obtained by other two magnetic SPE-based methods applied to cereals, which were limited to one or two mycotoxins, whereas in this work the investigated mycotoxins belonged to three different chemical classes.

Keywords: mycotoxins; aflatoxins; ochratoxin A; zearalenone; magnetic solid phase extraction; graphitized carbon black; liquid chromatography-tandem mass spectrometry; cereals; wheat; maize

1. Introduction

Mycotoxins are secondary metabolites produced by various filamentous fungi, mainly species of *Aspergillus*, *Fusarium*, and *Penicillium*, but also *Claviceps* and *Alternaria* [1]. These molds may grow under a wide range of climatic conditions on several agricultural commodities, including cereals, oleaginous seeds, spices, and coffee, both pre- and post-harvest (e.g., during storage) [2]. Some fungi produce a single mycotoxin, whereas others may produce many toxic compounds, which may be shared across fungal genera.

There are about 400 known mycotoxins that exhibit a great structural diversity [2]. However, only a few of them are considered to be of agricultural importance [3]. Indeed, the main three genera of fungi, *Aspergillus*, *Fusarium*, and *Penicillium*, produce mycotoxins belonging to five relevant groups to the food industry: aflatoxins (AFs) produced by *Aspergillus* species, ochratoxin A (OTA) produced

by both *Aspergillus* and *Penicillium* species, and fumonisins—specifically trichothecenes and resorcyclic lactones (zearalenones)—all produced mainly by *Fusarium* species [4–6].

Human exposure to mycotoxins occurs mostly through the intake of contaminated food and beverages, and to a minor extent through dermal contact and inhalation [1]. Mycotoxin occurrence in food is due to direct contamination of plant materials or products, or due to carry over of mycotoxins and their metabolites in foods (e.g., meat, milk, eggs) obtained from animals fed with contaminated feed [2].

The consumption of mycotoxin-contaminated food rarely determines acute toxicity [7]; nevertheless, a wide range of adverse effects for human and animal health, including carcinogenic, mutagenic, estrogenic, and immunosuppressive effects, has been demonstrated [8,9]. According to the International Agency for Research on Cancer system of classification, AFs are carcinogenic to humans (group 1) [5,10], whereas OTA is possibly carcinogenic to humans (group 2B), and zearalenone (ZEN) is not carcinogenic to humans (group 3) [5]. However, ZEN is well known for its estrogenic effect. Furthermore, OTA exposure has been related to nephropathies and other adverse health effects [11]. Finally, various mycotoxins might co-contaminate food, with possible detrimental additive and/or synergic effects on human and animal health [12].

In order to limit mycotoxin exposure and protect consumer and animal health from adverse effects, many countries have adopted regulations and maximum admissible levels (MLs) for the most prevalent and hazardous mycotoxins in certain commodities that are more prone to fungal proliferation [6,7]. Mycotoxins also have a negative impact on world trade: according to the annual report of the Rapid Alert System for Food and Feed [13], mycotoxins were the main hazard category for border rejection notifications in the European Union (EU) in 2015. Most of the notifications on mycotoxins in food were related to the presence of AFs (421/475 notifications), with a significant increase compared to 2014, whereas 42 notifications were due to OTA occurrence and the others (mainly to fusariotoxins).

Cereals, especially maize and bakery products, are one of the commercial categories frequently affected by mycotoxin presence [6]. The EU has fixed MLs for some mycotoxins in cereals and derived products, namely 2 and 4 μg kg^{-1} for AFB1 and the sum of the 4 AFs (AFB1, AFB2, AFG1, and AFG2), respectively, and 3 μg kg^{-1} for OTA. For ZEN, an ML of 75 μg kg^{-1} has been set for cereals and cereal flour, with the exception of maize, for which the ML is 100 μg kg^{-1} [14].

Methods for mycotoxin determination may roughly be classified as chromatographic-based, immunological-based, and sensor-based [3]. Chromatographic methods are generally used for confirmation purposes, whereas the other two method categories are often employed for screening analysis. Liquid chromatography, including ultra-high performance liquid chromatography (UHPLC), is generally preferred to gas chromatography for its versatility and is now considered the standard separation technique for mycotoxin analysis [3]. For detection, mass spectrometry (MS) and fluorescence (FD) are the gold standard against which all other methods are compared. In particular, tandem mass spectrometry (MS/MS) is the technique of choice for most authors [1].

Cereals are complex food matrices. Therefore, sample pre-treatment and/or clean-up and enrichment steps are generally required for most chromatographic methods. Although analysis time and cost are increased in this case, the analytical method benefits from increased sensitivity and robustness (e.g., reducing column blockage and contamination) [3]. After a preliminary solvent extraction from the solid matrix (generally with a mixture of acetonitrile or methanol and water), the mycotoxin extract is generally cleaned up/enriched by solid phase extraction (SPE). A wide variety of solid phases has been used, from the common C18-siclica bonded materials to the specific immunoadsorbent materials [3]. Recently, dispersive SPE, in a magnetic mode using nanoparticles, is attracting increasing scientific interest [15,16]. The mechanisms occurring in magnetic SPE (mSPE) are analogous to those observed in classical on-column SPE, where the interactions between target molecules and adsorbent functional groups determine the efficiency of the system. Certainly, the matrix composition also affects the selection of the best combination adsorbent—elution mixture [16]. In mSPE, the dispersion of the magnetic material into the solution containing the target molecules assures

a continuous and dynamic contact with the adsorbent surface, leading to a more efficient analyte retention. The separation of the magnetic material with the adsorbed analytes from the solution is then realized by applying a magnet outside the vessel (e.g., on the bottom), avoiding centrifugation or filtration steps. Finally, after eventual washings, analytes are eluted from the magnetic material by a proper solvent mixture.

The mSPE technique followed by HPLC-FD analysis was already employed to extract mycotoxins from cereals. Magnetic nanoparticles coated by 3-(trimethoxysilyl)-1-propanethiol and different functionalizations were used to extract OTA from cereals [17] and AFB1 and AFB2 [18] from corn and rice samples.

In a previous study [19], the capability of magnetic graphitized carbon black (mGCB) in mycoestrogen extraction from milk samples was successfully tested. To this study's authors' best knowledge, that was the first application of mGCB in milk. In the present work, the same magnetic adsorbing material was employed to extract the main and most dangerous mycotoxins (namely AFB1, AFB2, AFG1, AFG2, OTA, and ZEN) from corn (*Zea mays*) meal and durum wheat (*Triticum durum*) flour. The extract was then analyzed by UHPLC-MS/MS with an electrospray (ESI) source. The method was suitably modified and validated in the new complex matrices. It was rapid and provided satisfying process efficiency (PE) and suitable limits of quantification (LOQs).

2. Results and Discussion

2.1. The Magnetic GCB Adsorbent Material

mGCB was chosen based on previous work [19], where its suitability for the extraction of mycoestrogens from milk was demonstrated. Moreover, in the past, GCB was also used in classical SPE mode to extract ZEN [20] and the four AFs [21] from maize.

Generally, most of the carbon-based materials used in aqueous environment need an oxidation step to improve their wettability and limit aggregation phenomena [22,23]. GCB is easily dispersed in water because of the presence of polar heterogeneities in its structure; nevertheless, a mild oxidation (1%, *w*/*w*) helped in reducing the elution solvent volume [19]. However, when treated in stronger oxidizing conditions, up to obtaining a 10% (*w*/*w*) oxidation, GCB lost most of its retention capability (data not shown).

The GCB characteristics before and after magnetization, as well as batch-to-batch preparation reproducibility, were previously assessed [19].

2.2. Samples

Most steps of the analytical method were developed using corn meal samples, and only successively validated for wheat flour samples. These two cereal flours contain nearly the same lipid amount, whereas protein and carbohydrate contents are significantly different (see Supplementary Table S1). Therefore, s-lensthe ESI matrix effects (MEs) on the analytes could also differ. Furthermore, the maize plant is very susceptible to contamination by *Fusarium* species [24], so it is generally more affected by certain mycotoxin contamination than wheat (in particular ZEN contamination) [6]. Indeed, during method development, the problem of finding blank samples emerged, since large amounts of ZEN were detected in most corn meal samples, thus leading to recovery (RE) and ME overestimate (up to 300%). Therefore, before matrix spiking, blank analysis of every new sample batch was performed to verify the absence of the investigated mycotoxins. Moreover, for each new sample set, a blank sample was randomly injected to verify the absence of possible carryover.

2.3. Sample Preparation

Generally, acetonitrile/water mixtures in the *v*/*v* ratios ranging from 75:25 up to 85:15 are used to extract mycotoxins from cereal samples; in some cases, acetic or formic acid up to 1% (*v*/*v*) is added to the mixture [1]. In the present work, before the mSPE procedure, corn meal and durum wheat flour samples were extracted with acetonitrile/water/formic acid 80:19.8:0.2 (*v*/*v*/*v*). In preliminary

experiments, a neutral mixture and a mixture containing 1% formic acid were employed. However, high acid amounts had a detrimental effect on AF REs, whereas neutral mixture caused a slight decrease in the REs of OTA and ZEN (data not shown).

Different sample to mGCB ratios were tested, to obtain the highest overall PE values, i.e., taking into account both RE and ME. Keeping the magnetic material amount to 50 mg, experiments (three replicates for each condition) using 1000 mg, 500 mg, and 250 mg of maize meal were performed. In all three cases, the spiking level was 5 ng g^{-1} for the four AFs and OTA and 250 ng g^{-1} for ZEN. Results, which are reported in Supplementary Table S2, showed that the smallest sample amount allowed obtaining the best PE values for most mycotoxins. Moreover, with the lowest sample to adsorbent ratio, signal suppression for OTA was significantly reduced; the ME led to a moderate signal enhancement only for ZEN.

Five mL of dichloromethane/methanol 80:20 (*v*/*v*) containing 0.2% formic acid was used to elute the retained analytes from mGCB. A mixture dichloromethane/methanol, either neutral or containing up to 0.2% of a weak acid, was generally used to elute the analytes from GCB in classical SPE mode. In the analysis of maize samples, ZEN and its derivatives were eluted from GCB using a neutral dichloromethane/methanol 80:20 (*v*/*v*) mixture [25], while AFs were extracted with the same mixture containing acid [21]. In this case, the elution conditions were chosen to enhance the REs of the AFs, which are the natural least abundant and most toxic mycotoxins among the selected ones. Furthermore, the presence of acid in the elution mixture could also enhance the REs of compounds, such as OTA, establishing electrostatic interactions with the GCB surface [26].

At the end of the sample preparation procedure, the solvents were removed by evaporation from the eluate, and the residue was reconstituted with 250 µL of methanol/water 80:20 (*v*/*v*) containing 5 mmol L^{-1} ammonium formate. The choice of the reconstitution mixture appeared to be a critical point. Initially, a mixture of acetonitrile/water 50:50 (*v*/*v*) containing ammonium formate was used. However, REs were low for most of the analytes (see Table 1), very likely due to solubility issues. A larger amount of acetonitrile did not significantly improve such results, since the little RE increase was associated with worse MEs. Finally, the replacement of acetonitrile with methanol gave the best results, also in terms of chromatographic peak broadening.

Table 1. Recovery (RE, %) and matrix effect (ME, %) obtained from the extraction of corn meal samples spiked with 0.5 µg kg^{-1} of each AF, 1.5 µg kg^{-1} of ochratoxin A (OTA) and 375 µg kg^{-1} of zearalenone (ZEN). The residue was reconstituted with three different mixtures containing 5 mmol L^{-1} ammonium formate: (I) acetonitrile/water 50:50 (*v*/*v*); (II) acetonitrile/water 80:20 (*v*/*v* II); methanol/water 80:20 (*v*/*v*).

Analyte	ACN/H_2O 50:50		ACN/H_2O 80:20		MeOH/H_2O 80:20	
	RE	ME	RE	ME	RE	ME
AFG2	44	78	55	68	61	72
AFG1	53	74	50	73	61	70
AFB2	49	77	62	66	53	76
AFB1	45	73	67	56	64	65
OTA	61	63	64	68	56	66
ZEN	83	86	78	82	81	89

2.4. Liquid Chromatography-Tandem Mass Spectrometry Analysis

Analytes were separated onto a reversed-phase C18 chromatographic column. In the first experiments, a Thermo Fisher Hypersil Gold C18 column (50 × 2.1 mm i.d., 1.9 µm particle size) was used, and different elution gradients were tested to reduce signal suppression due to MEs. However, in corn samples, OTA ME was ca. 50% in all the chromatographic conditions (data not shown), even reducing the gradient rate or eluting the analyte in isocratic conditions. Such large ME was attributed to the effect of some coextracted lipids eluting at retention times very close to that of OTA. Indeed, in previous works on mycotoxin determination in maize meal [25,27], the extract was filtered through a C18 cartridge before GCB SPE clean-up to retain phospholipids and most triglycerides. In the present

work, lower MEs were obtained only by using a longer column, namely a 100 mm Cortecs UPLC C18+ column by Waters (see Supplementary Table S3). Figure 1 shows the extracted ion chromatograms relative to the six investigated analytes in a corn meal extract, whereas Figure S1 shows the extracted ion chromatograms relative to the six investigated analytes in a wheat flour extract.

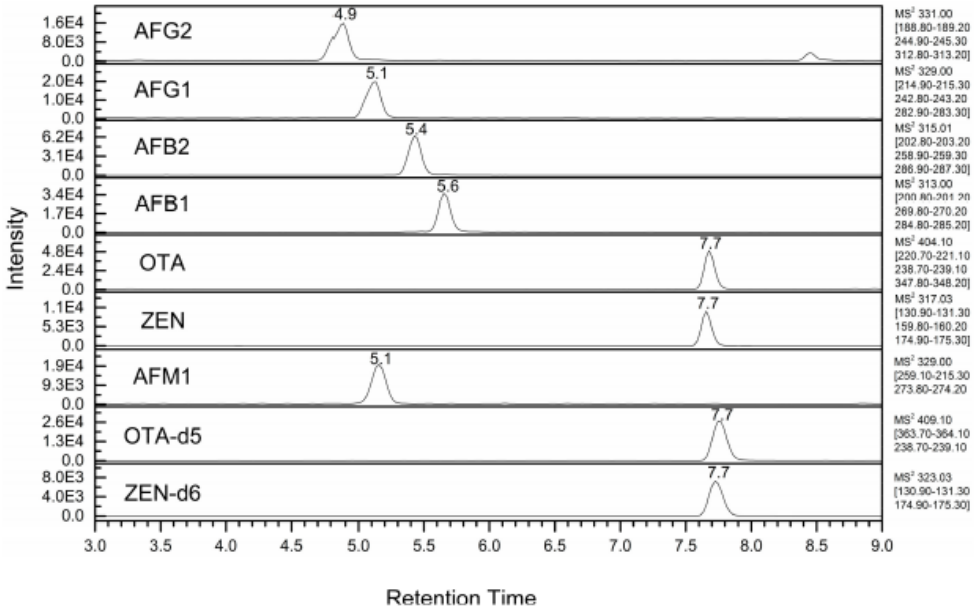


Figure 1. Extracted ion chromatograms (sum of three transition pairs for each analyte) of a corn meal sample spiked with the analytes at 0.5 × ML.

Analyte-specific acquisition parameters (e.g., S-lens, precursor ions, MRM transitions and collision energy) were optimized by directly infusing 1 ng μL^{-1} individual mycotoxin standard solution prepared in water/methanol 50:50 (*v*/*v*) at 10 µL min^{-1} (see Table 2). General ESI source tune parameters (i.e., spray voltage, gas pressures, and source temperatures) were optimized by simultaneously introducing, through a tee-junction, the same 1 ng μL^{-1} individual mycotoxin standard solution at 10 µL min^{-1} and water/methanol 50:50 (*v*/*v*) containing 5 mmol L^{-1} ammonium formate and 0.1% formic acid at 300 µL min^{-1} flow-rate. The chosen parameters were a compromise between the optimal ones determined for each analyte.

Table 2. Mycotoxin detection parameters.

Mycotoxin	Retention Time (Min)	Precursor Ion (*m*/*z*)	Product Ion [1] (*m*/*z*) (CE [2], eV)	S-Lens (V)
Positive polarity		$[M+H]^+$		
AFG2	4.9	331	189 (42), 245 (31), **313** (25)	145
AFM1 (IS)	5.1	329	259 (34), **273** (25)	145
AFG1	5.4	329	215 (33), **243** (27), 283 (24)	145
AFB2	5.6	315	203 (36), 259 (33), **287** (24)	155
AFB1	4.6	313	201 (30), 270 (26), **285** (23)	150
OTA-d5 (IS)	7.7	409	363 (25), **239** (27)	
OTA	7.7	404	221 (36), **239** (26), 348 (13)	110
Negative polarity		$[M-H]^-$		
ZEN-d6 (IS)	7.7	323	131 (33), **175** (25)	140
ZEN	7.7	317	131 (33), 160 (34), **175** (25)	150

[1] In bold, the most intense transition; [2] CE, collision energy.

For each analyte, the unambiguous identification was based on comparison with the authentic standard (retention times, relative intensity ratios of MRM transition pairs) following the criteria reported in the Decision 2002/657/EC [28].

2.5. Reuse of MGCB

As in a previous work on magnetic nanoparticles coated with polydopamine [29], the reuse of mGCB was evaluated, but it was not convenient due to the large amount of solvent required for mGCB effective washing and the loss of material during such operation.

2.6. Method Performance

The MLs fixed by EU for AFs in cereals are 2 μg kg^{-1} for the most dangerous and widespread AFB1 and 4 μg kg^{-1} for the sum of AFB1, AFB2, AFG1 and AFG2. For method development, the worst scenario was prospected, i.e., each of the four AFs at 1 μg kg^{-1} concentration.

For laboratory method validation, RE and ME were determined at three spiking levels in both corn meal and durum wheat flour samples, according to Equations (1) and (2) (see Table 3). The product of RE and ME provides the overall PE (Equation (3)). The RE relative standard deviations were below 17% for all the analytes.

Table 3. Recovery (RE, %, $n = 6$) and matrix effect (ME, %) for AFs, OTA and ZEN in corn meal and durum wheat flour samples. Fortification levels were maximum limit (ML, i.e., 1 μg kg^{-1} for each AF, 3 μg kg^{-1} for OTA and 750 μg kg^{-1} for ZEN), 0.5 × ML and 2 × ML.

Analyte	0.5 × ML				ML				2 × ML			
	Corn		Durum Wheat		Corn		Durum Wheat		Corn		Durum Wheat	
	RE	ME	RE	ME	RE	ME	RE	ME	RE	ME	RE	ME
AFG2	78	76	74	86	67	79	69	79	71	84	67	88
AFG1	71	75	74	79	66	73	74	75	68	78	73	84
AFB2	69	76	68	78	63	85	76	86	74	89	71	76
AFB1	73	68	71	70	74	69	73	68	72	69	69	73
OTA	67	72	73	67	83	68	76	71	79	69	81	68
ZEN	78	87	79	102	89	94	82	116	84	104	88	108

Recoveries were >67% at the lowest fortification level. The MEs of signal suppression affected in particular AFB1 and OTA. The only signal enhancement ME was observed for ZEN, however it cannot be excluded that this signal enhancement was due to a natural contamination below method detection limit (MLOD).

The trueness of the method was assessed by means of apparent REs (Equation (4)) at three different fortification levels, whereas intra-day and inter-day laboratory precision were determined by performing recovery experiments ($n = 6$) at 0.5 × ML in the same day and in six consecutive days, respectively (see Table 4). Apparent REs were higher for the two mycotoxins whose deuterated ISs were available, i.e., OTA and ZEN, whereas for the four AFs, the hydroxylated Phase I AFB1 metabolite, namely AFM1, was used as IS.

The equations and coefficient of determination (R^2) obtained for the standard and the two matrix-matched calibration graphs are reported in Supplementary Table S4.

The MLODs and MLOQs were determined as described in the Experimental section (see Table 5), since by operating in MRM mode with the last generation triple quadrupole mass spectrometers, it is quite common to obtain MRM signals without noise [8]. The mSPE technique was used to extract AFB1, AFB2 [18] and OTA [17] from cereals. Compared to the work by Hashemi et al. [18], which obtained MLOQs around 0.2 and 0.05 μg kg^{-1} for AFB1 and AFB2, respectively, the MLOQs obtained in the present work for AFB2 are slightly higher. Compared to the work by Mashhadizadeh et al. [17], the limits for OTA are three times higher. However, it should be considered that both these works used

HPLC-FD for determination, thus the criteria for MLOD and MLOQ estimation are different, and they analyzed from one up to two mycotoxins.

Table 4. Trueness, intra-day and inter-day laboratory precision obtained by analyzing corn meal and durum wheat flour samples spiked with aflatoxins (AFs), OTA, and ZEN at maximum limit (ML, i.e., 1 μg kg^{-1} for each AF, 3 μg kg^{-1} for OTA and 750 μg kg^{-1} for ZEN), 0.5 × ML and 2 × ML. Results are averaged from $n = 6$, performed in the same day and in six consecutive days.

Analyte	Trueness						Precision (0.5 × ML)			
	0.5 × ML		ML		2 × ML		Intra-Day		Inter-Day	
	Corn	Wheat	Corn	Wheat	Corn	Wheat	Corn	Wheat	Corn	Wheat
AFG2	94	98	95	96	92	91	9	10	17	14
AFG1	93	99	97	95	98	96	8	12	7	9
AFB2	89	91	94	89	103	99	6	3	10	8
AFB1	90	93	94	90	106	101	10	7	8	10
OTA	99	97	98	100	96	97	7	4	11	9
ZEN	103	99	101	98	104	106	11	3	16	20

Table 5. Method limits of detection (MLODs) and quantification (MLOQs) estimated (est.) according to Equations (5) and (6) and confirmed (conf.) according to Equations (7) and (8).

Analyte	MLODs (μg kg^{-1})				MLOQs (μg kg^{-1})			
	Corn		Durum Wheat		Corn		Durum Wheat	
	Est.	Conf.	Est.	Conf.	Est.	Conf.	Est.	Conf.
AFG2	0.11	0.05	0.12	0. 05	0.38	0.20	0.43	0.15
AFG1	0.08	0.10	0.13	0.05	0.27	0.20	0.23	0.15
AFB2	0.09	0.05	0.14	0.05	0.29	0.10	0.43	0.10
AFB1	0.11	0.10	0.23	0.10	0.36	0.10	0.23	0.10
OTA	0.48	0.10	0.25	0.20	1.60	0.30	0.79	0.30
ZEN	10.2	1.0	4.2	2.2	33.8	1.0	34.6	2.2

Moreno et al. [30] used magnetic nanoparticles coated with a layer of octadecyl group-modified silica containing multiwalled carbon nanotubes to extract ZEN and its derivative from maize, followed by LC-MS analysis. They used only 5 mg of magnetic nanoparticles; however, a comparison with this method is difficult, since it is not clear if 6 g or 10 g sample aliquot was used in the sample preparation procedure. Moreover, LODs and LOQs are provided in μg L^{-1}, thus the concentration is not clearly referred to maize amount. Ethylene glycol bis-mercaptoacetate modified silica coated magnetic nanoparticles were also used to extract AFs from wheat before spectrofluorometric analysis [31]; the MLODs (0.07 μg kg^{-1}) were comparable to those obtained in the present work, whereas the MLOQs (0.24 μg kg^{-1}) were higher.

2.7. Sample Analysis

One of the problems arising during method development was to find samples free of ZEN, which, indeed, contaminates most of the corn meal samples [24,25]. OTA was also detected in some samples, though at concentrations below MLOQ. In a short survey carried out on 10 corn meal samples, OTA was detected in one sample at 1.3 μg kg^{-1} level, whereas it was detected at values >MLOD but <MLOQ in the other two samples. In eight out of 10 samples, ZEN was detected at values >MLOD but <MLOQ, and in one sample at 72.9 μg kg^{-1} level. Quantification was made by matrix-matched calibration (results are shown in Table S5). None of the investigated mycotoxins were detected in the five durum wheat flour samples.

3. Conclusions

In this work, the suitability of mGCB for the extraction of AFB1, AFB2, AFG1, AFG2, OTA, and ZEN from corn meal and durum wheat flour samples was demonstrated. The overall process efficiency of the developed method was a compromise between performance and easiness and rapidity of application. Compared with a classic on-column SPE, the time required for a single extraction was about the same. Nevertheless, mSPE was less labor intensive, and more than ten extractions can be managed simultaneously.

Even if in the present work, the limits of quantification were comparable to or higher than those of other mSPE methods. However, this method allows for the simultaneous investigation of a larger number of mycotoxins. Moreover, due to the different detection technique (fluorescence in the other works and tandem mass spectrometry in the present one), MLOQ calculation modes are very different.

This is the first application of mGCB in an mSPE procedure for extraction of mycotoxins from cereals. The potentiality of this material has been exploited before for the extraction of mycotoxins belonging to the same chemical class, but from milk.

4. Materials and Methods

4.1. Chemicals and Reagents

Organic solvents of analytical grade, formic acid, nitric acid, ammonium formate, hydrochloric acid (reagent grade), iron (III) chloride hexahydrate, ethylene glycol, trisodium citrate, sodium acetate, poly(ethyleneglycol)-10k, and GCB (Supelclean ENVI-Carb, surface area: 100 m^2/g, particle size: 120/400 mesh) were obtained from Sigma-Aldrich (St. Louis, MO, USA). LC-MS grade methanol and ultrapure water (resistivity 18.2 MΩ cm) were obtained from Sigma Aldrich and used for LC mobile phase.

Pure (purity ≥98%, unless differently specified) standards of the analytes AFB1, AFB2, AFG1, AFG2, OTA, and ZEN (purity ≥99%) were purchased from Sigma-Aldrich. The standards of aflatoxin M1 (AFM1) and deuterated OTA (OTA-d5) acquired from Sigma-Aldrich, and deuterated ZEN (ZEN-d6) acquired from Wellington Laboratories (Toronto, ON, Canada) were used as internal standards (ISs).

Individual stock standard solutions of the analytes were prepared at 200 ng μL^{-1} in methanol. A composite working standard solution of the six analytes was prepared in methanol at 10 pg μL^{-1} for AFB1, AFB2, AFG1, and AFG2, 30 pg μL^{-1} for OTA, and 750 pg μL^{-1} for ZEN. This mixture was renewed every two weeks. All the solutions were stored in the dark at −20 °C and brought to room temperature before use.

Safety Considerations

AFs and OTA are carcinogenic and possibly carcinogenic compounds to humans, respectively. Therefore, handling standard solutions and extracts requires extreme care. Gloves and other protective clothing were worn as a safety precaution during the handling of the mycotoxins. Solid AF standards were handled in a glove box. Glassware used for standards or samples was soaked in 3% aqueous sodium hypochlorite to destroy mycotoxin residue before cleaning and re-use. When possible, disposable plastic material was used. To avoid degradation, mycotoxins were protected from daylight during sample preparation and the standard solutions were kept in amber vials.

4.2. Magnetic Graphitized Carbon Black Preparation

As in a previous work [19], mGCB was prepared by adapting a literature protocol for carbon nanotubes magnetization [32]. Briefly, 0.40 g of GCB was first added to 50 mL of concentrated nitric acid under agitation for 7 h at room temperature. After that, the material was washed with distilled water until the discarded water reached neutral pH, and then dried overnight at 50 °C. For magnetization, 0.30 g of GCB was dispersed into 80 mL of ethylene glycol and added with 1.62 g of iron (III) chloride

hexahydrate, 0.30 g of trisodium citrate, 7.20 g of sodium acetate, and 2.00 g of poly(ethylene glycol). After 3 h sonication (with a 37 kHz Elmasonic S 60 H by Elma, Singen, Germany), the obtained mixture was sealed in an autoclave for 10 h at 200 °C. The resulting mGCB microparticles were allowed to cool at room temperature for 3 h, then washed with 30 mL ethanol followed by 30 mL distilled water six times. Finally, the mGCB was dried in an oven at 80 °C for 3 h; after cooling, it was stored in a glass flask at room temperature until use.

4.3. Characterization of Graphitized Carbon Black Material

Transmission electron microscopy, Fourier transform infrared spectroscopy spectra, thermogravimetric analysis, porosimetry, and specific surface area analysis were used to characterize the GCB material at each preparation step, i.e., before and after treatment with nitric acid, and after magnetization [19].

4.4. Samples and Extraction Protocol

Samples of *Triticum durum* flour and *Zea mays* meal were obtained from local markets of the Lazio region (Italy).

Sample aliquot of 250 mg was placed in a 15 mL-polypropylene centrifuge tube and added with 2.5 mL of the extracting mixture constituted by acetonitrile/water/formic acid 80:19.8:0.2 (*v*/*v*/*v*). The tube was vortexed (with a Digital Vortex-Genie 2 by Scientific Industries, Bohemia, NY, USA) at 2000 rpm for 3 min, then placed in an ultrasonic bath for 10 min, and finally centrifuged at 12,500× *g* for 15 min at 4 °C. After that, the supernatant was transferred to a 50 mL-polypropylene centrifuge tube containing 50 mg of mGCB previously conditioned. The conditioning procedure consisted in adding to the weighted material 5 mL of methanol followed by 5 mL of ultrapure water. Each time, after vortexing for 30 s, the material was allowed to settle down by magnetic decantation (using a permanent magnetic disk Nd-Fe-B, 25 mm × 5 mm, by Supermagnete, Gottmadingen, Germany), and the solvent was removed.

The mixture extract-mGCB was diluted with 22.5 mL ultrapure water and vortexed at 900 rpm for 30 min to promote the analyte adsorption onto the magnetic microparticles. At this point, as well as during the following protocol steps, the solvent was removed by magnetic decantation with a permanent magnetic disk placed on the bottom of the tube; the mGCB was washed with 4 mL of ultrapure water by vortexing for 30 s at 2800 rpm. Water was removed and 250 µL of methanol was added to eliminate residual water. After 30 s-manual shaking, methanol was removed. Mycotoxins were eluted from mGCB by 5 mL of dichloromethane/methanol 80:20 (*v*/*v*) containing 0.2% formic acid under vortexing for 3 min at 2000 rpm. The supernatant was collected in a clean 20 mL round bottom glass vial and the solvent was removed by a gentle nitrogen stream in a water bath at 37 °C. When required, the three ISs were added to the extract before solvent removal. Finally, the residue was reconstituted with 250 µL of methanol/water 80:20 (*v*/*v*) containing 5 mmol L^{-1} ammonium formate; after 30 s vortexing, the solution was placed in an ultrasonic bath for 10 s (3 times). Before transferring the extract in an autosampler vial, it was centrifuged (by a MicroCL 21R centrifuge, Thermo Scientific, Waltham, MA, USA) at 21,000× *g* for 5 min to eventually sediment suspended particles and/or remove residual fats.

To artificially fortify mycotoxin-free samples before extraction (e.g., for preparation of the matrix-matched calibration solutions and RE experiments), and promote analyte dispersion onto the whole sample, the procedure was as follows. Two hundred and fifty mg of durum wheat flour or corn meal was soaked in 250 µL of acetone containing the required amount of the analyte working standard mixture. Then, the sample was placed for 1 h in a ventilated oven at 40 °C to let the organic solvent evaporate. Finally, the spiked sample was extracted as reported above. The fortification levels were at ML, i.e., 1 µg kg^{-1} for each AF, 3 µg kg^{-1} for OTA and 750 µg kg^{-1} for ZEN (obtained by adding 25 µL of the working solution, 10 pg μL^{-1} each AF, 30 pg μL^{-1} OTA, and 750 pg μL^{-1} ZEN), and its multiples or fractions.

4.5. Liquid Chromatography-Tandem Mass Spectrometry Conditions

The UHPLC/ESI-MS/MS system was from Thermo Fisher Scientific (Bremen, Germany) and was constituted by a triple quadrupole mass spectrometer, mod. TSQ (triple stage quadrupole) Vantage EMR^{TM} (enhanced mass range), coupled to an UHPLC system Ultimate 3000 binary pump via a heated ESI source. The software Xcalibur^{TM} 2.2 (Thermo Fisher Scientific, Bremen, Germany) was used for LC-MS data acquisition and processing.

Sample aliquots of 10 µL were injected via the UHPLC autosampler. The six mycotoxins were separated onto a reversed-phase Cortecs UPLC C18+ column (100 mm × 2.1 mm i.d., 1.6 µm particle size), preceded by a VanGuard pre-column (5 mm × 2.1 mm i.d.) packed with the same stationary phase (Waters Milford, MA, USA). The column was thermostatted at 40 °C. The mobile phase was water (A) and methanol (B), both containing 5 mmol L^{-1} ammonium formate and 0.1% formic acid; the flow-rate was 300 µL min^{-1}. The elution gradient was the following: after 0.5 min at 15%, B was linearly increased to 35% in 1 min, then to 68% in 3.5 min, and finally to 75% in 3 min. To rinse the column, B was brought to 98% in 1 min, and held constant for 3 min. After bringing B back to 15% in 0.5 min, the column was allowed to equilibrate for 5.5 min.

MS Data were acquired in multiple reaction monitoring (MRM) mode, by operating ESI source in both positive (for the four AFs, OTA, and the corresponding ISs) and negative (for ZEN and its IS) ionization mode. The tune parameters were set as follows: spray voltage, +3.2/−2.8 kV; vaporizer temperature, 280 °C; capillary temperature, 220 °C; sheath gas pressure, 50 (arbitrary units, a.u.); sweep ion gas pressure (+) 0/(−) 1 (a.u.); auxiliary gas pressure, 25 (a.u.). For each compound, from two to three MRM transitions were monitored (see Table 1).

Monthly, the calibration solutions provided by Thermo Fisher Scientific (range m/z 69-2800) were injected in infusion mode for mass calibration and resolution adjustments of the resolving lens and quadrupole.

4.6. Analytical Method Performance

To assess the performance of the developed analytical method, overall PE, trueness and precision (both intra-day and inter-day precision), MLODs, and MLOQs were considered.

4.6.1. Process Efficiency (Recovery and Matrix Effect)

RE, ME, and PE were evaluated as in a previous work [29]. Blank matrix solutions spiked with mycotoxin standards before and after extraction were named as sample set 1 and set 2, respectively, whereas neat standard solutions prepared in pure solvents were named as sample set 3. For each analyte, the absolute peak area was measured. The RE was assessed according to the following equation:

$$\text{RE (\%)} = (\text{Area set1}/\text{Area set2}) \times 100 \quad (1)$$

ME was estimated according to the following equation:

$$\text{ME (\%)} = (\text{Area set2}/\text{Area set3}) \times 100 \quad (2)$$

PE, which is the product between RE and ME, was estimated according to the following equation:

$$\text{PE (\%)} = (\text{Area set1}/\text{Area set3}) \times 100 \quad (3)$$

The fortification levels used to assess PE were three (n = 6 for each level). For matrix solutions (sample sets 1 and 2), the levels were: ML, 0.5 × ML and 2 × ML. For standards solutions (sample set 3), these levels corresponded to 1 pg μL^{-1} of the four AFs, 3 pg μL^{-1} of OTA, and 75 pg μL^{-1} of ZEN, defined as standard ML (sML), 0.5 × sML and 2 × sML.

4.6.2. Calibration Graphs

For each mycotoxin, both six-point standard and matrix-matched calibration graphs were constructed. Standard solutions were prepared in water/methanol (50:50, *v*/*v*) at 0.2 × sML, 0.3 × sML, 0.5 × sML, sML, 2 × sML, and 4 × sML concentration levels. Matrix-matched solutions were prepared by spiking analyte-free samples before extraction at 0.2 × ML, 0.3 × ML, 0.5 × ML, ML, 2 × ML, and 4 × ML levels. The same amount of the three ISs was added to all the solutions, i.e., 2 pg μL^{-1} for AFM1, 6 pg μL^{-1} for OTA-d5 and 150 pg μL^{-1} for ZEN-d6.

Each solution was prepared in duplicate and injected twice, starting from the lowest up to the highest concentration level; finally, the results were averaged to give rise to a single calibration graph for each mycotoxin.

For each analyte, the combined ion current profile for the selected transitions was extracted from the LC-MRM dataset; the resulting traces were smoothed (Gaussian type, 7 points) by applying the automatic processing smoothing of XcaliburTM software.

The analyte to the corresponding IS peak area ratio versus the analyte concentration was plotted, considering the sum of all the MRM transitions to measure the areas. Unweighted regression lines for standard and matrix-matched calibration graphs were calculated using XcaliburTM QuanBrowser (Thermo Fisher Scientific).

4.6.3. Trueness and Precision

To assess the trueness of the developed method, each mycotoxin apparent recovery was calculated according to the following equation:

$$\text{Apparent RE (\%)} = [(\text{Area analyte set 1}/\text{Area IS set 1})/(\text{Area analyte set 2}/\text{Area IS set 2})] \times 100 \quad (4)$$

i.e., by comparing the analyte to IS peak area ratios in free-analyte flour samples spiked before and after the extraction procedure. The samples were spiked at the same levels used for PE assessment; the amounts of the three ISs was the same used in the calibration experiments. For each spiking level, six replicates were performed.

Intra-day (repeatability) and inter-day (reproducibility) were used to evaluate within laboratory precision of the developed analytical method. The relative standard deviation of the apparent recovery values of six spiked samples, at 0.5 × ML concentration, analyzed in the same day (RSDr) and in six consecutive days (RSD_R) were used to estimate the intra-day and inter-day laboratory precision, respectively.

4.6.4. Method Limits of Detection and Quantification

MLODs and MLOQs of the analytes were assessed as reported in previous works [8,33]. Briefly, a first estimation was done in the classical way, i.e., according to the following equations:

$$\text{MLOD} = 3 \times \sigma/S \quad (5)$$

and

$$\text{MLOQ} = 10 \times \sigma/S \quad (6)$$

where σ is the standard deviation of the intercept and S the slope of the matrix-matched calibration graph. For MLOQ calculation, the peak area obtained by the sum of all the MRM transitions was used, whereas for MLOD calculation, the peak area obtained only by the second most intense MRM transition was considered.

After those calculations, the MLOD and MLOQ values obtained according to Equations (5) and (6), respectively, were verified. Therefore, corn meal and wheat flour samples were spiked with the

six mycotoxins at levels very close to the extrapolated MLOQ values, and subject to the whole analytical procedure. For limits confirmation, the following equations were used:

$$\text{MLOD} = 3 \times S/N \quad (7)$$

and

$$\text{MLOQ} = 10 \times S/N \quad (8)$$

where S/N is the signal to noise ratio manually estimated by the LC-MRM data set, since the S/N provided by Thermo Xcalibur Qual Browser software by both INCOS noise method and manual noise region selection, was unlikely high.

Supplementary Materials: The following are available online at www.mdpi.com/2072-6651/9/4/147/s1, Table S1: *Zea mays* meal and *Triticum durum* flour composition, Table S2: Recovery, Matrix Effect, and Process Efficiency using different corn meal amounts, Table S3: Matrix effects using two different C18 chromatographic columns, Table S4: Equations and coefficient of determination relative to standard and matrix-matched calibration curves, Table S5: Results of 10 corn meal sample survey, Figure S1: Extracted ion chromatograms (sum of three transition pairs for each analyte) of a wheat flour sample spiked with the analytes at 0.5 × ML.

Acknowledgments: This research was supported by University of Rome "La Sapienza" (Call for proposals for Scientific Research "Ateneo 2016").

Author Contributions: A.L.C., C.C., P.F. and A.L. conceived and designed the experiments; G.L.B. and C.M.M. performed the experiments; P.F., C.M.M. and R.Z.C. analyzed and evaluated the data; A.L. provided the LC-MS/MS instrument; C.C. and A.L.C. supervised the experimental work and data analysis; R.Z.C. drafted the manuscript and all authors contributed to the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Capriotti, A.L.; Caruso, G.; Cavaliere, C.; Foglia, P.; Samperi, R.; Lagana, A. Multiclass mycotoxin analysis in food, environmental and biological matrices with chromatography/mass spectrometry. *Mass Spectrom. Rev.* **2012**, *31*, 466–503. [CrossRef] [PubMed]
2. Zöllner, P.; Mayer-Helm, B. Trace mycotoxin analysis in complex biological and food matrices by liquid chromatography–atmospheric pressure ionisation mass spectrometry. *J. Chromatogr. A* **2006**, *1136*, 123–169. [CrossRef] [PubMed]
3. Turner, N.W.; Bramhmbhatt, H.; Szabo-Vezse, M.; Poma, A.; Coker, R.; Piletsky, S.A. Analytical methods for determination of mycotoxins: An update (2009–2014). *Anal. Chim. Acta* **2015**, *901*, 12–33. [CrossRef] [PubMed]
4. Cavaliere, C.; Foglia, P.; Samperi, R.; Laganà, A. Determination of Aflatoxins and Ochratoxin A in Olive Oil. In *Olives and Olive Oil in Health and Disease Prevention*; Elsevier: Amsterdam, The Netherlands, 2010; pp. 645–652.
5. International Agency for Research on Cancer; Working Group on the Evaluation of Carcinogenic Risks to Humans. Some traditional herbal medicines, some mycotoxins, naphthalene and styrene. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*; IARC: Lyon, France, 2002.
6. Marin, S.; Ramos, A.J.; Cano-Sancho, G.; Sanchis, V. Mycotoxins: Occurrence, toxicology, and exposure assessment. *Food Chem. Toxicol.* **2013**, *60*, 218–237. [CrossRef] [PubMed]
7. Anfossi, L.; Giovannoli, C.; Baggiani, C. Mycotoxin detection. *Curr. Opin. Biotechnol.* **2016**, *37*, 120–126. [CrossRef] [PubMed]
8. Belen Serrano, A.; Capriotti, A.L.; Cavaliere, C.; Piovesana, S.; Samperi, R.; Ventura, S.; Lagana, A. Development of a Rapid LC-MS/MS Method for the Determination of Emerging Fusarium mycotoxins Enniatins and Beauvericin in Human Biological Fluids. *Toxins* **2015**, *7*, 3554–3571. [CrossRef] [PubMed]
9. Edite Bezerra da Rocha, M.; da Chagas Oliveira Freire, F.; Erlan Feitosa Maia, F.; Izabel Florindo Guedes, M.; Rondina, D. Mycotoxins and their effects on human and animal health. *Food Control* **2014**, *36*, 159–165. [CrossRef]

10. International Agency for Research on Cancer; Working Group on the Evaluation of Carcinogenic Risks to Humans. IARC monographs on the evaluation of carcinogenic risks to humans. In *Chemical Agents and Related Occupations*; Weltgesundheitsorganisation, Ed.; IARC: Lyon, France, 2012; Volume 100 F.
11. Bui-Klimke, T.R.; Wu, F. Ochratoxin A and Human Health Risk: A Review of the Evidence. *Crit. Rev. Food Sci. Nutr.* **2015**, *55*, 1860–1869. [CrossRef] [PubMed]
12. Pinotti, L.; Ottoboni, M.; Giromini, C.; Dell'Orto, V.; Cheli, F. Mycotoxin Contamination in the EU Feed Supply Chain: A Focus on Cereal Byproducts. *Toxins* **2016**, *8*. [CrossRef] [PubMed]
13. European Commission. The Rapid Alert System for Food and Feed (RASFF)—2015 Annual Report. Health and Food Safety—2016. Available online: https://ec.europa.eu/food/safety/rasff_en (accessed on 9 January 2017).
14. European Commission. Commission Regulation (EC) No. 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs (Text with EEA relevance). *Off. J. Eur. Union* **2006**, *364*, 5–24.
15. Ríos, A.; Zougagh, M.; Bouri, M. Magnetic (nano)materials as an useful tool for sample preparation in analytical methods. A review. *Anal. Methods* **2013**, *5*, 4558–4573. [CrossRef]
16. Wierucka, M.; Biziuk, M. Application of magnetic nanoparticles for magnetic solid-phase extraction in preparing biological, environmental and food samples. *TrAC Trends Anal. Chem.* **2014**, *59*, 50–58. [CrossRef]
17. Mashhadizadeh, M.H.; Amoli-Diva, M.; Pourghazi, K. Magnetic nanoparticles solid phase extraction for determination of ochratoxin A in cereals using high-performance liquid chromatography with fluorescence detection. *J. Chromatogr. A* **2013**, *1320*, 17–26. [CrossRef] [PubMed]
18. Hashemi, M.; Taherimaslak, Z.; Rashidi, S. Application of magnetic solid phase extraction for separation and determination of aflatoxins B1 and B2 in cereal products by high performance liquid chromatography-fluorescence detection. *J. Chromatogr. B* **2014**, *960*, 200–208. [CrossRef] [PubMed]
19. Capriotti, A.L.; Cavaliere, C.; Foglia, P.; La Barbera, G.; Samperi, R.; Ventura, S.; Laganà, A. Mycoestrogen determination in cow milk: Magnetic solid-phase extraction followed by liquid chromatography and tandem mass spectrometry analysis: Sample Preparation. *J. Sep. Sci.* **2016**, *39*, 4794–4804. [CrossRef] [PubMed]
20. Cavaliere, C.; Foglia, P.; Pastorini, E.; Samperi, R.; Lagana, A. Development of a multiresidue method for analysis of major Fusarium mycotoxins in corn meal using liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2085–2093. [CrossRef] [PubMed]
21. Cavaliere, C.; Foglia, P.; Guarino, C.; Nazzari, M.; Samperi, R.; Laganà, A. A sensitive confirmatory method for aflatoxins in maize based on liquid chromatography/electrospray ionization tandem mass spectrometry. *Rapid Commun. Mass Spectrom. RCM* **2007**, *21*, 550–556. [CrossRef] [PubMed]
22. Amoli-Diva, M.; Pourghazi, K.; Hajjaran, S. Dispersive micro-solid phase extraction using magnetic nanoparticle modified multi-walled carbon nanotubes coupled with surfactant-enhanced spectrofluorimetry for sensitive determination of lomefloxacin and ofloxacin from biological samples. *Mater. Sci. Eng. C* **2016**, *60*, 30–36. [CrossRef] [PubMed]
23. Chen, X.H.; Pan, S.D.; Ye, M.J.; Li, X.P.; Zhao, Y.G.; Jin, M.C. Magnetic solid-phase extraction based on a triethylenetetramine-functionalized magnetic graphene oxide composite for the detection of ten trace phenolic environmental estrogens in environmental water: Sample Preparation. *J. Sep. Sci.* **2016**, *39*, 762–768. [CrossRef] [PubMed]
24. Jouany, J.P. Methods for preventing, decontaminating and minimizing the toxicity of mycotoxins in feeds. *Anim. Feed Sci. Technol.* **2007**, *137*, 342–362. [CrossRef]
25. Cavaliere, C.; D'Ascenzo, G.; Foglia, P.; Pastorini, E.; Samperi, R.; Lagana, A. Determination of type B trichothecenes and macrocyclic lactone mycotoxins in field contaminated maize. *Food Chem.* **2005**, *92*, 559–568. [CrossRef]
26. Capriotti, A.L.; Cavaliere, C.; Foglia, P.; Samperi, R.; Stampachiacchiere, S.; Ventura, S.; Lagana, A. Multiclass analysis of mycotoxins in biscuits by high performance liquid chromatography-tandem mass spectrometry. Comparison of different extraction procedures. *J. Chromatogr. A* **2014**, *1343*, 69–78. [CrossRef] [PubMed]
27. Faberi, A.; Foglia, P.; Pastorini, E.; Samperi, R.; Laganà, A. Determination of type B fumonisin mycotoxins in maize and maize-based products by liquid chromatography/tandem mass spectrometry using a QqQ linear ion trap mass spectrometer: LC/MS/MS of type B fumonisin mycotoxins in maize-based products. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 275–282. [CrossRef] [PubMed]

28. European Commission. Commission Decision 657/2002/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. *Off. J. Eur. Union* **2002**, *221*, 8–36.
29. Capriotti, A.L.; Cavaliere, C.; La Barbera, G.; Piovesana, S.; Samperi, R.; Zenezini Chiozzi, R.; Laganà, A. Polydopamine-coated magnetic nanoparticles for isolation and enrichment of estrogenic compounds from surface water samples followed by liquid chromatography-tandem mass spectrometry determination. *Anal. Bioanal. Chem.* **2016**, *408*, 4011–4020. [CrossRef] [PubMed]
30. Moreno, V.; Zougagh, M.; Ríos, Á. Hybrid nanoparticles based on magnetic multiwalled carbon nanotube-nanoC18SiO2 composites for solid phase extraction of mycotoxins prior to their determination by LC-MS. *Microchim. Acta* **2016**, *183*, 871–880. [CrossRef]
31. Manafi, M.H.; Allahyari, M.; Pourghazi, K.; Amoli-Diva, M.; Taherimaslak, Z. Surfactant-enhanced spectrofluorimetric determination of total aflatoxins from wheat samples after magnetic solid-phase extraction using modified Fe_3O_4 nanoparticles. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2015**, *146*, 43–49. [CrossRef] [PubMed]
32. Yan, Y.; Zheng, Z.; Deng, C.; Zhang, X.; Yang, P. Selective enrichment of phosphopeptides by titania nanoparticles coated magnetic carbon nanotubes. *Talanta* **2014**, *118*, 14–20. [CrossRef] [PubMed]
33. Capriotti, A.L.; Cavaliere, C.; Piovesana, S.; Stampachiacchiere, S.; Samperi, R.; Ventura, S.; Laganà, A. Simultaneous Determination of Naturally Occurring Estrogens and Mycoestrogens in Milk by Ultrahigh-Performance Liquid Chromatography-Tandem Mass Spectrometry Analysis. *J. Agric. Food Chem.* **2015**, *63*, 8940–8946. [CrossRef] [PubMed]

Chapter 4:

Suspected approach and food quality

1 Phytochemicals suspected analysis in food

In the second section of the thesis the application of a suspected screening for the analysis of phytochemicals in food is discussed. In the next paragraphs, three papers about a method development for the analysis of polyphenols (Paper III and IV) and GLSs (Paper V) in strawberries and cauliflowers, respectively, will be presented.

1.1 Background

As previously largely explained (Chapter 1, paragraph 3.1.1), phytochemicals are bioactive secondary metabolites synthesized by plants. Their characterization in complex biological matrices, such as food, can result very important for the assessment of a relationship between such metabolites and the bioactivity of a specific product. Polyphenols and GLSs are two classes of compounds that showed several positive effects on health, therefore, in Paper III, IV and V a metabolic profiling, i.e. the quantitative and/or qualitative determination of a defined group of related compounds or of specific metabolic pathways, of these classes of phytochemicals will be presented.

A suspected screening approach has been used in order to detect and identify the highest number of polyphenols in Paper III and IV and GLSs in Paper V. To achieve this goal, a chromatographic and mass spectrometric method development has been carried out. Indeed, in contrast to targeted methods where the identification is usually limited to metabolites for which authentic reference standards are available, suspected/untargeted profiling methods try to find analytical features of all detectable compounds to achieve the widest possible metabolic coverage[95]. Metabolite profiling is a difficult challenge because of the huge chemical diversity of substances to analyze and the wide variation in abundance of the metabolites present in a complex matrix.

With HRMS the metabolic profile of a complex biological sample can be significantly extended. However, MS should be coupled with a separation technique, which is necessary to separate the compounds in a complex mixture. A good separation technique provides the individual compounds RTs, which are later used to support their identification, and is especially needed for isomeric metabolites, which are otherwise indistinguishable [96]. In complex phytochemical mixtures there are a wide variety of positional isomers (e.g.,

flavonoid glycoconjugates [97]). Low-quality peaks could interfere with peak picking during data processing, strongly reducing the number of detected metabolites. Therefore, optimization of the chromatographic conditions in a suspected/untargeted analysis it is crucial for the coverage of the broad spectrum of metabolites.

However, metabolites separation is not the major bottleneck in suspected/untargeted metabolomics workflows. Once detected, metabolites have to be identified and, usually, tentative metabolite annotation by studying the *m/z* and deriving the molecular formulae does not provide a single confident identification, particularly when the sample contains a large number of isomers. To provide the real identity of the detected metabolite isomers, MS/MS experiments are used to ascertain structural information about the molecules[98]. Current metabolomics studies focus on the development of several approaches to obtain the MS/MS information, such as using DDA. However, DDA, as currently applied, is often not optimal to maximize the number of metabolites for which MS/MS data are acquired in a single run. Moreover, several parameters when setting a MS method should be taken in into account based on the class of analytes to analyze. Therefore, the optimization of the mass spectrometric conditions in a suspected/untargeted analysis could result in a less time-consuming method and more informative data for the identification of a larger number of compounds.

In Paper III, IV and V, a typical suspected analysis approach is presented, showing how different chromatographic and mass-spectrometric conditions can affect the metabolic profiling of polyphenols in strawberry for Papers III and IV and of GLSs in cauliflower in Paper V.

1.2 Sample extraction

One of the relevant points to assess food quality by MS-based metabolomics is the choice of proper sample preparation procedures. In any case, to prevent any substantial loss of possible relevant metabolites, minimum sample preparation is preferable. Moreover, when hundreds of compounds have to be analyzed, the extraction protocol should be the right compromise for the extraction of all the compounds belonging to the class of interest. For this reason, the protocols chosen for the suspected analysis of both polyphenols in strawberry and GLSs in cauliflower are quite simple and rapid protocols.

In Paper III and IV the extraction protocol of polyphenols from strawberry was chosen on the basis of the work by Kajdžanoska *et al.*[99] showing that the best extraction mixture for the whole spectra of antioxidant compounds in strawberries (i.e., flavonoids, cyanidins, ETs, etc.) is CH_3COCH_3/CH_3COOH (99:1 v/v). However, on the contrary of the mentioned work, in our case the starting material was lyophilized. Therefore, 30 % of water was added to extract the more polar compounds.

For GLSs extraction, a rapid comparison between two different extraction protocols was carried out in Paper V. Indeed, knowing that GLSs are naturally hydrolyzed by myrosinase when heated at 55–65 °C [100], it is important to choose an extraction protocol that could reduce myrosinase activity. The myrosinase is denatured in organic solvent and heating, so most part of extractions are performed at 70 °C [28, 101, 102]. However, several other extractions performed at room temperature have been published [103, 104]. In order to clarify this ambiguity we decided to test two simple extraction protocols varying from each other for the set temperature (70 °C either 25°C). In contrast to the most part of the papers but the work of Doheny-Adamns *et al.* [105], the comparison showed better performances for the protocol at 25 °C. Based on these results, the extraction at room temperature was chosen as the best extraction protocol.

1.3 Chromatographic and mass spectrometric methods evaluation

As previously mentioned, the chromatographic and mass spectrometric method could be crucial for a comprehensive profiling of a complex matrix. Paper III, IV and V are focused on different aspects of the method optimization. In paper III, both a chromatographic and mass spectrometric evaluation was accomplished onto an UHPLC-(ESI)-MS/MS system. Indeed, the coupling of a specific HPLC to a specific mass spectrometer requires a proper optimization in order to explore the whole potentialities of the coupling of such specific instruments. In paper IV, while keeping constant the spectrometric method optimized in paper III, a chromatographic comparison of different column systems was performed on both a standard mixture and a real sample. The aim of this evaluation was both showing the importance of the chromatographic evaluation prior to a suspected screening analysis, and showing the importance of performing this evaluation on a real sample rather than a simulated system such as a mixture of standards. Finally, in paper V a chromatographic

evaluation on a different class of compounds (GLSs) was accomplished by varying both phases and columns, to show how different conditions can affect both chromatographic and mass-spectrometric performances.

In all papers, a Vanquish UHPLC was coupled by means of ESI to a QExactive mass spectrometer. In Paper III, a Kinetex C18 column (100 x 2.1 mm^2, 2.6 µm particles) was used to test the conditions summarized in Table 4.1. In Paper IV, using the same chromatographic and mass spectrometric method optimized in Paper III (phases, temperature, gradient, mass spectrometers parameters), four chromatographic systems were tested: (a) one Kinetex C18 column (100 × 2.1 mm2 2.6 µm particles); (b) one Kinetex C18 column (100 × 2.1 mm2 ,1.7 µm particles); (c) two Kinetex C18 columns (100 × 2.1 mm2 , 2.6 µm particles) connected in series; (d) two Kinetex C18 columns (100 × 2.1 mm2 , 2.6 µm particles) connected in series. Finally, in Paper V, four different columns were tested (1) one column Luna Omega polar C_{18} (100 mm × 2.1 mm i.d., 1.6 µm particles); (2) one column Kinetex ™ Biphenyl (100 mm × 2.1 mm i.d., 1.7 µm particles); (3) one column Kinetex™ core-shell XB-C_{18} (100 mm × 2.1 mm i.d., 2.6 µm particles); (4) two columns Kinetex™ core-shell XB-C_{18} (100 mm × 2.1 mm i.d., 2.6 µm particles) connected in series with four different phases a) HCOOH 0.1% in both H_2O (A) and ACN(B); b) HCOOH 0.2% in both H_2O (A) and ACN(B); c) $HCOONH_4$ 5 mM in both H_2O (A) and ACN(B); d) $HCOONH_4$ 10 mM in both H_2O (A) and ACN(B). All the chromatographic systems were tested using a mass spectrometric method appropriate for the analysis of GLSs.

The performances of the several systems under investigation were evaluated by means of a suspected screening approach, in which the number of features was the parameter mainly used for the evaluation. In untargeted metabolomics, indeed, the number of detected features is the most commonly used parameter for the selection of an optimal LC–MS method. A feature is considered as a peak with a specific RT and *m/z* ratio that probably represents a real molecule. The feature number parameter is not free from misinterpretations, indeed, the main drawback of this approach is that false positive features (caused by poor chromatographic signals, MS signals not generated by proton adduct, in-source fragments, complex ions, such as dimers 2 M± H, and salt cluster ions) could be included in the count[96]. However, feature numbering allows rapid comparison of different methods, avoiding the time-consuming identification step. Moreover, several methods can be used to avoid false positives, mainly using some pre-processing data analysis tools such as MZmine software.

Table 4.1: LC-MS parameters for methods tested in Paper III

method	phase A, H_2O:HCOOH	phase B, ACN:HCOOH	gradient (%B)	t (°C)	*m/z range*	data dependent (N)	exclusion list	dynamic exclusion (sec)	NCE
a	99.9 : 0.1	99.9 : 0.1	5%- 60% in 19 min	40	150-1000	5	yes	1	40
b	99.9 : 0.1	99.9 : 0.1	5%-50% in 19 min	40	150-1000	5	yes	1	40
c	99.9 : 0.1	99.9 : 0.1	5%-40% in 19 min	40	150-1000	5	yes	1	40
d	99.9 : 0.1	99.9 : 0.1	5%-60% in 38 min	40	150-1000	5	yes	1	40
e	99.9 : 0.1	99.9 : 0.1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	40	150-1000	5	yes	1	40
f	99 ? 1	99 ? 1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	40	150-1000	5	yes	1	40
g	99 ? 1	99 ? 1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	50	150-1000	5	yes	1	40
h	99.9 : 0.1	99.9 : 0.1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	40	150-1000	5	yes	3	40
i	99.9 : 0.1	99.9 : 0.1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	40	150-1000	5	yes	1	25-55
l	99.9 : 0.1	99.9 : 0.1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	40	150-1000	10	yes	3	40
m	99.9 : 0.1	99.9 : 0.1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	40	150-1000	10	yes	1	40
n	99.9 : 0.1	99.9 : 0.1	5%-15% in 10 min, 15%-35% in 15 min, 35%-50% in 5 min	40	150-1000	10	no	3	40

For each chromatographic run, accurate mass ion chromatograms were imported into MZmine and, for each scan, a list of ions was generated. After noise filtering, a chromatogram was built for each mass that could be detected continuously over the scans. Later, the built chromatograms were deconvoluted into individual peaks and the isotopic peaks were removed from the peak list. The peak lists obtained for the replicates and the blank were aligned and peaks coming from the blank were subtracted. Only peaks found in all replicate analyzes were kept. LC–MS features were filtered out if the RSD of the integrated peak areas was greater than 25%. An additional step to remove in-source

fragments and adducts was also carried out in Paper IV and V. The data pre-treatment allowed to obtain the final number of features. Whereas the number of features was the parameter mainly used to evaluate the different methods in all the three papers, some other different strategies were used in each case.

In Paper III, the highest number of features was compared with the average number of data points counted for each detected feature. The condition with the highest number of features and highest minimum number of data point per peak was chosen as the best chromatographic condition. Once chosen the best chromatographic method, several mass spectrometric conditions were tested and compared my means of the number of fragmented features. The method that allowed to obtain the highest number of fragmented features was chosen as the best method. The final optimized chromatographic and mass spectrometric method was applied for the final identification that will be discussed in the next sections.

In Paper IV, at first an evaluation of chromatographic systems was carried out by comparing the shape and the repeatability of the peaks of the analyzed standard compounds. A second evaluation was carried out in a suspected screening approach taking into account the highest number of features and also the highest number of tentatively identified compounds to choose the best chromatographic system.

Finally, in Paper V, a first evaluation between the different phases used on each column was carried out by taking into account the highest number of features. Once chosen the best phase for all the columns, a comparison among columns was accomplished by taking into account the highest number of tentatively identified compounds.

1.4 Results of the methods evaluation

As mentioned in the previous paragraphs, Paper III consisted in the evaluation of several chromatographic and mass spectrometric conditions for the profiling of polyphenols in strawberries. Operating in DDA mode, the instrument alternates between MS and MS/MS scans, and it is necessary to ensure the highest number of detected peaks while simultaneously acquiring fragmentation data required for metabolite identification [106]. To achieve this, the acquisition of a sufficient number of data points across a chromatographic peak is necessary. Since the Orbitrap mass analyzer does not possess the high scan speed typical of a TOF mass analyzer, it was necessary to find a compromise between the performance of the UHPLC and MS systems.

Several chromatographic conditions, as summarized in Table 4.1, were tested. In chromatographic methods a–e the elution gradient was progressively slowed to increase the number of data points per chromatographic peak. The minimum number of data points required to properly define the chromatographic peak apex and shape was fixed at ten, and it was obtained with method (e). This minimum number of data points was considered for all compounds but the anthocyanins, which exhibit a peculiar chromatographic behaviour[107]. Anthocyanins exhibit a typical chromatographic peak broadening due to the coexistence of different forms[108], and thus more acidic mobile phases containing 1% (v/v) HCOOH and a higher column temperature were tested in methods (f) and (g) to stabilize the flavylium cationic form. As shown in Figure 4.1, however, the number of features decreased probably because the higher acidity had a negative effect on the intensity of the signals of non-anthocyanin compounds, and also a higher column temperature (50 °C) did not significantly increase the chromatographic efficiency of anthocyanin separation.

After the method (e) was chosen as the best method, the capability of the UHPLC–MS/MS system was tested to simultaneously acquire the required fragmentation data. Several parameters were varied to obtain the highest number of fragmented precursor ions (i.e. fragmented features). Firstly, the maximum number of the most abundant precursors to be selected for data-dependent MS/MS acquisition was set to 5 (TOP 5 acquisition) and 10 (TOP 10 acquisition), and the other parameters were kept unchanged. The TOP 5 acquisition gave the best result; in the TOP 10 acquisition mode, the scan rate is slower, which in turn implies that a smaller number of ions are fragmented. Secondly, the dynamic exclusion parameter was set to prevent the same precursor ion from being redundantly selected for a DDA scan within a certain time span. The dynamic exclusion duration was changed between 1 s and 3 s to promote the fragmentation of the less abundant ions, and the best result was obtained with 3 s. Thirdly, the normalized collision energy (NCE) was studied. If more than one NCE value is set, the QExactive mass spectrometer performs a stepwise fragmentation on the precursor at several NCEs. Although this could allow having a more comprehensive fragmentation pattern, it results in a lower number of fragmented features probably due to a short time delay in switching from one NCE value to another, which in turn leads to a slower scan rate. Fourthly, an exclusion list created by the user containing the most intense 500 ions detected in the blank was used. These ions were selected for the MS/MS scan, obviously leading to an increase of the number of fragmented features. At the end of the optimization, the best mass-spectrometric method (h), consisting of a TOP5 acquisition, with a dynamic

exclusion of 3 s, a one-step NCE fragmentation, and an exclusion list, allowed to avoid performing redundant MS/MS experiments on the same ions, maximizing the number of fragmented features.

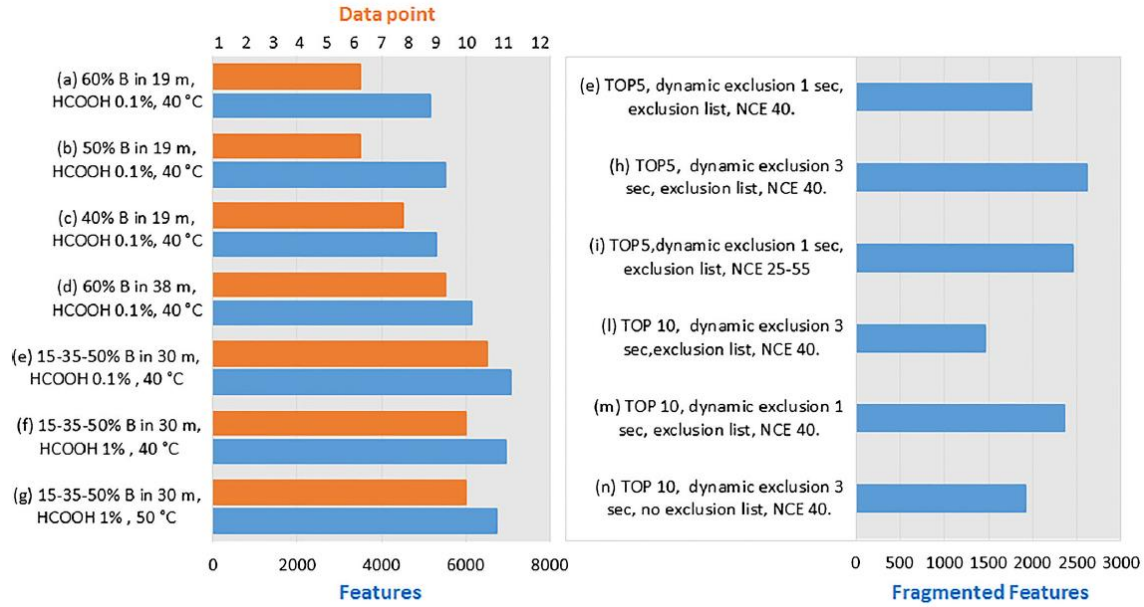


Figure 4.1: Left: Minimum number of data points per peak and detected features for methods (a)-(g). Right: Number of fragmented featured for method (h)-(n)

In Paper IV, starting from the optimized chromatographic and mass spectrometric conditions obtained in Paper III, four different chromatographic systems were compared to evaluate the effect of column length and particle size on the analysis of a complex phytochemical mixture. At first, a targeted approach was used for the comparison. The mixture of standards contained a range of flavonoids, chosen because representative of the phenolic compounds usually found in vegetable matrices. For each compound, the repeatability of the RT and the peak area were calculated, and an additional visual inspection of peak shapes was performed for each chromatographic condition. All metabolites were categorized into "unacceptable" if the shift of RTs and peak area was >25% and if the peak shape was affected by broadening. Peaks were classified as "acceptable" if the shift of RT and peak area was between 10 and 25% and if the peak was splitting or tailing. Finally, peaks were classified as "good" if the shift of RT and peak area was <10% and if they showed narrow and symmetric peaks.

At first glance, the performance of all the chromatographic systems was poor for most anthocyanins due to the typical peak broadening of anthocyanins. On the contrary, a very

strong difference among the chromatographic systems could be appreciated for some flavonoids belonging to the class of flavonols. These compounds resulted unacceptable in both the chromatographic systems with 1.7 µm sized particles compared to the 2.6 µm sized particles. Moreover, there was not such a difference between the one-column and the two-columns systems, regardless of the particle size.

To confirm or disprove the targeted analysis results and evaluate the effects on the analysis of a more complex matrix, we proceeded with the untargeted approach. The systems were compared at several levels as shown in Figure 4.2. Whereas a first result pointed out that the total number of detected features was larger for the chromatographic systems with 1.7 µm particle size than for the 2.6 µm particle sized ones, when the repeatable features have been taken into account the 1.7 sized columns showed a dramatic cut-off of detected features. This suggest that, regardless of the column lengths, the 1.7 µm sized particles columns performed worse than the 2.6 µm sized particles columns. Moreover, repeatability greatly benefits from a longer column setup. To confirm these results, however, the different systems were evaluated at a third more reliable level, that of the tentatively identified features. Non proton adducts complex ions and in-source fragments were identified using MZmine removed and assigned to the related molecular ion. The remaining features were searched against a home-made database then further investigated by a closer inspection of the MS/MS spectra resulting in the tentative identification of several compounds. Even at this level, the 2.6 µm sized particle columns performed better than the 1.7 µm sized particle columns, and the performance could be improved by the coupling of two columns. In conclusion, two main results are shown: the longer the column, the higher the number of repeatable detected and identified features; the smaller the particles size, the lower the number of repeatable detected and identified features. The first result was not highlighted in the targeted analysis because of the low complexity of the standard mixture compared to the strawberry extract, implying the strong importance of a chromatographic evaluation on a real matrix before an untargeted profiling. In the case of the second totally unexpected result, both targeted and untargeted analyzes suggested that 2.6 µm sized core-shell particles columns performed better than the 1.7 µm sized core-shell particles columns. This could be explained by the effect of temperature gradients through the columns. Indeed, when columns packed with very fine particles are operated at high mobile phase velocities, the friction of the mobile phase percolating through the column bed generates heat, and the band velocity is not constant across the column resulting in band broadening[109, 110]. Band broadening

could be accounted for the worse result observed for the Kinetex 1.7 µm sized particles packed columns in this study.

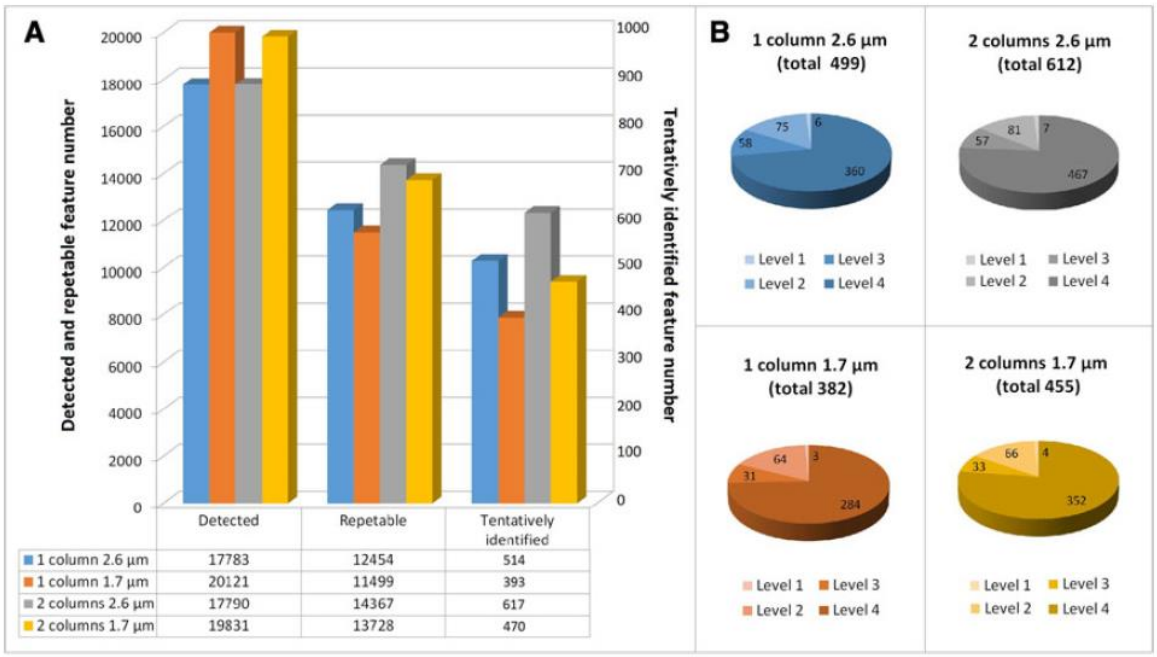
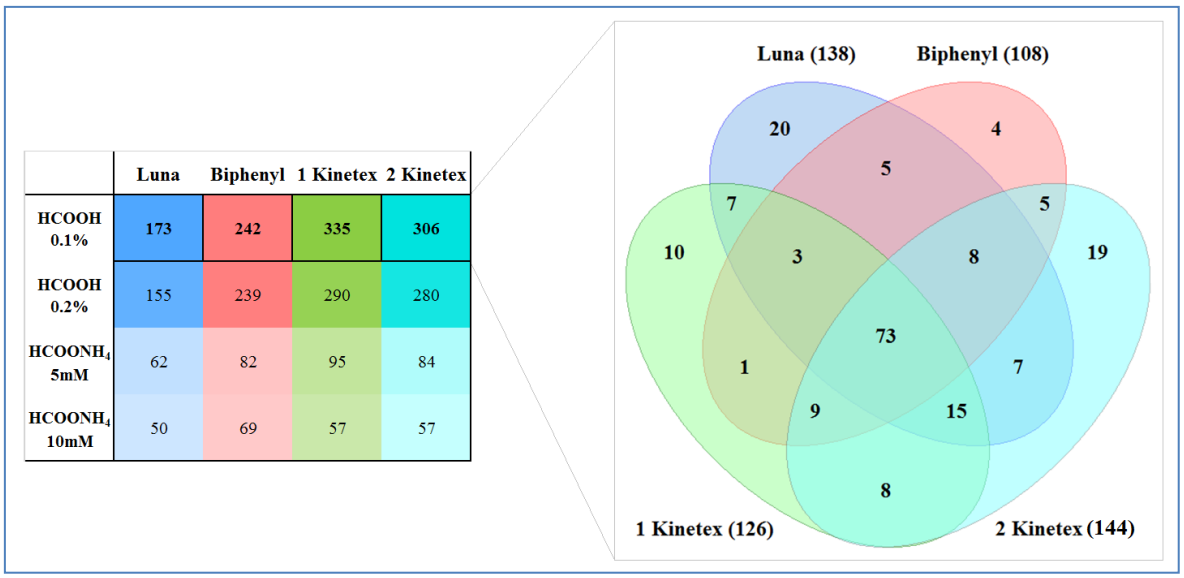


Figure 4.2: (A) Number of detected, repeatable, and tentatively identified features for the four chromatographic conditions; and (B) tentatively identified features after manual inspection with relative identification confidence level for the four chromatographic conditions: 1) match with standards, 2) match with MS/MS from literature, 3) match with some fragments from literature, 4) match with the accurate mass.

Finally, in Paper V, an evaluation of several chromatographic conditions was carried out on a completely different matrix and on a different class of compounds, i.e. GLSs, because, although widely studied, it is a class of compounds still generating increasing interest. Indeed, chromatographic separation of intact GLSs was historically a technical challenge that required several techniques such as ion pairing HPLC [111] or HILIC chromatography [112, 113], because of the poor retention of GLSs with hydrophilic side chains. In the last years, it has been shown that both trifluoroacetic acid (TFA) [114] and formic acid [28] induce sufficient adsorption of GLSs to reversed phase HPLC columns. However, when coupling HPLC with MS, besides a good chromatographic separation, effects such as ion suppression should be taken into account when using this compounds as additive in the phases. While it is known that TFA can give ion suppression during the ionization process, formic acid and ammonium formate, below certain concentration limits, should be compatible with MS. Therefore, we wanted to test formic acid and ammonium formate at

different concentrations in the chromatographic phases, for different columns, to obtain the best compromise between an efficient separation and a good ionization. Four different columns were evaluated with four different phases for a total of 16 different chromatographic conditions. The comparison of the different phases on each column was accomplished by the number of detected features that matched against a home-made database (i.e. level 4 identification) [102]. Looking at the table shown in Figure 4.3, it is possible to see that, in all cases, the phases with a percentage of formic acid give better results compared to the phases with ammonium formate. The obtained result, repeated for all the columns, suggests that the improved number of features is mostly due to an improvement in the ionization process rather than a real gain in chromatographic separation. When comparing different concentration of formic acid, it was shown a slightly increase of features number when using 0.1% compared to 0.2% of formic acid. Apparently, weather the peak shape was slightly better with 0.2% formic acid, the intensity of the peaks was sometimes decreased, losing the detection of the less abundant compounds. Based on these results, we decided that 0.1% formic acid phase was the good compromise between separation improvement and effective ionization. Once chosen the best phase for the four columns, we proceeded with the evaluation of the chromatographic performances among the columns. The four columns were compared after a visual inspection of the MS/MS spectra, thus based on the number of compounds identified at level 3. As shown in the Venn diagram, the number of tentatively identified GLSs was higher for the 2 Kinetex in series compared to the other systems. By a visual inspection of the chromatograms it could also be observed that the Luna column and the 2 Kinetex showed the best results in term of peak shape and repeatability. The similar peak shape together with the slightly difference in the number of features obtained with the Luna column and the 2 Kinetex pose a question about the choice between these two chromatographic systems for the analysis of GLSs. The obtained better result for the 2 Kinetex columns it is likely due to the higher retention of two coupled columns in series rather than an effective interaction of the analytes and the stationary phase. Indeed, when comparing one column Kinetex with one column Luna polar, the number of the identified compounds is higher and the shape of the peaks is better in Luna column compared to Kinetex column. This could be due to the fact that, while the C18 chains provide hydrophobic interactions, a polar modifier on the Luna column particle surface enhances the retention of the highly polar compounds. Luna Polar C18, allowing a more balanced retention of polar and hydrophobic compounds, is considered to be suitable for a class of

compounds such as GLSs, showing both high polarity (due to the sulfate moiety) and high heterogeneity in terms of hydrophobicity due to the variability of the –R side chain and the possible acylation of the thioglucose.



	Luna	Biphenyl	1 Kinetex	2 Kinetex
HCOOH 0.1%	173	242	335	306
HCOOH 0.2%	155	239	290	280
$HCOONH_4$ 5mM	62	82	95	84
$HCOONH_4$ 10mM	50	69	57	57

Figure 4.3: Table showing the number of repeatable tentatively identified features at confidence level 4 for all the conditions, Venn diagram showing the number of tentatively identified GLSs at confidence level 3 for four columns in the best chromatographic conditions, i.e. phases with HCOOH 0.1%.

This would suggest that the coupling of two Luna columns could give much better performances then the coupling of two Kinetex columns. In general, coupling two columns could be advantageous for a complete and comprehensive profiling of a matrix. However, when dealing with a large number of samples, and, when the differences between column performances are so slightly, a system implying the use of a single column should be preferred, in order to reduce the time and costs of the analysis. Due to the purpose of Paper V, namely obtaining a comprehensive profiling of GLSs in cauliflower, however, the two columns Kinetex in series were chosen as the best system for the final identification.

1.5 Identification

The identification is the bottleneck of suspected approaches. Several information have to be taken into account for the unambiguous identification of compounds, in the absence of standards. In Paper III, IV and V the identification of polyphenols and GLSs has been

accomplished by means of the investigation of the acquired MS/MS spectra. In Paper III and V, the method resulted to show the best performances was used for a final accurate identification, whereas in paper IV for all the tested conditions a tentative identification was performed in order to give a reliable method comparison.

In Paper III, on the basis of previous evaluations, the best chromatographic methods and mass spectrometric method was chosen for the identification. However, other settings needed to be optimized for each chemical class under investigation, since the mass-spectrometric and fragmentation behaviour strictly depend on the molecular structure. In this regard, the compounds reported in strawberry extracts, mainly belonging to the chemical classes of flavonoids, phenolic acids, dihydrochalcones, PAs and ETs[115] were treated differently by use of different mass-spectrometric parameters.

Flavonoids, usually present as glycoconjugates, show a particular behaviour in MS. The general trait of their mass spectra in full scan mode consists of ion signals corresponding both to the protonated glycosylated molecule and to the protonated aglycone formed during the ionization process. Some source parameters can promote in-source fragmentation before the ions reach the analyzer [97]. Because of this phenomenon the analysis of flavonoid glycoconjugates was performed with an NCE of 40 to break the glycosidic bond and obtain the aglycone residue. To investigate the structure of the aglycone, further acquisitions were performed with use of the source ion dissociation (SID) option and an NCE of 80. In this way, the glycoside group was removed before the ions entered the mass analyzer, allowing the aglycone moiety to be further fragmented to obtain a pseudo MS^3 spectrum and elucidate its structure. Finally, the alignment between the standard mass spectra and the SID mass spectra allowed the complete structure reconstruction.

The other two important classes of compounds found in strawberries are PAs and ETs. For these two classes, spectra were acquired in the m/z 500–1500 range under milder collision conditions (i.e., NCE of 20) to avoid loss of structural information. Moreover, an NCE of 20 was found to yield diagnostic ions similar to those in studies commonly found in literature, in which usually PAs and ETs are fragmented by CID. Identification of the phenolic compounds was based on the chromatographic behaviour, RT, accurate mass measurements, and MS/MS spectra. If authentic standards were not commercially available, the aforementioned parameters were compared with those in the MassBank database, the mzCloud database, and a homemade database; in all other cases, data were compared with literature data [115, 116]. Three identification levels were associated at each metabolite

according to Schymanski et al. [56]. On the contrary of previous works on strawberry profiling, the chromatographic and mass spectrometric optimization allowed the identification and tentative identification of 18 and 113 compounds respectively, including anthocyanins, other flavonoids, phenolic acids, PAs, and ETs.

Whereas in Paper III a comprehensive metabolic profiling of strawberry was accomplished, the aim of Paper IV was giving a reliable comparison of 4 systems. Therefore, the identification step, in this case, was carried out using a single acquisition for the analysis of all the classes of compounds present in strawberry. The acquisition consisted in a DDA acquisition mode, in the range 100-1000 m/z with a NCE 80. Using a NCE of 80 for all compounds, the typical fragments of ETs and PAs were not found. However, the identification was carried out exploiting the information obtained by the in source fragments found in the MS spectra. The best chromatographic condition, i.e. in series 2.6 μm particles sized packed columns (c), provided the largest number of identified compounds including anthocyanins, dihydrochalcones, dihydroflavonols, dihydroflavanols, flavanones, flavan-3-ols, PAs, and ETs. Also in this case identification levels according to Schymanski were assigned to the tentatively identified compounds (Figure 4.2).

In Paper V, on the basis of previous evaluations, the best chromatographic methods and mass spectrometric method was chosen for the identification. GLSs consist of a common glycone group and a variable aglycone side chain (R1). The glycone is characterized by a sulfonated oxime group and a thioglucose that can be conjugated to some acyl groups (R2) namely phenolic acids such as cinnamic acid, coumaric acid, etc. Due to the presence of a sulfonate moiety, glucosinolates, naturally occur in anionic form as [M-H], thus being identified in negative mode. Typical MS fragmentation patterns of glucosinolates have been widely investigated in CID [102, 104, 117], showing the presence of both product ions common to all GLSs and highly specific fragments due to the variability of the side chain (R1) and of the acyl group (R2). The major common fragmentation pathway [117] includes ions $[HSO_4]$-, $[C_2H_3O_5S]$-, $[HO_4S_2]$-, $[C_2H_2O_4NS]^-$, $[SO_3]^{-\cdot}$, $[C_2H_3OS]^-$ and $[C_2H_4O_5NS]^-$ at *m/z* 97, 139, 129, 136, 80, 75, 154, respectively. The three most intense ions (97, 75, 80) typical of the GLSs fragmentation pattern were used for a rapid screening of features in the step of columns comparison. Once chosen the best system, consisting of two Kinetex 2.6 in series, a deeper investigation on GLSs structure was carried out. Features selected by the rapid screening were studied and fragments resulting specific of the R1 or R2 side chains of GLSs were searched in the MS/MS spectra. GLSs were identified at different levels of

confidence: level 1 was associated to compounds identified by comparison with an authentic reference standard, level 2 was associated to compounds tentatively identified by means of typical MS/MS fragments already reported in the literature[102] (fragments derived from the R1 or R2 side chain), level 3 is associated to compounds that show some informative fragments, not already reported, or not enough specific to tentatively identify the compound but, sufficient to recognize the class the compound belongs to (level 3a and 3b if they could be associated to non-acylated or acylated GLSs, respectively.) On the contrary of previous works, in which a maximum of about 30 GLSs was found at the same time[28], in this work we could tentatively identify (at level 2, 3a and 3b) 51 GLSs, 24 of which have never been identified in cauliflower. This result show how a careful evaluation of the chromatographic system, coupled with an accurate identification, allows to extend the metabolic coverage of complex matrices.

1.6 Conclusions

Paper III, IV and V were focused on different aspects of the method optimization. As it was shown, several parameters have to be taken into account when performing a suspected or an untargeted screening. Firstly, the sample preparation should be simple in order to avoid eventual loss of metabolites and, at the same time, suitable for the extraction of the widest range of compounds. Secondly, as shown in all the presented papers, the chromatographic behaviour strictly affects the detection of compounds, therefore, phases, gradient, temperature should be optimized in order to achieve the best separation of the compounds present in a complex mixture. Thirdly, several factors have to be taken into account considering the coupling of the chromatographic system to the ESI source to allow an effective ionization, such as phases, as it was shown in Paper III and Paper V. Fourthly, the right chromatographic parameters and mass spectrometric parameters have to chosen to obtain the better outcome from the coupling of the separative instrumentation and the detector, such as the minimum number of data points or the number of fragmented precursors, as shown in Paper III. Fifthly, when dealing with different classes of compounds, particular attention has to be paid on the mass spectrometric parameters in order to allow the best fragmentation performances, as shown in Paper III and Paper V. Moreover, two main considerations can be done: an evaluation of the real system is always preferable compared

to an evaluation on a simulated systems such as a mixture of standards, as shown in Paper IV; a compromise between performances of the system and time and cost consume has always to be taken into account, of course depending on the aim of the study.

As clearly shown, the complexity of the food matrix, together with huge diversity of the metabolites in term of chemical properties pose serious difficulties in the accomplishment of a comprehensive metabolic profiling. Moreover, the lack of available databases and mass libraries on natural products, and software, adds major difficulties. However, several improvements in the techniques, and in the available software is showing the increasing potentialities of suspected approach for metabolic profiling (see also Appendix).

2 Paper III

Comprehensive polyphenol profiling of a strawberry extract (Fragaria × ananassa) by ultra-high-performance liquid chromatography coupled with high-resolution mass spectrometry

Anal Bioanal Chem (2017) 409:2127–2142
DOI 10.1007/s00216-016-0159-8

RESEARCH PAPER

Comprehensive polyphenol profiling of a strawberry extract (*Fragaria* × *ananassa*) by ultra-high-performance liquid chromatography coupled with high-resolution mass spectrometry

Giorgia La Barbera[1] · Anna Laura Capriotti[1] · Chiara Cavaliere[1] · Susy Piovesana[1] · Roberto Samperi[1] · Riccardo Zenezini Chiozzi[1] · Aldo Laganà[1]

Received: 17 September 2016 / Revised: 25 November 2016 / Accepted: 15 December 2016 / Published online: 11 January 2017

Abstract The aim of metabolic untargeted profiling is to detect and identify unknown compounds in a biological matrix to achieve the most comprehensive metabolic coverage. In phytochemical mixtures, however, the complexity of the sample could present significant difficulties in compound identification. In this case, the optimization of both the chromatographic and the mass-spectrometric conditions is supposed to be crucial for the detection and identification of the largest number of compounds. In this work, a systematic investigation of different chromatographic and mass-spectrometric conditions is presented to achieve a comprehensive untargeted profiling of a strawberry extract (*Fragaria* × *ananassa*). To fulfill this aim, an ultra-high-pressure liquid chromatography system coupled via an electrospray source to a hybrid quadrupole–Orbitrap mass spectrometer was used. Spectra were acquired in data-dependent mode, and several parameters were investigated to acquire the largest possible number of both mass spectrometry (MS) features and MS^2 mass spectra for unique metabolites. The main classes of polyphenols studied were flavonoids, phenolic acids, dihydrochalcones, ellagitannins, and proanthocyanidins. Method optimization allowed to us identify and tentatively identify 18 and 113 compounds, respectively, among which 74 have never been reported before in strawberries and, to the best of our knowledge, 22 of them have never been reported before. The results show the importance of an extended investigation of the chromatographic and mass-spectrometric method before a complete untargeted profiling of complex phytochemical mixtures.

Keywords Liquid chromatography · High-resolution mass spectrometry · Polyphenols · Strawberry · Untargeted profiling

Electronic supplementary material The online version of this article (doi:10.1007/s00216-016-0159-8) contains supplementary material, which is available to authorized users.

✉ Chiara Cavaliere
chiara.cavaliere@uniroma1.it

[1] Department of Chemistry, University of Rome "La Sapienza", Piazzale Aldo Moro 5, 00185 Rome, Italy

Introduction

Metabolomics is one of the youngest "omics" sciences [1], and in the last decade it has found application in several fields, such the life sciences, environmental science, forensic science, and food science [2]. In the past few years, metabolomics has gained an important role in food authentication, safety, and quality assessment [3] and in the investigation of the health effects of food intake [4].

In contrast to targeted methods, where the identification is usually limited to metabolites for which authentic reference standards are available, untargeted profiling methods try to find analytical features of all detectable compounds to achieve the widest possible metabolic coverage [5]. Untargeted metabolite profiling is a major challenge in metabolomics because of the huge chemical diversity of unknown substances and the wide variation in abundance of the metabolites in a complex matrix.

Nowadays, because of the recent rapid development of high-resolution mass spectrometers, which are able to provide the accurate mass of unknown compounds, the metabolic profile of a complex biological sample can be significantly extended. High-resolution mass spectrometry (HRMS) is usually coupled with a separation technique, which is necessary to separate the compounds in a complex mixture and provide their individual retention times, which are later used to support their identification, and is especially needed for isomeric metabolites, which are otherwise indistinguishable [6]. In

complex phytochemical mixtures there are a wide variety of positional isomers (e.g., flavonoid glycoconjugates [7]). Among the different separation techniques, liquid chromatography (LC) demonstrates the best characteristics in terms of the requirements for sample preparation, reproducibility, and metabolic coverage range [6]; it is usually used as a separation technique in HRMS-based metabolomics studies.

In untargeted metabolomics, the number of detected metabolites is the most important parameter in the selection of an LC–mass spectrometry (MS) method. Low-quality peaks could interfere with peak picking and alignment during data processing, consequently strongly affecting the resulting number of detected features. Therefore, optimization of the chromatographic conditions in an untargeted analysis could be crucial for maximization of the number of compounds detected in a complex matrix. In this way, untargeted metabolomics analysis allows the detection of thousands of metabolite features from a single experiment. However, the complexity of the sample could represent a significant difficulty in compound identification, and this process is a major bottleneck in metabolomics workflows. Tentative metabolite annotation by study of the mass-to-charge (m/z) ratio (measured with a mass accuracy of less than 5 ppm) by application of the seven golden rules [8] and by the deriving of the molecular formulae usually does not provide a single confident identification, particularly when the sample contains a large number of isomers. To provide the real identity of the detected metabolite isomers, MS^2 experiments are used to ascertain structural information about the molecules [9]. Current metabolomics studies typically consist of a multistep workflow in which a first profiling is done in MS mode to reveal the features of interest. These features are then fragmented in a second step to complete the identification. This time-consuming process has recently been shortened in a single workflow, in which data-dependent acquisition (DDA) allows the simultaneous acquisition of HRMS and MS^2 data, ensuring an efficient profiling and identification. DDA consists in the acquisition of MS^2 data for the most intense ions observed during the full scan profiling in the selected mass range; specific parameters can be set to avoid redundant MS^2 experiments. However, DDA, as currently applied, is often not optimal to maximize the number of unique metabolites for which MS^2 data are acquired, requiring further acquisition to elucidate the structure of not-fragmented compounds. Improvements in DDA have the potential to provide a larger amount of fragmentation data to be compared with reference MS^2 data sets, such as spectral libraries [9]. On the basis of these considerations, optimization of the mass-spectrometric conditions in an untargeted analysis could result in a less time-consuming method and more informative data for the identification of a larger number of compounds. Also the chromatogram quality has a crucial role in data processing.

Few studies on column evaluation and chromatographic method optimization have been presented in LC–HRMS-based metabolomics [6, 10–12]. Moreover, there are few articles focused on the investigation of the optimal DDA setting for routine applications by Orbitrap mass spectrometers [9].

In this work, a systematic investigation of different chromatographic and mass-spectrometric conditions is presented to optimize the untargeted analysis of a complex phytochemical mixture. Strawberry extract was used as a test system because of the known complexity in terms of isomeric antioxidant compounds [7]. We investigated the chromatographic conditions and the integrated acquisition of full-scan and higher-energy collisional dissociation (HCD) fragmentation data by an ultra-high-performance LC (UHPLC) system coupled via electrospray ionization (ESI) source to a hybrid quadrupole–Orbitrap mass spectrometer. From the results of the investigation, the best chromatographic and mass-spectrometric conditions were chosen for the final acquisition. Finally, the data were analyzed by our searching both home-made and online spectral databases. In contrast to previous work on identification of strawberry antioxidant compounds [13, 14], in which about 70 compounds were tentatively identified, in the present study the identification and tentative identification of 18 and 113 compounds respectively was achieved, 74 of which have never been found before in strawberries, showing the importance of a thorough investigation of the best instrumental parameters before the untargeted profiling of complex phytochemical mixtures.

Materials and methods

Chemicals and reagents

Strawberries (*Fragaria* × *ananassa* 'Elsanta') were bought from a local farm. The fruits were washed with distilled water, cut into quarters, freeze-dried and kept at −80 °C until use. LC–MS grade water was purchased from Fisher Scientific (Rodano, Italy). LC–MS grade methanol (MeOH) and acetonitrile (ACN) were purchased from VWR International (Milan, Italy). Formic acid and acetone were purchased from Sigma-Aldrich (St. Louis, MO, USA). Callistephin chloride (pelargonidin 3-*O*-glucoside), cyanidin chloride, epicatechin, flavone, kaempferol, kuromanin chloride (cyanidin 3-*O*-glucoside), naringenin, pelargonidin chloride, peonidin 3-*O*-glucoside chloride, procyanidin B1, quercetin dihydrate, quercetin 3-*O*-glucoside, rutin (quercetin 3-*O*-rutinoside), and taxifolin analytical standards were purchased from Sigma-Aldrich. Catechin, isorhamnetin, morin, and phloretin analytical standards were purchased from Extrasynthese (Lyon, France).

Sample preparation

Phenolic compounds were extracted following the protocol of Kajdžanoska et al. [15] with some modifications. Three

hundred milligrams of freeze-dried strawberry samples was extracted in 9 mL of CH_3COCH_3–H_2O–CH_3COOH (70:29.5:0.5 v/v/v). The extracts were sonicated for 15 min in an ice bath and then centrifuged for 10 min at 2000 *g* at 15 °C. The supernatant was collected. The extraction was repeated once. The supernatants were mixed and concentrated to 2.5 mL under a gentle nitrogen flow in a water bath at 37 °C. The sample was then added to 250 µL of MeOH. The final extract solution (H_2O–MeOH, 90:10 v/v) was centrifuged for 5 min at 2300 *g* at 25 °C. Before analysis, the supernatant was filtered through a 13-mm Acrodisc syringe filter with a 0.2-µm GH Polypro membrane (Pall, Ann Arbor, MI, USA). The extract was divided into aliquots and stored at −20 °C for further analysis.

UHPLC–ESI-MS^2 analysis

A Vanquish binary pump H (Thermo Fisher Scientific, Bremen, Germany), equipped with a thermostatted autosampler and a thermostatted column compartment, was used for sample chromatographic separation on a Kinetex core–shell C_{18} column (100 mm × 2.1 mm) with particles size of 2.6 µm (Phenomenex, Torrance, CA, USA) at a flow rate of 600 µL min^{-1}.

Several elution gradients and conditions and two mobile phases were tested to analyze strawberry extracts and the mixture of 18 standards at 0.5 ng μL^{-1} in H_2O–MeOH (90:10 v/v). For details, see methods a–g in Table S1. The injection volume was 5 µL. All samples were run in triplicate followed by the injection of the standard mix and a blank sample of H_2O–MeOH (90:10 v/v). Method e was chosen as the best chromatographic method on the basis of the number of features and the number of data points (see "Optimization of chromatographic and mass-spectrometric conditions"). For method e, the column was thermostatted at 40 °C and the mobile phase was H_2O–HCOOH (99.9:0.1 v/v; solvent A) and ACN–HCOOH (99.9:0.1 v/v; solvent B); the elution gradient was as follows: 5% solvent B to 15% solvent B in 10 min, 15% solvent B to 35% solvent B in 15 min, 35% solvent B to 50% solvent B in 5 min. The chromatographic system was coupled to a Q Exactive hybrid quadrupole–Orbitrap mass spectrometer (Thermo Fisher Scientific) with a heated ESI source. The ESI source parameters were as follows: capillary temperature 275 °C in positive mode, 350 °C in negative mode; sheath gas 50 arbitrary units (a.u.) in positive mode, 55 a.u. in negative mode; auxiliary gas 15 a.u.; sweep gas 2.25 a.u. in positive mode, 3 a.u. in negative mode; spray voltage 3500 V in positive mode, 2500 V in negative mode; auxiliary gas heater temperature 450 °C in positive mode, 300 °C in negative mode; S-lens RF level 50%.

For all the chromatographic methods investigated, detection was conducted in HRMS DDA mode. MS data were acquired in the *m/z* 150–1000 range in positive ionization mode with a resolution (full width at half maximum, FWHM, at *m/z* 200) of 35,000. The automatic gain control (AGC) target value was 200,000 in full-scan mode. The maximum ion injection time was 100 ms. The isolation window width was 2 *m/z*. MS^2 fragmentation was performed on the five most intense ions detected in full-scan mode with a resolution FWHM of 17,500. The AGC target value was 100,000. Dynamic exclusion was set to 1 s. An exclusion list was set containing the ions most commonly detected in the blank. Fragmentation was achieved in the HCD cell at a normalized collision energy (NCE) of 40.

After the best chromatographic method (method e) had been chosen, several parameters were investigated to obtain the largest number of fragmented features. For more details, see methods h–n in Table S1. For each method, the sample was run in triplicate followed by the injection of the standard mix and the blank H_2O–MeOH (90:10 v/v). Method h was chosen as the best method on basis of the number of fragmented features (see "Optimization of chromatographic and mass-spectrometric conditions"). It consists of a TOP5 DDA, with an exclusion list, the dynamic exclusion set at 3 s, and a one-step fragmentation at an NCE of 40.

On the basis of the chromatographic and mass-spectrometric method optimization, the best conditions (method h) were chosen and used for the final acquisitions. Four final acquisitions at different mass ranges, NCEs, and polarities were performed to identify all classes of compounds under investigation: *m/z* 150–1200, NCE 40, positive and negative polarities (method o); *m/z* 150–1200, source-induced dissociation (SID) fragmentation (40 eV), NCE 80, positive and negative polarities (method p); *m/z* 50021500, NCE 20, positive and negative polarities (method q); (r) *m/z* 1500–2500, NCE 20, negative polarity (method q).

Data analysis

For each chromatographic method, accurate mass ion chromatograms obtained from three consecutive injections and from the blank were processed by the open source software program MZmine version 2.9 (http://mzmine.sourceforge.net/) [16]. Briefly, the accurate mass LC–MS data were imported, and for each scan a list of ions was generated with use of a mass-detection module (Exact mass algorithm with noise of 10^4). A filter was applied to correct the false signals around *m/z* peaks called "shoulder peaks" that are residues of the Fourier transform function (Gaussian peak model function). Then mass lists generated and filtered for each MS scan were used to build a chromatogram for each mass that could be detected continuously over the scans by means of the chromatogram builder tool (minimum time span 0.1 min; minimum height 10^5; *m/z* tolerance 0.001 or 2 ppm). Later, the deconvolution module deconvoluted these chromatograms into individual peaks (local minimum search algorithm with chromatographic threshold 30; minimum retention time range 0.03 min; minimum relative height 0%; minimum absolute

height 10^5; minimum top/edge 1.5; peak duration 2.5 min). With the isotopic peak grouper algorithms, the isotopic peaks were removed from the peak list, keeping only the highest peak (*m/z* tolerance 0.001 or 2 ppm; retention time tolerance 0; maximum charge 2; most intense representative isotope). The peak lists obtained for the replicates and the blank were normalized for retention time (*m/z* tolerance 0.001 or 5 ppm; retention time tolerance 0.05 min; minimum intensity standard 10^6) and aligned with the Ransac alignment module (*m/z* tolerance 0.001 or 5 ppm; retention time tolerance 0.05 min; retention time after correction 0.05 min; Ransac iterations 0; minimum number of points 20; threshold value 4). Peaks coming from the blank were subtracted. Only peaks found in all replicate analyses were kept. LC–MS features were filtered out if the relative standard deviation (RSD) of the integrated peak areas was greater than 25%. The resulting peaks represent the number of features as shown in Fig. 1 for each method.

For each mass-spectrometric method, raw data obtained from three consecutive injections and from the blank were processed to generate a list containing all the fragmented precursor ions (i.e., the fragmented features). By use of the MS^2 peak picker module (*m/z* tolerance 0.005 or 5 ppm; retention time tolerance 0.1 min) a list of chromatographic peaks for which at least a fragmentation scan had been acquired was built. Then, the peak extender module extended the peaks in both directions of the retention time to add missing data points (*m/z* tolerance 0.005 or 5 ppm; minimum height 10^5). Once the chromatograms had been obtained, the deconvolution, the deisotoping, and the alignment steps were repeated as previously described. Again the interference from the blank was removed, and only peaks found in all replicate analyses with an RSD on the integrated areas less than 25% were kept. The resulting peaks, characterized by MS^2 fragmentation spectra, represent the number of fragmented features, as shown in Fig. 1 for each method.

The analysis of the final acquisition spectra and the calculation of elemental compositions were performed with Qual Browser of Xcalibur 2.2 (Thermo Fisher Scientific). The compounds were putatively identified on the basis of accurate mass (±5 ppm) by both homemade and online spectral databases (MassBank, mzCloud). When databases were lacking, phenolic compounds were identified according to the characteristic fragmentation spectra and retention time by comparison with literature data.

Finally, Qual Browser of Xcalibur 2.2 was used to visualize standard mix solution runs for each method tested. Data points across the chromatographic peaks were observed for all the flavonoid standards but not the anthocyanin ones.

Results and discussion

Extraction procedure

The extraction protocol was chosen on the basis of the work by Kajdžanoska et al. [15] showing that the best extraction mixture for the whole spectra of antioxidant compounds in strawberries (i.e., flavonoids, cyanidins, ellagitannins, etc.) is CH_3COCH_3–CH_3COOH (99:1 v/v). However, in our case the starting material was lyophilized strawberries instead of fresh ones (this, in turn, was so we could homogenize the starting sample and use it for all experiments). Therefore, considering that the freeze-dried sample lost 95% of its initial weight, 300 mg of strawberries was extracted in 9 mL of CH_3COCH_3–H_2O–CH_3COOH (70:29.5:0.5 v/v/v), where H_2O was needed to extract the more polar compounds.

Optimization of chromatographic and mass-spectrometric conditions

For both positive ionization mode, and negative ionization mode, the ESI parameters were previously optimized by direct injection into the source of the individual standard solutions of six analytes (belonging to either aglycone and glycosylated flavonoids) at 2 ng μL^{-1} in H_2O–MeOH (90:10 *v/v*) together with H_2O–ACN–HCOOH (79.9:20:0.1 v/v/v) at a rate of 600 $\mu L\ min^{-1}$ through a high-performance LC tee splitter.

Operating in DDA mode, the instrument alternates between MS and MS^2 scans, and it is necessary to ensure the highest number of detected peaks while simultaneously acquiring fragmentation data required for metabolite characterization [17]. To achieve this, the acquisition of a sufficient number of data points across a chromatographic peak is necessary.

Since the Orbitrap mass analyzer does not possess the high scan speed typical of a time-of-flight mass analyzer, it was necessary to find a compromise between the performance of the UHPLC and MS systems. For the Orbitrap, the higher the resolution is set, the slower the scan rate is, even if the effective scan rate also depends on other settings (e.g., AGC intensity and ion accumulation time). Therefore, the resolution was set at 35,000 (FWHM at *m/z* 200) for the MS scan and at 17,500 for the MS^2 scan; the mass error was less than 3 ppm. Higher resolution settings for the MS scan did not significantly reduce the mass error.

Starting from this mass-spectrometric configuration, we tested several chromatographic conditions on the same core–shell C_{18} column, as described in "UHPLC–ESI-MS^2 analysis." In chromatographic methods a–e the elution gradient was progressively slowed to increase the number of data points per chromatographic peak. The minimum number of data points required to properly define the chromatographic peak apex and shape was fixed at ten, and it was obtained with method e. This minimum number of data points was considered for all compounds but the anthocyanins, which exhibit a peculiar chromatographic behavior [18]. Anthocyanins exhibit a typical chromatographic peak broadening due to the coexistence of different forms [19], and thus more acidic mobile phases containing 1% (v/v) HCOOH and a higher column

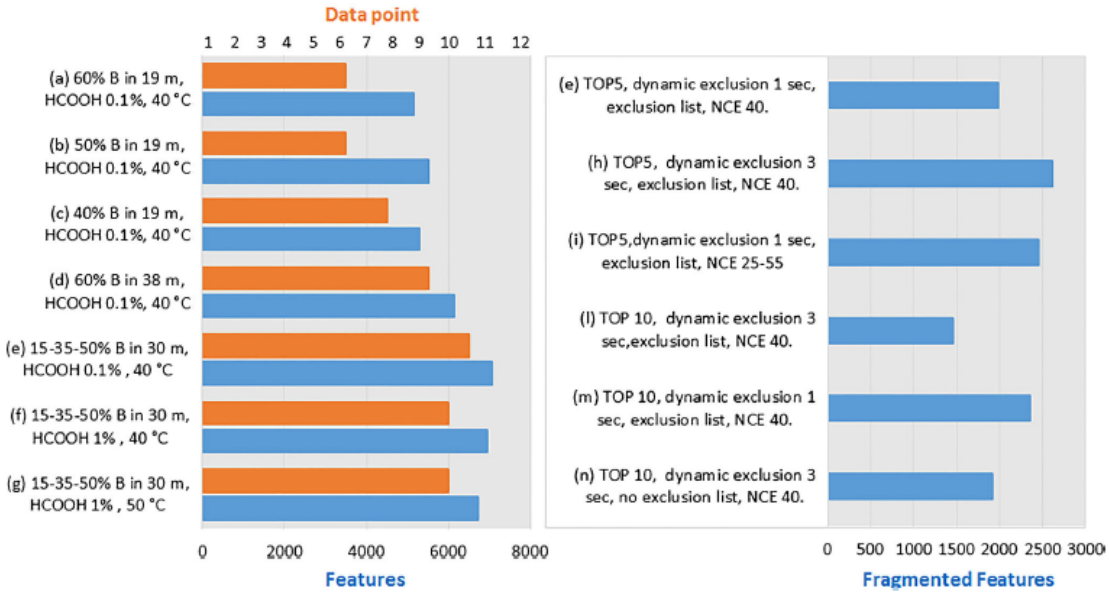


Fig. 1 *Left*: Summary of the minimum number of data points to define a peak and the detected features for methods a–g. *Right*: Number of fragmented features for methods e and h–n. *NCE* normalized collision energy

temperature were tested in methods f and g to stabilize the flavylium cationic form.

Figure 1 shows, for each of the LC–MS^2 methods tested, the number of data points per peak and the number of features that showed an RSD on the peak area of less than 25% for the three technical replicates. The feature number parameter is not free from misinterpretations; indeed, the main drawback of this approach is that false positive features (caused by poor chromatographic signals, MS signals not generated by proton adduct, in-source fragments, complex ions, such as dimers 2 M ± H, and salt cluster ions) could be included in the count [6]. However, feature numbering allows rapid comparison of different methods, avoiding the time-consuming identification step. As shown in Fig. 1, the best results were obtained for method e. The number of data points per peak and the number of features increased when the gradient slope was lowered from method a to method d, and further improvement was obtained with method e, where the gradient slope was lowered in the richest part of the chromatogram. In contrast, the higher acidity had a negative effect on the intensity of the signals of non-anthocyanin compounds, and also a higher column temperature (50 °C) did not significantly increase the chromatographic efficiency of anthocyanin separation. After the method had been assessed with the optimized number of detected features, the capability of the UHPLC–MS^2 system was tested to simultaneously acquire the required fragmentation data. Several parameters were varied to obtain the highest number of fragmented precursor ions (i.e. fragmented features). Firstly, the maximum number of the most abundant precursors to be selected for data-dependent MS^2 acquisition was set to 5 (TOP 5 acquisition) and 10 (TOP 10 acquisition), and the other parameters were kept unchanged. The TOP 5 acquisition gave the best result; in the TOP 10 acquisition mode, the scan rate is slower, which in turn implies that a smaller number of ions are fragmented. Secondly, the dynamic exclusion parameter was set to prevent the same precursor ion from being redundantly selected for a DDA scan within a certain time span. The dynamic exclusion duration was changed between 1 s and 3 s to promote the fragmentation of the less abundant ions, and the best result was obtained with 3 s; longer dynamic exclusion periods were incompatible with the narrow chromatographic peak width. Thirdly, the NCE was studied. This dimensionless value is approximately equivalent to the HCD collision energy (in electronvolts) for a reference ion of mass 500 u and charge 1. If more than one NCE value is set, the Q Exactive mass spectrometer performs a stepwise fragmentation on the precursor ion and all fragments created in the steps are analyzed during the same scan. Although setting several NCEs during a single scan should not affect the number of fragmented precursor ions, our results showed that a two-step fragmentation gives a lower number of fragmented features compared with the single-step fragmentation. This could be due to a short time delay in switching from one NCE value to another, which in turn leads to a slower scan rate. Fourthly, the opportunity of using an exclusion list was investigated. The exclusion list, created by the user, contains the most intense

500 ions detected in the blank; these ions will be not selected for the MS^2 scan. Obviously, the presence of the exclusion list increased the number of fragmented features. At the end of the optimization, the best mass-spectrometric method (method h) allowed us to avoid performing redundant MS^2 experiments on the same ions, maximizing the number of fragmented features, and consisted of a TOP5 acquisition, with a dynamic exclusion of 3 s, a one-step NCE fragmentation, and an exclusion list. The strawberry extract chromatogram acquired with method h is shown in Fig. 2.

Compound identification

On the basis of previous evaluations, the best mass-spectrometric method was chosen. However, other settings needed to be optimized for each chemical class under investigation, since the mass-spectrometric and fragmentation behavior strictly depend on the molecular structure. In this regard, the compounds reported in strawberry extracts, mainly belonging to the chemical classes of flavonoids (flavanols, flavonols, anthocyanidins), phenolic acids, dihydrochalcones, proanthocyanidins and ellagitannins [13] were treated differently by use of different mass-spectrometric parameters and interpretations of the spectra. Flavonoids are present as glycoconjugates. Different sugars (mainly hexose, deoxyhexose, pentose, polyglycoside), often acylated with aliphatic or aromatic acidic groups, might be linked to flavonoid aglycones. The general trait of glycoconjugate mass spectra in full-scan mode consists of ion signals corresponding both to the protonated glycosylated molecule and to the protonated aglycone formed during the ionization process. Some source parameters (e.g., temperature, voltage, gas pressure) can promote in-source fragmentation before the ions reach the analyzer [7]. Because of this phenomenon, for the analysis of flavonoid glycoconjugates, acquisition in the m/z 150–1200 range was performed with an NCE of 40 to break the glycosidic bond and obtain the aglycone residue. To investigate the structure of the aglycone further, acquisitions were performed in the m/z 150–1200 range, with use of the SID option (set to 40 eV) and an NCE of 80. In this way, the glycoside group was totally removed before the ions entered the mass analyzer, allowing the aglycone moiety to be further fragmented to obtain a pseudo MS^3 spectrum and elucidate its structure. Finally, the alignment between the standard mass spectra and the SID mass spectra allowed the complete (tentative) structure reconstruction. The same mass-spectrometric conditions were used for the analysis of phenolic acids and dihydrochalcones. Spectra were also acquired in negative ionization mode to provide additional information on phenolic acids. The other two important classes of compounds found in strawberries are proanthocyanidins and ellagitannins. For these two classes, spectra were acquired in the m/z 500–1500 range under milder collision conditions (i.e., NCE of 20) to avoid loss of structural information. Moreover, an NCE of 20 was found to yield diagnostic ions similar to those in studies commonly found in literature, in which usually proanthocyanidins and ellagitannins are fragmented by collision-induced dissociation. Spectra were acquired in both positive ionization mode and negative ionization mode to check the molecular masses. Two distinct runs for each polarity mode were preferred to polarity switching mode, to obtain a larger number of data points per peak. Finally, an acquisition in negative ionization mode in the m/z 1500–2500 range was performed to study some high molecular mass ellagitannins.

Final spectra acquisition allowed the identification and tentative identification of 18 and 113 compounds respectively, including anthocyanins, dihydrochalcones, dihydroflavonols, dihydroflavanols, flavanones, flavan-3-ols, phenolic acids, proanthocyanidins, and ellagitannins. Identification of the phenolic compounds was based on

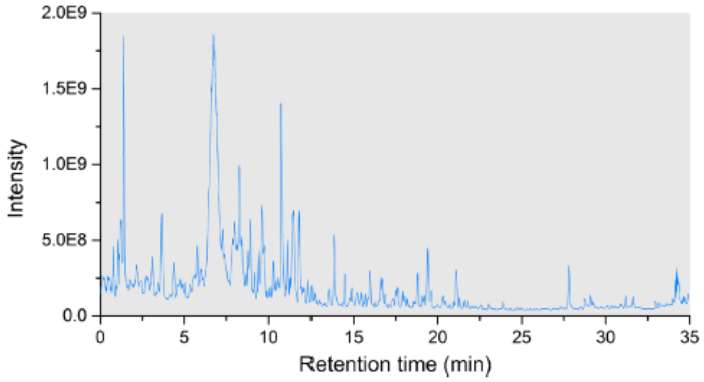


Fig. 2 Mass chromatogram of a strawberry extract acquired in positive ionization mode in the m/z 150–1000 range with method h

Table 1 Flavonoids, phenolic acids and dihydrochalcones identified in strawberry extract.

t_R (min)	$[M+H]^+/[M-H]^-$	Δ (ppm)	Molecular formula	Polarity	$[M+H]^+$ confirming peaks in MS[a]	MS^2/pseudo MS^3 fragments [b]	Identity	Identification confidence level[c]	Reference or library
1.70	741.2233	0.40	$C_{33}H_{41}O_{19}^+$	+	433.1129; 271.0600	MS^2: 271.0600 Pseudo MS^3: 121.0286; 65.0392; 141.0699; 115.0546; 197.0599; 169.0649; 271.0602	Pelargonidin hexosyl-deoxyhexosyl hexoside	2b	
3.10	291.0861	0.69	$C_{15}H_{14}O_6$	+		Pseudo MS^3: 123.0443; 139.0391; 147.0441; 119.0495; 91.0547; 68.9977	Epicatechin	1	mzCloud
3.44	447.0922	0.00	$C_{21}H_{19}O_{11}^+$	+	271.0600	MS^2: 271.0600 Pseudo MS^3: 121.0286; 65.0392; 141.0699; 115.0546; 197.0599; 169.0649; 271.0602	Pelargonidin glucuronide	2a	[32][d]
3.64	325.0925	2.17	$C_{15}H_{18}O_8$	-	163.0401	MS^2: 163.0401 Pseudo MS^3: 145.8898; 119.0493; 93.0335	Coumaroyl hexoside	2a	[27]
4.40	325.0925	2.17	$C_{15}H_{18}O_8$	-	163.0401	MS^2: 163.0401 Pseudo MS^3: 145.8898; 119.0493; 93.0335	Coumaroyl hexoside	2a	[27]
5.64	449.1079	−0.14	$C_{21}H_{21}O_{11}^+$	+	287.0550	MS^2: 287.0549 Pseudo MS^3: 121.0286; 56.9655; 84.9602; 81.0338; 146.9967	Cyanidin hexoside	1	[24]
6.24	291.0863	0.00	$C_{15}H_{14}O_6$	+		Pseudo MS^3: 123.0443; 139.0391; 147.0441; 119.0495; 91.0547; 68.9977	Catechin	1	[27]
6.66	433.1130	−0.23	$C_{21}H_{21}O_{10}^+$	+	271.0600	MS^2: 271.0602 Pseudo MS^3: 121.0286; 141.07; 115.0546; 65.0392; 169.0649; 197.0599	Pelargonidin hexoside	1	[29]
7.10	287.0549	0.35	$C_{15}H_{11}O_6^+$	+		Pseudo MS^3: 121.0286; 56.9655; 84.9602; 81.0338; 146.9967	Cyanidin	1	mzCloud
7.24	773.2127	1.03	$C_{33}H_{40}O_{21}$	+	627.1544; 465.1022; 303.0494	MS^2: 303.0499 Pseudo MS^3: 153.0184; 68.9978; 137.0234; 155.0493; 229.0497; 201.0549; 173.0601; 84.9602	Quercetin hexosyl-deoxyhexosyl hexoside	2a	[26][d]
7.80	463.1235	−0.02	$C_{22}H_{23}O_{11}^+$	+	301.0706	MS^2: 301.0706 Pseudo MS^3: 84.9603; 56.9655; 102.9708; 120.9812; 229.0496	Peonidin 3-*O*-glucoside	1	[13]
8.24	579.1706	0.34	$C_{27}H_{31}O_{14}^+$	+	417.1180; 271.0600	MS^2: 271.0600 Pseudo MS^3: 121.0286; 65.0392; 141.0699; 115.0546; 197.0599; 169.0649; 271.0602	Pelargonidin hexosyl-deoxyhexoside	2a	[24]
8.43	611.1610	−0.49	$C_{27}H_{30}O_{16}$	+	287.055; 449.1085	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol dihexoside	2a	[23]
8.78	757.2184	0.26	$C_{33}H_{40}O_{20}$	+	611.1609; 595.1663; 449.1077; 287.0548	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol deoxyhexose dihexoside	2b	
9.15	403.1023	0.25	$C_{20}H_{19}O_9^+$	+	271.0599	MS^2: 271.0600 Pseudo MS^3: 121.0286; 65.0392; 141.0699; 115.0546; 197.0599; 169.0649; 271.0602	Pelargonidin pentoside	2b	
9.20	305.0656	−0.03	$C_{15}H_{12}O_7$	+		Pseudo MS^3: 123.0444; 153.0185; 121.0287; 65.0393; 149.0236; 68.93; 157.0653	Taxifolin	1	[20][d]
9.80	271.0599	0.74	$C_{15}H_{11}O_5^+$	+		Pseudo MS^3: 121.0286; 65.0392; 141.0699; 115.0546; 197.0599; 169.0649	Pelargonidin	1	mzCloud
10.50	611.1247	−0.65	$C_{26}H_{26}O_{17}$	+	479.0824; 303.0499	MS^2: 303.0499 Pseudo MS^3: 153.0184; 68.9978; 137.0234; 155.0493; 229.0497; 201.0549; 173.0601; 84.9602	Quercetin glucuronyl pentoside	2b	
11.77	479.0824	−0.83	$C_{21}H_{18}O_{13}$	+	303.0498	MS^2: 303.0499 Pseudo MS^3: 153.0184; 68.9978; 137.0234; 155.0493; 229.0497; 201.0549; 173.0601; 84.9602	Quercetin glucuronide	2a	[13]
11.80	611.1609	−0.33	$C_{27}H_{30}O_{16}$	+		MS^2: 303.0502 Pseudo MS^3: 153.0184; 71.0498; 85.029; 137.0237; 229.0497; 201.055	Quercetin rutinoside	1	[24]
11.89	465.1033	−1.29	$C_{21}H_{20}O_{12}$	+	303.0499	MS^2: 303.0498 Pseudo MS^3: 153.0182; 69.0342; 137.0238; 85.029; 81.0341; 201.0549	Quercetin hexoside	1	[13]
12.17	581.1509	−1.55	$C_{26}H_{28}O_{15}$	+	449.1082; 287.0551	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol pentosyl hexoside	2b	
12.40	475.1228	1.47	$C_{23}H_{23}O_{11}$	+	271.0596	MS^2: 271.0600 Pseudo MS^3: 121.0286; 65.0392; 141.0699; 115.0546; 197.0599; 169.0649	Pelargonidin acetyl hexoside	2a	[29]
12.87	437.1447	−1.14	$C_{21}H_{24}O_{10}$	+		MS^2: 107.0495 Pseudo MS^3: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	Phloretin hexoside	2a	[26]
13.78	463.0870	0.22	$C_{21}H_{18}O_{12}$	+	287.0548	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol glucuronide	2a	[24]
13.85	449.1080	−0.45	$C_{21}H_{20}O_{11}$	+	287.0548	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol hexoside	2a	[24]
13.90	407.1337	0.00	$C_{20}H_{22}O_9$	+	275.0913	MS^2: 107.0495 Pseudo MS^3: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	Phloretin pentoside	2b	
14.19	581.1882	−2.93	$C_{27}H_{32}O_{14}$	+	435.1275; 419.1342; 273.0759	MS^2: 273.0756 Pseudo MS^3: 153.0185; 91.0548; 119.0495; 69.0341; 147.0443	Naringenin hexosyl-deoxyhexoside	2a	[26][d]
14.40	625.1770	−1.12	$C_{28}H_{32}0_{16}$	+	479.1187; 317.0655	MS^2: 317.0655 Pseudo MS^3: 153.0185; 217.0498; 245.0449; 92.0262; 229.0498; 203.0341	Isorhamnetin rutinoside	2a	[23]

Table 1 (continued)

t_R (min)	$[M+H]^+/[M-H]^-$	Δ (ppm)	Molecular formula	Polarity	$[M+H]^+$ confirming peaks in MS[a]	MS^2/pseudo MS^3 fragments [b]	Identity	Identification confidence level[c]	Reference or library
14.47	507.1141	−1.58	$C_{23}H_{22}O_{13}$	+	303.0499	MS^2: 303.0499 Pseudo MS^3: 153.0184; 68.9978; 137.0234; 155.0493; 229.0497; 201.0549; 173.0601; 84.9602	Quercetin acetyl hexoside	2a	[13]
14.61	479.1189	−1.04	$C_{22}H_{22}O_{12}$	+	317.0658	MS^2: 317.0655 Pseudo MS^3: 153.0185; 217.0498; 245.0449; 92.0262; 229.0498; 203.0341	Isorhamnetin hexoside	2a	[22][d]
14.83	493.0982	−1.01	$C_{22}H_{20}O_{13}$	+	317.0655	MS^2: 317.0655 Pseudo MS^3: 153.0185; 217.0498; 245.0449; 92.0262; 229.0498; 203.0341	Isorhamnetin glucuronide	2a	[26]
14.96	179.0341	1.20	$C_9H_8O_4$	-		MS^2: 135.0442	Caffeic acid	2a	[20][d]
15.40	535.1084	−0.37	$C_{24}H_{22}O_{14}$	+	449.1076; 287.0549	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol malonyl hexoside	2a	[23]
15.40	303.0497	−0.65	$C_{15}H_{10}O_7$	+		Pseudo MS^3: 68.9977; 153.0184; 219.0287; 137.0236; 121.0287; 84.9602	Morin	1	mzCloud
15.85	435.1284	0.46	$C_{21}H_{22}O_{10}$	+	273.0756	MS^2: 273.0756 Pseudo MS^3: 153.0185; 91.0548; 119.0495; 69.0341; 147.0443	Naringenin hexoside	2a	[23]
16.54	491.1189	−1.02	$C_{23}H_{22}O_{12}$	+	287.0548	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol acetyl hexoside	2a	[14]
16.89	583.2028	−1.20	$C_{27}H_{34}O_{14}$	+	437.1451; 421.1506; 275.0916	MS^2: 107.0495 Pseudo MS^3: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	Phloretin hexosyl-deoxyhexoside	2b	
17.23	303.0499	0.00	$C_{15}H_{10}O_7$	+		Pseudo MS^3: 153.0184; 68.9978; 137.0234; 155.0493; 229.0497; 201.0549; 173.0601; 84.9602	Quercetin	1	[14]
19.42	595.1449	−0.50	$C_{30}H_{26}O_{13}$	+	287.0545	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol *p*-coumaroyl hexoside	2a	[24]
19.97	595.1449	−0.50	$C_{30}H_{26}O_{13}$	+	287.0545	MS^2: 287.0548 Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol *p*-coumaroyl hexoside	2a	[24]
20.10	287.0552	−0.70	$C_{15}H_{10}O_6$	+		Pseudo MS^3: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	Kaempferol	1	[24]
26.16	223.0755	−0.45	$C_{15}H_{10}O_2$	+		Pseudo MS^3: 178.0776; 165.0699; 152.0619; 129.0334; 121.0285; 105.0449; 103.0545; 95.0495; 65.0391	Flavone	1	mzCloud

MS mass spectrometry

[a] Confirming peaks are ions found in MS because of spontaneous in-source fragmentation.

[b] MS^2 fragments were acquired by use of higher-energy collisional dissociation (HCD) fragmentation set to a normalized collision energy (NCE) of 40. Pseudo MS^3 fragments were acquired by use of source-induced dissociation (SID) fragmentation set to 40 eV followed by HCD fragmentation set to an NCE of 80.

[c] Level 1, confirmed structures where a reference standard is available; level 2a, evidence by spectrum matching with a spectrum from the literature or a library; level 2b, diagnostic evidence where no other structure fits the experimental MS^2 information

[d] Compound not found in strawberry extracts but found in strawberry secondary products or in other fruits and vegetables

Table 2 Ellagitannins identified in strawberry extract

t_R (min)	$[M+H]^+$/ $[M-H]^-$	Δ (ppm)	Molecular formula	Polarity	$[M+H]^+$ confirming peaks in MS[a]	MS^2 fragments[b]	Identity	Identification confidence level[c]	Reference
1.10	785.0819	1.66	$C_{34}H_{24}O_{22}$	+	483.0774; 333.0816	483.0757; 477.0562; 465.0676; 429.0455; 303.0136; 277.0344	DiHHDP-hexose	2a	[20][d]
1.25	951.0747	−0.21	$C_{41}H_{28}O_{27}$	-		907.0849; 781.0537; 605.0788; 479; 425.0251; 298; 273.0042;	Galloyl-HHDP-DHHDP-hexose (geraniin)	2a	[13]
1.42	951.0747	−0.21	$C_{41}H_{28}O_{27}$	-		907.0849; 781.0537; 605.0788; 479; 425.0251; 298; 273.0042;	Galloyl-HHDP-DHHDP-hexose (geraniin)	2a	[13]
1.60	301.0566	−0.33	$C_{12}H_{14}O_9$	-		169.0132; 149.9948; 125.0233	Galloyl pentose	2a	[13]
1.78	633.0725	−0.47	$C_{27}H_{22}O_{18}$	-		633.0725; 481.0619; 463.0512; 331. 0667; 300.9988; 275.0201	Galloyl-HHDP-hexose	2a	[22][d]
2.16	633.0725	−0.47	$C_{27}H_{22}O_{18}$	-	481.0629	633.0725; 481.0619; 463.0512; 331. 0667; 300.9988; 275.0201	Galloyl-HHDP-hexose	2a	[22][d]
2.28	951.0747	−0.21	$C_{41}H_{28}O_{27}$	-	633.0722	907.0849; 781.0537; 605.0788; 479; 425.0251; 298; 273.0042;	Galloyl-HHDP-DHHDP-hexose (geraniin)	2a	[13]
2.82	483.0765	0.83	$C_{20}H_{18}O_{14}$	+		303.0136; 277.0342	HHDP-hexose	2b	[33][d]
3.44	785.0851	1.2	$C_{34}H_{26}O_{22}$	-		633.0735; 615.0626; 463.0528; 483.0778; 313.0570; 300.9991	Digalloyl-HHDP-hexose	2a	[21][d]
4.75	633.0725	−0.47	$C_{27}H_{22}O_{18}$	-	481.0629	633.0725; 481.0619; 463.0512; 331. 0667; 300.9988; 275.0201	Galloyl-HHDP-hexose	2a	[22][d]
4.87	933.0644	−0.54	$C_{41}H_{26}O_{26}$	-		915.0509; 631.0575; 479.0464; 461.0377 300.9991	Galloyl-HHDP-hexose-HHDP (castalagin, vescalagin)	2a	[31][d]
5.27	933.0645	−0.64	$C_{41}H_{26}O_{26}$	-		915.0509; 631.0575; 479.0464; 461.0377; 300.9991	Galloyl-HHDP-hexose-HHDP (castalagin, vescalagin)	2a	[31][d]
6.11	951.0747	−0.21	$C_{41}H_{28}O_{27}$	-		907.0849; 781.0537; 605.0788; 479; 425.0251; 298; 273.0042;	Galloyl-HHDP-DHHDP-hexose (geraniin)	2a	[13]
6.24	933.0651	−1.29	$C_{41}H_{26}O_{26}$	-		915.0509; 631.0575; 479.0464; 461.0377; 300.9991	Galloyl-HHDP-hexose-HHDP (castalagin, vescalagin)	2a	[31][d]
6.26	785.0851	1.2	$C_{34}H_{26}O_{22}$	-		633.0735; 615.0626; 463.0528; 483.0778; 313.0570; 300.9991	Digalloyl-HHDP-hexose	2a	[21][d]
6.48	951.0747	−0.21	$C_{41}H_{28}O_{27}$	-		907.0849; 781.0537; 605.0788; 479; 425.0251; 298; 273.0042;	Galloyl-HHDP-DHHDP-hexose (geraniin)	2a	[13]
6.72	951.0747	−0.21	$C_{41}H_{28}O_{27}$	-		907.0849; 781.0537; 605.0788; 479; 425.0251; 298; 273.0042;	Galloyl-HHDP-DHHDP-hexose (geraniin)	2a	[13]
7.04	483.0784	−0.83	$C_{20}H_{20}O_{14}$	-		331.0674; 313.0574	Digalloyl hexose	2a	[21][d]
7.86	635.0899	−1.42	$C_{27}H_{24}O_{18}$	-		465.0666; 313.0562; 483.0768;	Trigalloyl hexose	2a	[13]
7.93	951.0747	−0.21	$C_{41}H_{28}O_{27}$	-		907.0849; 781.0537; 605.0788; 479; 425.0251; 298; 273.0042;	Galloyl-HHDP-DHHDP-hexose (geraniin)	2a	[13]
8.64	933.0653	−1.50	$C_{41}H_{26}O_{26}$	-		915.0509; 631.0575; 479.0464; 461.0377; 300.9991	Galloyl-HHDP-hexose-HHDP (castalagin, vescalagin)	2a	[31][d]
8.91	935.0786	1.07	$C_{41}H_{28}O_{26}$	-	783.0686	935.0807; 783.0688; 765.0568; 633.0729; 463.0518; 300.9989	Galloyl-diHHDP-hexose (casuarictin, potentillin)	2a	[13]
9.60	935.0786	1.07	$C_{41}H_{28}O_{26}$	-	783.0687	935.0807; 783.0688; 765.0568; 633.0729; 463.0518; 300.9989	Galloyl-diHHDP-hexose (casuarictin, potentillin)	2a	[13]
9.80	785.0816	2.04	$C_{34}H_{24}O_{22}$	+		483.0757; 477.0562; 465.0676; 429.0455; 303.0136; 277.0344	DiHHDP-hexose	2a	[20][d]
10.60	937.0951	0.11	$C_{41}H_{30}O_{26}$	-	767.0738	937.0951, 785.0786, 767.0738, 635.1984; 465.0669, 300.9990	Trigalloyl-HHDP hexose	2a	[20][d]
10.79	937.0951	0.11	$C_{41}H_{30}O_{26}$	-	767.0738	937.0951, 785.0786, 767.0738, 635.1984; 465.0669, 300.9990	Trigalloyl-HHDP hexose	2a	[20][d]
11.12	1567.1431	0.96	$C_{68}H_{48}O_{44}$	-		1415.1411; 1265.1367; 1103.0593; 1085.0763; 633.0740; 935.0805; 783.0688; 481.0610	Sanguiin h10	2a	[20][d]
11.44	1567.1431	0.96	$C_{68}H_{48}O_{44}$	-		1415.1411; 1265.1367; 1103.0593; 1085.0763; 633.0740; 935.0805; 783.0688; 481.0610	Sanguiin h10	2a	[20][d]
12.07	935.0786	1.07	$C_{41}H_{28}O_{26}$	-	783.0686	935.0807; 783.0688; 765.0568; 633.0729; 463.0518; 300.9989	Galloyl-diHHDP-hexose (casuarictin, potentillin)	2a	[13]
12.33	787.1008	−2.54	$C_{34}H_{28}O_{22}$	-	617.0777	787.0963; 635.0881; 617.0777; 465.0676; 300.9992; 169.0134	Tetragalloyl hexose	2a	[21][d]

Table 2 (continued)

t_R (min)	$[M+H]^+$/ $[M-H]^-$	Δ (ppm)	Molecular formula	Polarity	$[M+H]^+$ confirming peaks in MS[a]	MS^2 fragments[b]	Identity	Identification confidence level[c]	Reference
13.54	1869.1426	4.39	$C_{82}H_{54}O_{52}$	-	483.0769	1869.1426; 1567.1518; 1265.1403; 1085.0763; 935.0815; 783.0677; 633.0740; 613.0485; 481.0643; 300.9995	Galloyl-diHHDP-hexose dimer (agrimoniin)	2a	[13]
14.30	939.1141	−3.41	$C_{41}H_{32}O_{26}$	-		939.1118; 787.0982; 769.0890; 617.0796; 599.0702; 447.0588; 300.9988; 169.0135	Pentagalloyl hexose	2a	[13]

DHHDP dehydrohexahydroxydiphenic acid, *HHDP* hexahydroxydiphenic acid

[a] Confirming peaks are ions found in MS because of spontaneous in-source fragmentation

[b] MS^2 fragments are referred to acquisition with HCD set to an NCE of 20.

[c] Level 1, confirmed structures where a reference standard is available; level 2a, evidence by spectrum matching with a spectrum from the literature or a library; level 2b, diagnostic evidence where no other structure fits the experimental MS^2 information

[d] Compound not found in strawberry extracts but found in strawberry secondary products or in other fruits and vegetables

the chromatographic behavior, retention time, accurate mass measurements, and MS^2 spectra. If authentic standards were not commercially available, the aforementioned parameters were compared with those in the MassBank database, the mzCloud database, and a homemade database; in all other cases, data were compared with literature data [13, 14, 20–33]. The identification data for the tentatively identified compounds are discussed in the following sections and summarized in Tables 1, 2, and 3 with the related confidence level, according to Schymanski et al. [34]. In particular, level 1 refers to compounds identified by the checking of the standard retention time and MS^2 spectra; level 2a refers to compounds tentatively identified by the matching of MS^2 spectra data in the literature or in online spectral libraries; level 2b refers to compounds tentatively identified by study of the diagnostic fragments in MS^2 spectra but not supported by literature data. Coherently, the compounds tentatively identified with identification level 1 or level 2a have already been reported in the literature, whereas the compounds tentatively identified with identification level 2b have never been reported in the literature, to the best of our knowledge.

Interpretation of flavonoid, phenolic acid, and dihydrochalcone spectra

Seventeen standard compounds belonging to the classes of flavonoids and dihydrochalcones were used to assist the identification. The retention times of the aglycones were used to confirm the identity of both the detected aglycones and the respective glycosylated forms, which were expected to be eluted earlier from the C_{18} column since they are more polar. MS^2 spectra were acquired at an NCE of 80 for aglycone standards to confirm the detected aglycones and compare their fragmentation patterns with those of the aglycones generated from glycosylated compounds by SID plus HCD. Retention times and MS^2 spectra of standards are summarized in Table S2. All compounds for which authentic reference standards were available were identified on the basis of their m/z ratio (mass error less than 5 ppm) and retention time. When standards were not available, compounds were tentatively identified on the basis of whether (1) their m/z ratio was measured with a mass error less than 5 ppm, (ii) the accurate mass of their aglycone was found at the same retention time as for the precursor, and (iii) the fragmentation pattern of the aglycone, obtained in the SID plus HCD acquisition (pseudo MS^3), was confirmed by comparison with standard MS^2 spectra or with spectra found in MassBank, mzCloud, and in the literature, or by our following the spectra interpretation guidelines for flavonoids [7, 35–37]. Identified and tentatively identified

Table 3 Proanthocyanidins identified in strawberry extract

t_R (min)	$[M+H]^+$/ $[M-H]^-$	Δ (ppm)	Molecular formula	Polarity	$[M+H]^+$ confirming peaks in MS[a]	MS^2 fragments[b]	Identity	Identification confidence level[c]	Reference
0.90	867.2122	−1.04	$C_{45}H_{38}O_{18}$	+		697.1500; 579.1495; 427.1021; 409.0915; 289.0709; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
1.29	849.2035	1.18	$C_{45}H_{38}O_{17}$	-		831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
1.85	833.2094	2.16	$C_{45}H_{38}O_{16}$	-		707.1757; 561.1394; 543.1241; 435.1081; 407.0775; 289.0719; 271.0613	Proanthocyanidin trimer (Afz-Afz-Cat)	2a	[38][d]
2.14	867.2124	−0.81	$C_{45}H_{38}O_{18}$	+		697.1535; 579.1495; 427.1028; 409.0930; 289.0710; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
2.20	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-		983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
2.72	579.1497	1.04	$C_{30}H_{26}O_{12}$	+		453.1182; 427.1025; 409.0915; 301.0706; 291.0864; 289.0708; 163.0391; 127.0392	Procyanidin B1 (Cat-Cat)	1	[13]
3.18	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-		983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
3.57	867.2113	−2.08	$C_{45}H_{38}O_{18}$	+	579.1495	697.1535; 579.1495; 427.1028; 409.0930; 289.0710; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
3.63	849.2036	1.30	$C_{45}H_{38}O_{17}$	-		831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
3.80	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-	865.1988	983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
3.94	1137.2703	3.87	$C_{60}H_{50}O_{23}$	-		1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2b	
4.20	1137.2694	3.08	$C_{60}H_{50}O_{23}$	-		1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2b	
4.69	561.1399	1.43	$C_{30}H_{26}O_{11}$	-		543.1305; 435.1083; 407.0772; 289.0717; 271.0615	Proanthocyanidin dimer (Afz-Cat)	2a	[20][d]
4.74	867.2119	−1.38	$C_{45}H_{38}O_{18}$	+	579.1495	697.1535; 579.1495; 427.1028; 409.0930; 289.0710; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
4.97	561.1401	1.78	$C_{30}H_{26}O_{11}$	-		543.1305; 435.1083; 407.0772; 289.0717; 271.0615	Proanthocyanidin dimer (Afz-Cat)	2a	[20][d]
5.15	849.2040	1.77	$C_{45}H_{38}O_{17}$	-	577.1354	831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
5.19	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-	575.1345; 865.1988	983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
5.63	1137.2682	2.02	$C_{60}H_{50}O_{23}$	-		1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2b	
5.69	849.2047	2.59	$C_{45}H_{38}O_{17}$	-	561.1397	831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
5.90	849.2037	1.41	$C_{45}H_{38}O_{17}$	-	577.1354	831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.0600 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
6.03	867.2117	−1.61	$C_{45}H_{38}O_{18}$	+		697.1535; 579.1495; 427.1028; 409.0930; 289.0710; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
6.44	1137.2665	0.53	$C_{60}H_{50}O_{23}$	-	849.2025; 577.1354			2b	

Table 3 (continued)

t_R (min)	$[M+H]^+/[M-H]^-$	Δ (ppm)	Molecular formula	Polarity	$[M+H]^+$ confirming peaks in MS[a]	MS^2 fragments[b]	Identity	Identification confidence level[c]	Reference
						1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)		
6.60	1441.3282	2.78	$C_{75}H_{62}O_{30}$	-		1153.2584; 865.1946; 739.1681; 695.1397; 575.1201; 449.1038; 431.0983; 287.0567; 269.0468; 125.0232	Procyanidin pentamer (Cat-Cat-Cat-Cat-Cat)	2a	[38][d]
6.73	1137.2682	2.02	$C_{60}H_{50}O_{23}$	-		1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2b	
6.90	1137.2685	2.29	$C_{60}H_{50}O_{23}$	-		1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2a	[38][d]
6.96	1441.3256	0.97	$C_{75}H_{62}O_{30}$	-		1153.2584; 865.1946; 739.1681; 695.1397; 575.1201; 449.1038; 431.0983; 287.0567; 269.0468; 125.0232	Procyanidin pentamer (Cat-Cat-Cat-Cat-Cat)	2a	[38][d]
7.30	833.2097	2.52	$C_{45}H_{38}O_{16}$	-	561.1394	707.1757; 561.1394; 543.1241; 435.1081; 407.0775; 289.0719; 271.0613	Proanthocyanidin trimer (Afz-Afz-Cat)	2a	[38][d]
7.50	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-		983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
7.60	1137.2672	1.14	$C_{60}H_{50}O_{23}$	-		1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2a	[38][d]
7.96	1123.2839	−2.40	$C_{60}H_{50}O_{22}$	+	851.2177	851.2078;715.1606; 681.1587; 579.1505; 409.0923; 289.0709; 273.0761	Proanthocyanidin tetramer (Cat-Cat-Afz-Afz)	2b	
8.13	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-	575.1345; 865.1988	983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
8.30	833.2080	0.48	$C_{45}H_{38}O_{16}$	-		707.1757; 561.1394; 543.1241; 435.1081; 407.0775; 289.0719; 271.0613	Proanthocyanidin trimer (Afz-Afz-Cat)	2a	[38][d]
8.40	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-	865.1988	983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
8.64	561.1401	1.78	$C_{30}H_{26}O_{11}$	-		543.1305; 435.1083; 407.0772; 289.0717; 271.0615	Proanthocyanidin dimer (Afz-Cat)	2a	[20][d]
8.70	1153.2620	1.04	$C_{60}H_{50}O_{24}$	-	865.1988	983.2017; 865.1994; 739.1658; 695.1382; 577.1345; 575.1345; 449.0878; 425.0875; 423.0720; 287.0563.	Procyanidin tetramer (Cat-Cat-Cat-Cat)	2a	[38][d]
9.11	547.1600	0.18	$C_{30}H_{26}O_{10}$	+		393.0970; 273.0769	Propelargonidin dimer (Afz-Afz)	2b	
9.15	1409.3370	1.84	$C_{75}H_{62}O_{28}$	-		865.1984; 831.1897; 577.1367; 543.1305; 451.1024; 425.0887; 407.0768; 287.0656; 271.0616	Proanthocyanidin pentamer (Afz-Afz-Cat-Cat-Cat)	2b	
9.19	849.2046	2.47	$C_{45}H_{38}O_{17}$	-		831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
9.28	1137.2676	1.49	$C_{60}H_{50}O_{23}$	-		1011.2367; 967.2087; 865.1984; 695.1413; 577.1357; 559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2a	[38][d]
9.52	867.2120	−1.27	$C_{45}H_{38}O_{18}$	+		697.1535; 579.1495; 427.1028; 409.0930; 289.0710; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
9.85	1441.3282	2.78	$C_{75}H_{62}O_{30}$	-		1153.2584; 865.1946; 739.1681; 695.1397; 575.1201; 449.1038; 431.0983; 287.0567; 269.0468; 125.0232	Procyanidin pentamer (Cat-Cat-Cat-Cat-Cat)	2a	[38][d]
10.00	867.2125	−0.69	$C_{45}H_{38}O_{18}$	+		697.1535; 579.1495; 427.1028; 409.0930; 289.0710; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
10.04	1393.3391	−0.22	$C_{75}H_{62}O_{27}$	-	1123.2871	409.0909; 287.0554; 163.0392; 139.0391; 127.0393; 123.0443		2b	

Table 3 (continued)

t_R (min)	$[M+H]^+$/ $[M-H]^-$	Δ (ppm)	Molecular formula	Polarity	$[M+H]^+$ confirming peaks in MS[a]	MS^2 fragments[b]	Identity	Identification confidence level[c]	Reference
							Proanthocyanidin pentamer (Afz-Afz-Afz-Cat-Cat)		
10.19	1137.2675	1.41	$C_{60}H_{50}O_{23}$	-	849.2931; 561.1397	1011.2367; 967.2087; 865.1984; 695.1413; 577.1357;559.1246; 287.0564; 271.0613	Proanthocyanidin tetramer (Afz-Cat-Cat-Cat)	2a	[38][d]
10.52	867.2124	−0.81	$C_{45}H_{38}O_{18}$	+		697.1535; 579.1495; 427.1028; 409.0930; 289.0710; 127.0392	Procyanidin trimer (Cat-Cat-Cat)	2a	[13]
10.60	1393.3391	−0.22	$C_{75}H_{62}O_{27}$	-		409.0909; 287.0554; 163.0392; 139.0391; 127.0393; 123.0443	Proanthocyanidin pentamer (Afz-Afz-Afz-Cat-Cat)	2b	
10.85	561.1398	1.25	$C_{30}H_{26}O_{11}$	-		543.1305; 435.1083; 407.0772; 289.0717; 271.0615	Proanthocyanidin dimer (Afz-Cat)	2a	[38][d]
11.00	833.2105	3.48	$C_{45}H_{38}O_{16}$	-		707.1757; 561.1394; 543.1241; 435.1081; 407.0775; 289.0719; 271.0613	Proanthocyanidin trimer (Afz-Afz-Cat)	2a	[38][d]
11.60	849.2041	1.88	$C_{45}H_{38}O_{17}$	-	577.1354	831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
11.88	1123.2864	−0.18	$C_{60}H_{50}O_{22}$	+		851.2078;715.1606; 681.1587; 579.1505; 409.0923; 289.0709; 273.0761	Proanthocyanidin tetramer (Cat-Cat-Afz-Afz)	2b	
12.00	849.2041	1.06	$C_{45}H_{38}O_{17}$	-	561.1397	831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
12.30	849.2029	−0.35	$C_{45}H_{38}O_{17}$	-	561.1397	831.1924; 723.1717; 697.1569; 577.1354; 559.1240; 451.1033; 433.0922; 407.0764; 287.0560; 271.06; 125.0234	Proanthocyanidin trimer (Afz-Cat-Cat)	2a	[13]
12.68	819.2288	0.61	$C_{45}H_{38}O_{15}$	+	547.1572	819.2288; 547.1592; 273.075; 421.0907; 285.0756	Proanthocyanidin trimer (Afz-Afz-Afz)	2b	
13.22	1123.2862	−0.36	$C_{60}H_{50}O_{22}$	+		851.2078; 715.1606; 681.1587; 579.1505; 409.0923; 289.0709; 273.0761	Proanthocyanidin tetramer (Cat-Cat-Afz-Afz)	2b	
14.21	1089.2834	1.01	$C_{60}H_{50}O_{20}$	-	817.2132	815.1866; 545.1450; 543.1297; 407.0765; 271.0614; 300.9996; 125.0235	Propelargonidin tetramer (Afz-Afz-Afz-Afz)	2b	
16.10	547.1608	1.64	$C_{30}H_{26}O_{10}$	+		393.0970; 273.0769	Propelargonidin dimer (Afz-Afz)	2b	

Afz afzelechin, *Cat* catechin

[a] Confirming peaks are ions found in MS because of spontaneous in-source fragmentation

[b] MS^2 fragments are referred to acquisition with HCD set to an NCE of 20.

[d] Compound not found in strawberry extracts but found in strawberry secondary products or in other fruits and vegetables

flavonoids, phenolic acids, and dihydrochalcones are listed in Table 1.

The interpretation of the spectra of glycosylated flavonoids and dihydrochalcones reported in Table 1 needs further explanation. The SID fragments of the polyglycosylated compounds were used to speculate on the position of the sugar moieties in the aglycone. The precursor ion $[M + H]^+$ at *m/z* 741.2233 ($C_{33}H_{41}O_{19}^+$) at a retention time of 1.70 min, for example, was identified as pelargonidin hexosyl-deoxyhexosyl hexoside. At the same retention time, $[M + H - \text{hexose-deoxyhexose}]^+$ at *m/z* 433.1129 and $[M + H - \text{hexose-deoxyhexose-hexose}]^+$ at *m/z* 271.0600 were found, suggesting that the first sugar of the glycosyl moiety linked to the aglycone is a hexose, or that a hexose is linked to a hydroxyl group, and the dimer hexose–deoxyhexose is linked to another hydroxyl group [35]. Detailed information on the other polyglycosylated flavonoids and dihydrochalchones is given in the electronic supplementary material.

Interpretation of ellagitannin spectra

Ellagitannins, commonly found in strawberries [13], are esters of hexahydroxydiphenic acid (HHDP) and glucose (in some cases gallic acid). Typical neutral losses during fragmentation of ellagitannins are a galloyl unit (152 u), gallic acid (170 u), HHDP (302 u), galloyl-glucose (332 u), HHDP-glucose (482 u), and galloyl-HHDP-glucose (634 u) [13]. Sometimes, two often reported fragment ions—*m/z* 303 (+)/301 (−) and *m/z* 277 (+)/ 275 (−)—are observed. The first ion is the result of spontaneous dilactonization of HHDP to ellagic acid. The second ion is the result of the partial lactonization of HHDP followed by subsequent decarboxylation. Often the ions *m/z* 171 (+)/169 (−) and *m/z* 127 (+)/125 (−), characteristic of gallic acid and its fragment, are detected [38]. Ellagitannins were identified on the basis of the typical losses in the MS^2 scan and reported in Table 2. The detected in-source fragments of the ellagitannins were used to further confirm their structure. Detailed explanations of the interpretation of ellagitannins spectra are given in the electronic supplementary material.

Interpretation of proanthocyanidin spectra

Proanthocyanidins are polymers of flavanols, mainly catechin, afzelechin, and gallocatechin, forming procyanidins, propelargonidins, and prodelphinidins respectively. As reported by Lin et al. [39], the most important MS^2 fragments of proanthocyanidins are formed by quinone methide fission: that is, the breaking of the interflavan bond between the monomers to form the terminal unit fragment $[M_T - H]^-$ or $[M_T + H]^+$ ions and the extension unit fragment $[M_E - 3H]^-$ or $[M_E + 3H]^+$ ions for B-type proanthocyanidins. Other typical proanthocyanidin fragments are formed by retro-Diels–Alder fission—that is, loss of the B–ring with C2 − C3 part of the C-ring (loss of 152, 136, and 168 u for catechin, afzelechin, and gallocatechin respectively)—and by heterocyclic ring fission—that is, loss of the A-ring (loss of 126 u). Other product ions formed by the loss of H_2O, CO, etc., are also observed. On he basis of the typical fragmentation pattern, 55 proanthocyanidins were tentatively identified in strawberries and one, procyanidin B1, was identified by the relative standard confirmation. However, many of them were found at several retention times because proanthocyanidins have many regioisomeric and stereoisomeric forms. Each flavan-3-ol unit has two chiral centers, which can result in four stereoisomers. Moreover, units can connect to each other in two ways, providing 64 possible isomers for a B-type proanthocyanidin dimer. The formation of trimers and tetramers leads to an increase in the possible number of isomers that, at present, are indistinguishable by means of $LC–MS^n$ methods [39]. All isomeric forms of the tentatively identified proanthocyanidins are reported in Table 3. The SID fragments of proanthocyanidins were used to confirm their structures. Details on the interpretation of the spectra details are given in the electronic supplementary material.

Conclusion

This is the first study in which a systematic investigation of both chromatographic and mass-spectrometric conditions was performed to optimize the untargeted analysis of a complex phytochemical mixture. Exploring the capability of the coupling of an UHPLC system to an $HRMS^2$ system allowed us to select the best chromatographic and mass-spectrometric parameters for a comprehensive profiling of a strawberry extract. In contrast to previous work on the identification of strawberry polyphenol compounds, the identification and tentative identification of 18 and 113 compounds respectively was achieved. The UHPLC–HRMS analysis identified the presence of a variety of anthocyanins, flavan-3-ols, dihydroflavonols, proanthocyanidins, and ellagitannins. The tentative identification of phenolic compounds revealed 74 interesting compounds newly found in strawberry and, as far as we know, 22 of these compounds have never been reported before. The results clearly show the importance of an extensive optimization of chromatographic and mass-spectrometric methods before the untargeted profiling of complex phytochemical mixtures, suggesting this meticulous approach should be used in all untargeted metabolomics analyses.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

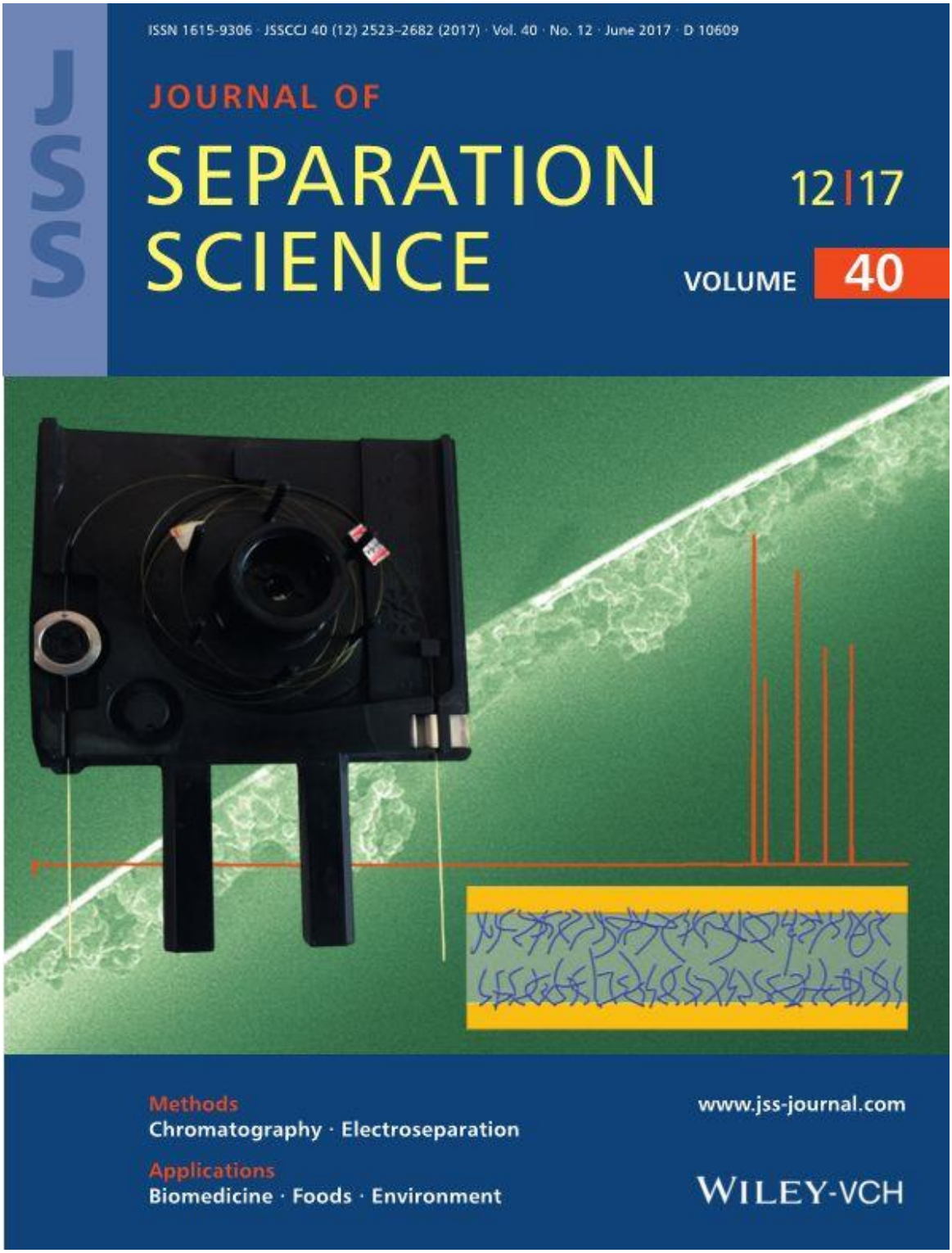
References

1. Gika HG, Wilson ID, Theodoridis GA. LC–MS-based holistic metabolic profiling. Problems, limitations, advantages, and future perspectives. J Chromatogr B. 2014;966:1–6. doi:10.1016/j.jchromb.2014.01.054.
2. Kaufmann A. Combining UHPLC and high-resolution MS: a viable approach for the analysis of complex samples? Trends Anal Chem. 2014;63:113–28. doi:10.1016/j.trac.2014.06.025.
3. Riedl J, Esslinger S, Fauhl-Hassek C. Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. Anal Chim Acta. 2015;885:17–32. doi:10.1016/j.aca.2015.06.003.
4. Kim S, Kim J, Yun EJ, Kim KH. Food metabolomics: from farm to human. Curr Opin Biotechnol. 2016;37:16–23. doi:10.1016/j.copbio.2015.09.004.
5. Schuhmacher R, Krska R, Weckwerth W, Goodacre R. Metabolomics and metabolite profiling. Anal Bioanal Chem. 2013;405:5003–4. doi:10.1007/s00216-013-6939-5.
6. Zhang R, Watson DG, Wang L, Westrop GD, Coombs GH, Zhang T. Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based untargeted metabolite profiling of Leishmania parasites. J Chromatogr A. 2014;1362:168–79. doi:10.1016/j.chroma.2014.08.039.
7. Stobiecki M, Kachlicki P, Wojakowska A, Marczak Ł. Application of LC/MS systems to structural characterization of flavonoid glycoconjugates. Phytochem Lett. 2015;11:358–67. doi:10.1016/j.phytol.2014.10.018.
8. Kind T, Fiehn O. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. BMC Bioinformatics. 2007;8:105. doi:10.1186/1471-2105-8-105.
9. Mullard G, Allwood JW, Weber R, Brown M, Begley P, Hollywood KA, et al. A new strategy for MS/MS data acquisition applying multiple data dependent experiments on Orbitrap mass spectrometers in non-targeted metabolomic applications. Metabolomics. 2015;11:1068–80. doi:10.1007/s11306-014-0763-6.
10. Kloos D-P, Lingeman H, Niessen WMA, Deelder AM, Giera M, Mayboroda OA. Evaluation of different column chemistries for fast urinary metabolic profiling. J Chromatogr B. 2013;927:90–6. doi:10.1016/j.jchromb.2013.02.017.
11. Zhang T, Creek DJ, Barrett MP, Blackburn G, Watson DG. Evaluation of coupling reversed phase, aqueous normal phase, and hydrophilic interaction liquid chromatography with Orbitrap mass spectrometry for metabolomic studies of human urine. Anal Chem. 2012;84:1994–2001. doi:10.1021/ac2030738.
12. Bajad SU, Lu W, Kimball EH, Yuan J, Peterson C, Rabinowitz JD. Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. J Chromatogr A. 2006;1125:76–88. doi:10.1016/j.chroma.2006.05.019.
13. Sun J, Liu X, Yang T, Slovin J, Chen P. Profiling polyphenols of two diploid strawberry (Fragaria vesca) inbred lines using UHPLC-HRMSn. Food Chem. 2014;146:289–98. doi:10.1016/j.foodchem.2013.08.089.
14. Kårlund A, Hanhineva K, Lehtonen M, Karjalainen RO, Sandell M. Nontargeted metabolite profiles and sensory properties of strawberry cultivars grown both organically and conventionally. J Agric Food Chem. 2015;63:1010–9. doi:10.1021/jf505183j.
15. Kajdžanoska M, Petreska J, Stefova M. Comparison of different extraction solvent mixtures for characterization of phenolic compounds in strawberries. J Agric Food Chem. 2011;59:5272–8. doi:10.1021/jf2007826.
16. Pluskal T, Castillo S, Villar-Briones A, Orešič M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics. 2010;11:395. doi:10.1186/1471-2105-11-395.
17. Benton HP, Ivanisevic J, Mahieu NG, Kurczy ME, Johnson CH, Franco L, et al. Autonomous metabolomics for rapid metabolite identification in global profiling. Anal Chem. 2015;87:884–91. doi:10.1021/ac5025649.
18. Willemse CM, Stander MA, de Villiers A. Hydrophilic interaction chromatographic analysis of anthocyanins. J Chromatogr A. 2013;1319:127–40. doi:10.1016/j.chroma.2013.10.045.
19. de Villiers A, Cabooter D, Lynen F, Desmet G, Sandra P. High performance liquid chromatography analysis of wine anthocyanins revisited: effect of particle size and temperature. J Chromatogr A. 2009;1216:3270–9. doi:10.1016/j.chroma.2009.02.038.
20. Álvarez-Fernández MA, Hornedo-Ortega R, Cerezo AB, Troncoso AM, García-Parrilla MC. Determination of nonanthocyanin phenolic compounds using high-resolution mass spectrometry (UHPLC-Orbitrap-MS/MS) and impact of storage conditions in a beverage made from strawberry by fermentation. J Agric Food Chem. 2016;64:1367–76. doi:10.1021/acs.jafc.5b05617.
21. McDougall G, Martinussen I, Stewart D. Towards fruitful metabolomics: high throughput analyses of polyphenol composition in berries using direct infusion mass spectrometry. J Chromatogr B. 2008;871:362–9. doi:10.1016/j.jchromb.2008.06.032.
22. Newsome AG, Li Y, van Breemen RB. Improved quantification of free and ester-bound gallic acid in foods and beverages by UHPLC-MS/MS. J Agric Food Chem. 2016;64:1326–34. doi:10.1021/acs.jafc.5b04966.
23. Álvarez-Fernández MA, Cerezo AB, Cañete-Rodríguez AM, Troncoso AM, García-Parrilla MC. Composition of nonanthocyanin polyphenols in alcoholic-fermented strawberry products using LC–MS (QTRAP), high-resolution MS (UHPLC-Orbitrap-MS), LC-DAD, and antioxidant activity. J Agric Food Chem. 2015;63:2041–51. doi:10.1021/jf506076n.
24. Hanhineva K, Rogachev I, Kokko H, Mintz-Oron S, Venger I, Kärenlampi S, et al. Non-targeted analysis of spatial metabolite composition in strawberry (Fragaria × ananassa) flowers. Phytochemistry. 2008;69:2463–81. doi:10.1016/j.phytochem.2008.07.009.
25. D'Urso G, d'Aquino L, Pizza C, Montoro P. Integrated mass spectrometric and multivariate data analysis approaches for the discrimination of organic and conventional strawberry (Fragaria ananassa Duch.) crops. Food Res Int. 2015;77:264–72. doi:10.1016/j.foodres.2015.04.028.
26. Spínola V, Pinto J, Castilho PC. Identification and quantification of phenolic compounds of selected fruits from Madeira Island by HPLC-DAD–ESI-MSn and screening for their antioxidant activity. Food Chem. 2015;173:14–30. doi:10.1016/j.foodchem.2014.09.163.
27. Aaby K, Mazur S, Nes A, Skrede G. Phenolic compounds in strawberry (Fragaria x ananassa Duch.) fruits: composition in 27 cultivars and changes during ripening. Food Chem. 2012;132:86–97. doi:10.1016/j.foodchem.2011.10.037.
28. Cerezo AB, Cuevas E, Winterhalter P, Garcia-Parrilla MC, Troncoso AM. Isolation, identification, and antioxidant activity of anthocyanin compounds in Camarosa strawberry. Food Chem. 2010;123:574–82. doi:10.1016/j.foodchem.2010.04.073.

29. Lopes-da-Silva F, de Pascual-Teresa S, Rivas-Gonzalo J, Santos-Buelga C. Identification of anthocyanin pigments in strawberry (cv Camarosa) by LC using DAD and ESI-MS detection. Eur Food Res Technol. 2002;214:248–53. doi:10.1007/s00217-001-0434-5.
30. Sadilova E, Carle R, Stintzing FC. Thermal degradation of anthocyanins and its impact on color and in vitro antioxidant capacity. Mol Nutr Food Res. 2007;51:1461–71. doi:10.1002/mnfr.200700179.
31. IM de Tavares C, Lago-Vanzela ES, Rebello LPG, Ramos AM, Gómez-Alonso S, García-Romero E, et al. Comprehensive study of the phenolic composition of the edible parts of jambolan fruit (Syzygium cumini (L.) Skeels). Food Res Int. 2016;82:1–13. doi:10.1016/j.foodres.2016.01.014.
32. Carazzone C, Mascherpa D, Gazzani G, Papetti A. Identification of phenolic constituents in red chicory salads (*Cichorium intybus*) by high-performance liquid chromatography with diode array detection and electrospray ionisation tandem mass spectrometry. Food Chem. 2013;138:1062–71. doi:10.1016/j.foodchem.2012.11.060.
33. Álvarez-Fernández MA, Hornedo-Ortega R, Cerezo AB, Troncoso AM, García-Parrilla MC. Effects of the strawberry (*Fragaria ananassa*) purée elaboration process on non-anthocyanin phenolic composition and antioxidant activity. Food Chem. 2014;164:104–12. doi:10.1016/j.foodchem.2014.04.116.
34. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ Sci Technol. 2014;48:2097–8. doi:10.1021/es5002105.
35. Vukics V, Guttman A. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. Mass Spectrom Rev. 2010;29:1–16. doi:10.1002/mas.20212.
36. Cuyckens F, Claeys M. Mass spectrometry in the structural analysis of flavonoids. J Mass Spectrom. 2004;39:1–15. doi:10.1002/jms.585.
37. de Villiers A, Venter P, Pasch H. Recent advances and trends in the liquid-chromatography–mass spectrometry analysis of flavonoids. J Chromatogr A. 2016;1430:16–78. doi:10.1016/j.chroma.2015.11.077.
38. Hooi Poay T, Sui Kiong L, Cheng Hock C. Characterisation of galloylated cyanogenic glucosides and hydrolysable tannins from leaves of Phyllagathis rotundifolia by LC-ESI-MS/MS: characterisation of galloylated cyanogenic glucosides and tannins. Phytochem Anal. 2011;22:516–25. doi:10.1002/pca.1312.
39. Lin L-Z, Sun J, Chen P, Monagas MJ, Harnly JM. UHPLC-PDA-ESI/HRMS^n profiling method to identify and quantify oligomeric proanthocyanidins in plant products. J Agric Food Chem. 2014;62:9387–400. doi:10.1021/jf501011y.

3 Paper IV

Evaluation of column length and particle size effect on the untargeted profiling of a phytochemical mixture by using UHPLC coupled to high-resolution mass spectrometry.


ISSN 1615-9306 · JSSCCJ 40 (12) 2523–2682 (2017) · Vol. 40 · No. 12 · June 2017 · D 10609
JOURNAL OF
SEPARATION SCIENCE
12|17
VOLUME 40
Methods
Chromatography · Electroseparation
Applications
Biomedicine · Foods · Environment
www.jss-journal.com
WILEY-VCH

Received: 5 February 2017 | Revised: 6 April 2017 | Accepted: 12 April 2017

DOI: 10.1002/jssc.201700135

RESEARCH ARTICLE

JOURNAL OF SEPARATION SCIENCE

Evaluation of column length and particle size effect on the untargeted profiling of a phytochemical mixture by using UHPLC coupled to high-resolution mass spectrometry

Riccardo Zenezini Chiozzi | Anna Laura Capriotti | Chiara Cavaliere | Francesca Ferraris | Giorgia La Barbera | Susy Piovesana | Aldo Laganà

Department of Chemistry, University of Rome "La Sapienza", Rome, Italy

Correspondence
Dr. Giorgia La Barbera, Dipartimento di Chimica, Università di Roma "La Sapienza" Piazzale Aldo Moro 5, Rome 00185, Italy.
Email: giorgia.labarbera@uniroma1.it

Liquid chromatography coupled to high-resolution mass spectrometry is the technique of choice for the untargeted profiling of food matrices. Despite the high potential of high-resolution mass spectrometry, when dealing with complex mixtures, an efficient separation technique is also needed. The novel core-shell chromatographic columns packed with sub-2 µm sized particles are claimed to show very good resolution. However, the analytes retention can be significantly altered when working under ultra-high performance chromatographic conditions. In this work, an evaluation of four chromatographic systems, with either a single or two in-series Kinetex™ C_{18} columns, either packed with 2.6 or 1.7 µm particles, is presented for the targeted analysis of a standard mixture and the untargeted analysis of a strawberry extract. An ultra-high performance chromatographic system coupled via an electrospray source to a hybrid quadrupole-Orbitrap mass spectrometer was used. From the extensive comparison, a surprising result was obtained, namely, that the system identifying the largest number of features was the one with two in-series connected columns with the larger particle size. The inconsistency among the theoretical assumptions and the applicative findings points out the importance of an extensive chromatographic evaluation for the comprehensive untargeted profiling of complex real samples.

KEYWORDS
core-shell particles, high-resolution mass spectrometry, strawberry, untargeted profiling

1 | INTRODUCTION

Over the past few years, metabolomics has found applications in several fields, such as in life sciences, environment, and food [1–4]. Recently, metabolomics is playing an increasingly important role in food quality assessment, due to the strong relationship between the bioactive compounds found in food and their positive health effects [5]. LC–MS is suitable for qualitative and quantitative measurements of individual metabolites, and has been widely employed in natural product analysis [4,6–9]. However, only few of such investigations rely on an untargeted molecular approach due to the high complexity of natural products matrices. Some critical issues, such as the concomitant presence of thousands of metabolites with high chemodiversity and the lack of available secondary metabolites databases, make the natural products untargeted profiling a challenging task [10]. Efficient analytical protocols are thus required for both exhaustive detection and identification of novel bioactive compounds.

High-resolution mass spectrometry (HRMS) provides high sensitivity together with the possibility to distinguish

Abbreviation: HRMS, high-resolution mass spectrometry

Conflict of interest: The authors declare that they have no conflict of interest.

compounds by their accurate *m/z* signals and to study their fragmentation patterns in tandem HRMS mode. The Orbitrap mass spectrometer, in particular, is able to provide excellent resolution and mass accuracy <2 ppm [11].

Despite the high potentialities of HRMS, a major issue in natural products analysis arises in the upstream process of compounds separation. A separation technique, such as UHPLC, should be coupled with HRMS to differentiate metabolites by their retention times and provide additional information for their identification. In the analysis of a complex system, an efficient separation technique is especially needed for the presence of a wide range of isomeric metabolites. Positional isomers present in complex phytochemical mixtures, such as glycoconjugated flavonoids [12], generally elute together because of the similarity of their chemical properties. The coelution of isomeric metabolites makes them indistinguishable and could generate ionization suppression of the less abundant compounds. Therefore, optimizing the chromatographic conditions in an untargeted analysis could be crucial for maximizing the number of detected and consequently identified compounds [13].

A critical point in the use of UHPLC–HRMS analytical platforms is the choice of a suitable chromatographic system. Among the several different high-efficiency chromatographic columns now commercially available, the particle morphology of the novel core-shell RP LC materials has been recently attracting an increasing interest. The main advantage of core-shell particles is a minor band broadening compared to fully porous particles with the same diameter that leads to an important efficiency improvement. Moreover, the possibility of operating at higher flow rates at manageable backpressures gives the opportunity to obtain higher peak capacities together with shorter run times. The much lower backpressure allows then to upgrade the efficiency and performance of the chromatographic system by using sub-2 µm particle packed columns. In this regard, the Kinetex™ C_{18} with 1.7 µm sized core-shell particles were the first of such columns shipped to the market, and are claimed to show better resolution, higher peak capacity, and greater efficiency compared to the Kinetex™ C_{18} 2.6 µm sized core-shell particles packed columns [14]. However, several studies showed that under UHPLC conditions the analytes retention can be significantly altered by pressure and frictional heating related events [15,16]. The effect of high pressures and high flow rates on sub-2 µm sized particles packed columns could give some unexpected results. Based on these considerations, evaluating from a practical point of view the performance of core-shell columns with different lengths and particle sizes operated under different pressure drops, could provide different information leading to the identification of a larger number of compounds and, in turn, to an improved comprehensive metabolic profiling.

To the best of our knowledge, no study, investigating the effect of column length and particle size on the profiling of complex phytochemical mixtures, was published, yet. To this end, we present a thorough evaluation of several chromatographic systems by means of an UHPLC system coupled via an ESI source to a hybrid quadrupole-Orbitrap mass spectrometer. The four systems tested were the following: RP C_{18} columns with 2.6 µm sized core-shell particles or with 1.7 µm sized core-shell particles, either single column or two connected in series. The performance of the selected column systems was examined employing a test mixture of 45 polyphenol standards in a targeted approach. The feasibility of the qualitative and quantitative analyses was evaluated by a deep investigation on peak shapes, retention times, and peak areas repeatability. Moreover, to provide more insight into the performance of the column systems, nontargeted profiling was carried out on strawberry extracts, due to their well-known complexity in isomeric antioxidant compounds content [13]. The chromatographic systems were evaluated based on the number of detected features and, finally, on the putative identification of some classes of polyphenol compounds.

2 | MATERIALS AND METHODS

2.1 | Chemicals and reagents

Strawberries (*Fragaria* × *ananassa* cv.elsanta) were bought in a local farm. The fruits were washed with distilled water, cut into quarters, freeze-dried and kept at −80°C until use. LC–MS-grade water was purchased from Fisher Scientific. LC–MS-grade methanol and acetonitrile were purchased by VWR International (Milan, Italy). Formic acid and acetone were purchased from Sigma–Aldrich (St. Louis, MO, USA). Analytical standards of apigenin, apigenin 7-*O*-glucoside, biochanin A, callistephin chloride (pelargonidin 3-*O*-glucoside), catechin gallate, chlorogenic acid hemihydrate, coumestrol, cyanidin chloride, daidzein, daidzin (daidzein 7-*O*-glucoside), diosmetin, epicatechin, epicatechin gallate, equol, flavone, formononetin, genistin (genistein 7-*O*-glucoside), glycitein, hesperetin, kaempferol, kuromanin chloride (cyanidin 3-*O*-glucoside), malvidin galactoside, myricetin, naringenin, pelargonidin chloride, peonidin 3-*O*-glucoside chloride, procyanidin B1, procyanidin B2, primuletin, quercetin dihydrate, quercetin 3-*O*-glucoside, resveratrol, rutin (quercetin 3-*O*-rutinoside), taxifolin, and trihydroxyisoflavone were purchased from Sigma–Aldrich. Analytical standards of catechin, eriodictyol, genistein, hesperidin (hesperetin 7-*O*-rutinoside), isorhamnetin, luteolin, luteolin 7-*O*-glucoside, malvidin chloride, phloretin, and syringetin were purchased from Extrasynthese (Lyon, France).

2.2 | Standard sample preparation

Authentic standard stocks and the mix of 45 standards were prepared in H_2O/MeOH (50:50, v/v) and diluted to H_2O/MeOH (90:10, v/v) to obtain a concentration of 0.5 ng/µL before LC–MS analysis.

2.3 | Sample preparation

Phenolic compounds were extracted following the protocol by Kajdžanoska et al. [17] with some modifications. Three hundred milligrams of freeze-dried strawberry sample was extracted in 9 mL of CH_3COCH_3/H_2O/CH_3COOH (70:29.5:0.5, v/v/v). The extract was sonicated for 15 min in an ice-bath and then centrifuged for 10 min at 2000 × *g* at 15°C. The supernatant was collected. The extraction was repeated once. The supernatants were mixed and concentrated to 2.5 mL under a gentle nitrogen flow in a water bath at 37°C. The sample was then added with 250 µL of methanol. The final extract solution H_2O/MeOH (90:10, v/v) was centrifuged for 5 min at 2500 × *g* at 25°C. Before analysis, the supernatant was filtered through 13 mm Acrodisc syringe filter with 0.2 µm GHP membrane (Pall, Ann Arbor, MI, USA). The extract was divided in aliquots and stored at −20°C for further analysis.

2.4 | UHPLC–ESI-MS/MS analysis

The UHPLC system was a Vanquish binary pump H (Thermo Fisher Scientific, Bremen, Germany), equipped with a thermostatted autosampler and a thermostatted column compartment. The separation was carried out on four different chromatographic systems, all constituted by Kinetex™ core-shell XB-C_{18} (100 Å pore size, Phenomenex, Torrance, CA, USA) packing material: (a) one column (100 × 2.1 mm^2 id) with 2.6 µm particles size; (b) one column (100 × 2.1 mm^2 id) with 1.7 µm particles size; (c) two columns (100 × 2.1 mm^2 id) with 2.6 µm particles size connected by means of a virtually zero dead volume UHPLC fingertight fitting (Thermo Scientific); and (d) two columns (100 × 2.1 mm^2 id) with 1.7 µm particles size connected by means of a virtually zero dead volume UHPLC fingertight fitting. All columns were operated at a flow rate of 600 µL/min.

The chromatographic method for systems (a) and (b) was chosen based on the work of La Barbera et al. [13] Phase B gradient was the following: 5–15% in 10 min, 15–35% in 15 min, 35–50% in 5 min, at 40°C (still air option), with ACN/HCOOH (99.9:0.1, v/v) as phase B and H_2O/HCOOH (99.9:0.1, v/v) as phase A. The same conditions were kept for systems (c) and (d) but with the following gradient: phase B 5–15% in 20 min, 15–35% in 30 min, and 35–50% in 10 min. For each condition, there is included a pre-equilibration step of 5% B in 3 min, a washing step of 99% B in 5 min at the end of the gradient, and a re-equilibration step of 5% B in 8 min after the washing step. The injection volume was 5 µL for both the strawberry extract and the standard mix. Both the strawberry extract and the standard mix were run in triplicate followed by the blank injection H_2O/MeOH (90:10, v/v).

The chromatographic system previously described was coupled using a heated ESI source to a hybrid quadrupole-Orbitrap mass spectrometer QExactive (Thermo Fisher Scientific). ESI source parameters in positive modes were as follows: capillary temperature 275°C; sheath gas 50 (arbitrary units); auxiliary gas 15 (arbitrary units); sweep gas 2.25 (arbitrary units); spray voltage 3500 V; auxiliary gas heater temperature 450°C; S-Lens RF level 50 (%). The source was cleaned before running the mixture of standards and the strawberry samples for each column setup. The ionization efficacy was monitored by running a mix of standards before and after the samples for each chromatographic setup.

For all the investigated chromatographic conditions, the detection was conducted in HRMS data-dependent acquisition mode, setting the appropriate parameters according to La Barbera et al. [13]. Mass spectrometer data were acquired in the range *m/z* 200–1200 in positive ionization mode with a resolution (full width at half maximum, *m/z* 200) of 35 000. Automatic gain control target value was 200 000 in full scan. Max ion injection time was 100 ms. Isolation window width was 2 *m/z*. MS/MS fragmentation was performed on the five most intense ions detected in full scan with a resolution (full width at half maximum, @*m/z* 200) of 17 500. Automatic gain control target value was 100 000. Dynamic exclusion was set to 3 s. An exclusion list containing the most common ions detected in the blank was set. Fragmentation was achieved in the higher energy collisional-dissociation cell at a normalized collision energy of 80%.

2.5 | Data analysis

2.5.1 | Standard mix data analysis

The raw data obtained from the three replicates of the standard mix solution were visualized on Qual Browser of Xcalibur 2.2 (Thermo Fisher Scientific). The average and the SD of each peak were calculated, and the peak shape of each compound was observed to establish the goodness of the four chromatographic systems.

2.5.2 | Strawberry extracts data analysis

2.5.2.1 | Features extraction

For each chromatographic system, accurate mass ion chromatograms obtained from three consecutive injections of strawberry extract and blank were processed by the software MZmine version 2.19, an open-source software (http://mzmine.sourceforge.net/) usually employed in the preprocessing of metabolomics data [18]. Briefly, the accurate-mass LC–MS data were imported, and a list of ions for each

scan was generated using a mass detection module (Exact mass algorithm with noise 1E4). A filter was applied to correct the false signals around *m/z* peaks, called "shoulder peaks" that are residues of the Fourier Transform function (Gaussian peak model function). Then, mass lists generated and filtered for each mass spectrometer scan were used to build a chromatogram for each mass that could be detected continuously over the scans by means of the chromatogram builder tool (minimum time span 0.1 min; minimum height 1E5; *m/z* tolerance 0.001 or 2 ppm). Later, the deconvolution module deconvoluted these chromatograms into individual peaks (local minimum search algorithm with chromatographic threshold 30; minimum retention time 0.05 min for systems (a) and (b), 0.1 min for systems (c) and (d); minimum relative height 0%; minimum absolute height 1E5; minimum top/edge 1.3; peak duration 4 min). With the isotopic peak grouper algorithms, the isotopic peaks were removed from the peak list keeping only the highest peak (*m/z* tolerance 0.001 or 2 ppm; retention time tolerance 0; maximum charge 2; most intense representative isotope). The peak lists obtained for the replicates and the blank were normalized for retention time (*m/z* tolerance 0.001 or 5 ppm; retention time tolerance 0.05 min; minimum intensity standard 10^6) and aligned with the Ransac alignment module (*m/z* tolerance 0.001 or 5 ppm; retention time tolerance 0.5 min; retention time after correction 0.5 min; Ransac iteration 0; minimum number of points 20; threshold value 4). Peaks coming from the blank were subtracted. Only peaks common to all three replicates were maintained.

2.5.2.2 | Identification

The first identification step consisted in the application of the Fragment search module (Retention time tolerance 0.0 min; *m/z* tolerance of ms2 data 0.001 or 5 ppm; maximum fragment peak height 10000%; minimum ms2 peak height 5E3). This module identifies fragment peaks present in the mass spectrometer spectra due to the "in source fragmentation" process [19]. In-source fragments are identified by two conditions: (a) the retention time of the original ion and of the in-source fragment ion should be the same; and (b) the MS/MS pattern of the original ion must contain a peak with the same mass as the in-source fragment ion. Such in-source fragment ions were removed from the list of the detected features because they do not represent a compound but they were used as additional information to identify their precursor.

The second identification step consisted in the Adduct search ($[M+Na]^+$, $[M+K]^+$, $[M+acetonitrile+H]^+$, retention time tolerance 0.0 min; *m/z* tolerance 0.001 or 5 ppm; maximum adduct peak height 100%). This module identifies adducts if the retention time of the original ion and the adduct ion are the same.

The third identification step consisted in the Complex search (ionization mode $[M+H]^+$, retention time tolerance 0.0 min; *m/z* tolerance 0.001 or 5 ppm; maximum complex peak height 100%). This module attempts to identify ion complexes, i.e. pairs of ions, which appear together at the same retention time such as dimers adducts.

Finally, the obtained features were searched against a custom database (*m/z* tolerance 0.001 or 5 ppm, retention time not set) created by looking through the typical reported compounds in fruit and vegetable matrices and combining all the possible conjugated compounds which might be expected. The tentatively identified compounds were filtered by considering retention time and MS/MS spectra information following the rules reported in Sections 3.2 and 3.3.

3 | RESULTS AND DISCUSSION

In this work, an evaluation of the effect of column length and particle size on the analysis of complex phytochemical mixtures is proposed by means of both a targeted and an untargeted approach on a standard mixture and on a strawberry extract, respectively. The chromatographic and mass spectrometric method, previously optimized by La Barbera et al. [13] on a single Kinetex™ XB-C_{18} column with 2.6 µm core-shell sized particle, was kept unchanged for systems (a) and (b) and was adapted to systems (c) and (d) by doubling gradient time. The temperature was set to 40°C and the flow rate was set to 600 µL/min for all the chromatographic systems to achieve the best efficiency, as suggested by the specification sheet. The backpressures in the chromatographic starting condition (5% B) were 450, 600, 800, and 1100 bar for the systems (a), (b), (c), and (d), respectively.

3.1 | Chromatographic evaluation on standards solution

The mixture of standards contained a range of flavonoids, glycosylated flavonoids, anthocyanins, glycosylated anthocyanins, and proanthocyanidins, chosen because representative of the phenolic compounds usually found in vegetable matrices and, particularly, in red fruits samples. The mixture was lacking some glycosylated conjugates, proanthocyanidins, and ellagitannins, since they are not commercially available. For each compound, the repeatability of the retention time and the peak area were calculated, and an additional visual inspection of peak shapes was performed for each chromatographic condition. Finally, all metabolites were categorized into three groups based on the indication of Zhang et al. [20]. The peaks relative to each compound were classified as "unacceptable" if the shift of retention times and peak area was >25% and if the peak shape was affected by broadening. Peaks were classified as "acceptable" if the shift of retention time and peak area was between 10 and 25% and if the peak was splitting or tailing. Finally, peaks were classified as "good" if the shift of retention time and peak area

FIGURE 1 Classification of polyphenolic standards into good, acceptable, and unacceptable levels for the four chromatographic conditions

was <10% and if they showed narrow and symmetric peaks. Figure 1 shows the results for each chromatographic system.

At first glance, the performance of all the chromatographic systems was poor for most anthocyanins. This can be explained by the typical peak broadening of anthocyanins due to the coexistence of different forms [21]. A more acidic phase and a higher temperature could positively affect their chromatographic behavior but, at the same time, negatively affect the other flavonoids retention [22,23]. Therefore, phase acidity and temperature were chosen as the best compromise for all the classes of compounds [13]. In these conditions, a very strong difference among the chromatographic systems could be appreciated for some flavonoids belonging to the class of flavonols, such as kaempferol, quercetin, isorhamnetin, myricetin, and syringetin. These compounds resulted unacceptable in terms of shape and repeatability in both the chromatographic systems with 1.7 µm sized particles and 2.6 µm sized particles ones. This result was really unexpected taking into account the general better performance of sub-2 µm chromatographic columns. Moreover, there was not such a difference between the one-column and the two-columns systems, regardless of the particle size. This result might suggest that the length of the column did not affect the polyphenols analysis; however, the relatively low complexity of the examined standard mixture should be taken into consideration and can be accounted for such result. To confirm or disprove the targeted analysis results and evaluate the effects on the analysis of a more complex matrix, such as a phytochemical mixture, we proceeded on to the untargeted approach investigation.

3.2 | Chromatographic evaluation on the strawberry extract

To achieve a reasonable evaluation for the purpose of untargeted comprehensive polyphenolic profiling the chromatographic systems were compared at several levels, as shown in Fig. 2. The number of features (Section 2.5.2.1) represents the number of peaks characterized by a unique *m/z* and retention time and that probably correspond to real molecules. From this analysis, 20 121 and 19 831 features were detected for the single and double 1.7 µm particle size columns set up against 17 783 and 17 790 detected features for the 2.6 µm particle sized one, respectively. In this regard, the results pointed out that the total number of detected features was larger for the chromatographic systems with 1.7 µm particle size than for the 2.6 µm particle sized ones, regardless of the column length (Fig. 2A). To achieve reliable chromatographic results, however, the repeatable features were further taken into account. In fact, in untargeted approaches, when dealing with thousands of unknown signals, it is necessary to consider the reproducibility of the whole UHPLC–HRMS system, to select and study only the signals corresponding to the analytes. The repeatable features were calculated by filtering the features with an RSD > 25% of the integrated peak areas. After application of this repeatability filter, 30, 43, 19, and 31% of the detected features were removed from the features ascribed to the systems (a), (b), (c), and (d), respectively (Fig. 2A). Thus, the feature number sensibly decreased for all the conditions. However, both the chromatographic systems (b) and (d) showed a more remarkable decrease than the corresponding systems (a) and (c). The dramatic cutoff of detected features for 1.7 µm sized particles columns could be ascribed to the inaccurate integration of chromatographic peaks with poor shapes, suggesting that, regardless of the column lengths, the 1.7 µm sized particles columns performed worse than the 2.6 µm sized particles columns, as far as the peak shapes and repeatability was concerned. One final consideration about column coupling: other than for the simple detected features, where no significant improvement could be appreciated by a longer columns setup, in the case of repeatable features the coupling of two columns added 1913 and 2229 detected features for 2.6 µm particle sized columns and 1.7 µm particle

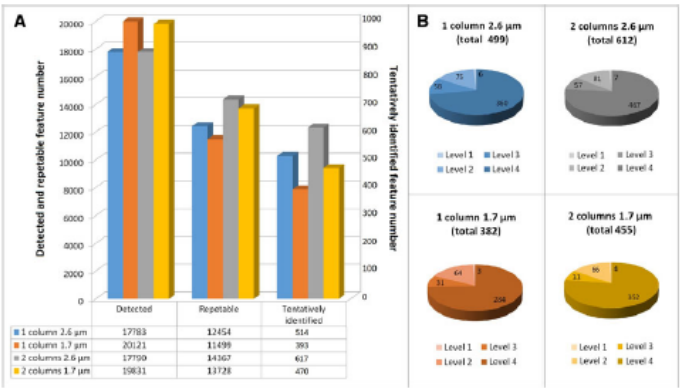


FIGURE 2 (A) Number of detected, repeatable, and tentatively identified features for the four chromatographic conditions; and (B) tentatively identified features after manual inspection with relative identification confidence level for the four chromatographic conditions

sized columns, respectively. Thus, even though the total number of features was similar for the single-column and double-columns setups at the detection level, repeatability greatly benefits from a longer column setup.

To confirm these results, however, the different systems were evaluated at a third more reliable level, that of the tentatively identified features. In fact, the feature number is not free from misinterpretations. Many false-positive features can be caused both by poor chromatographic signals and by mass spectrometer signals which are not generated by proton adducts, such as the case of complex ions (i.e. dimers $2M \pm H$), salt cluster ions, and in-source fragments. In this study, nonproton adducts complex ions and in-source fragments were identified using MZmine (Section 2.5.2.2). Such nonmolecular ion features were removed and assigned to the related molecular ion. The remaining features were searched against a home-made database resulting in the tentative identification of 514, 393, 617, and 470 compounds for systems (a), (b), (c) and (d), respectively. Even at this level, the 2.6 µm sized particle columns performed better than the 1.7 µm sized particle columns, and the performance could be improved by the coupling of two columns.

Finally, the compounds identity was further investigated by a closer inspection of the mass spectrometer spectra, the MS/MS spectra, and the retention times. Figure 2B shows the total number of the tentatively identified features and their classification into four different identification confidence levels, according to Schymanski et al. [24]. In particular, level 1 refers to compounds identified by checking standards retention time and MS/MS spectra; level 2 to compounds tentatively identified by checking MS/MS spectra data in the literature, in on-line spectral libraries or by studying diagnostic fragments in MS/MS spectra; level 3 to compounds identified by their accurate mass and some fragments, which are not sufficient for an univocal assignment; level 4 to compounds identified by accurate mass allowing to assign unambiguously a molecular formula. As shown in Fig. 2A, the number of the total tentatively identified features (514, 393, 617, and 470 for systems a, b, c, and d, respectively) is larger compared to the total number of compounds identified after the visual inspection (499, 382, 612, and 455 for systems a, b, c, and d, respectively, shown in Fig. 2B). In fact, after a closer inspection of raw data, it was possible to identify several false metabolites being minor fragments or adducts, which were not removed by the previous MZmine filtering process. Nevertheless, the overall data trend is confirmed.

In conclusion, after a detailed investigation on the strawberry extract by means of features extraction and tentative identification, two main results are shown: the longer the column, the higher the number of repeatable detected and identified features; the smaller the particles size, the lower the number of repeatable detected and identified features. The first result was not highlighted in the targeted analysis because of the low complexity of the standard mixture compared to the strawberry extract, implying the strong importance of a chromatographic evaluation on a real matrix before an untargeted profiling. Most importantly, however, was the second unexpected result. In this case, both targeted and untargeted analyses suggested that 2.6 µm sized core-shell particles columns performed better than the 1.7 µm sized core-shell particles columns. The Kinetex™ C_{18} with 1.7 µm sized core-shell particles should result in better resolution, higher peak capacity, and greater efficiency than the Kinetex™ C_{18} 2.6 µm sized core-shell particles packed columns. However, based on our evidences, when working on real systems, under UHPLC conditions, the analytes retention can significantly be altered. This could be explained by the effect of temperature gradients through the columns. Indeed, when columns packed with very fine particles are operated at high mobile phase velocities, the friction of the mobile phase percolating through the

TABLE 1 Tentatively identified compounds by means of two columns packed with 2.6 μm sized particles

t_R	[M+H]$^+$	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
0.85	483.0769	$C_{20}H_{18}O_{14}$	0.03	HHDP hexoside	465.0665; 303.0138		2a	[13]
1.39	303.0711	$C_{12}H_{14}O_9$	−0.33	Galloyl pentose		MS2: 127.0393; 99.0446; 71.0498; 81.0000	2a	[13]
1.77	433.1132	$C_{21}H_{21}O_{10}^+$	−0.61	Pelargonidin hexoside derivative	271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
1.92	867.2131	$C_{45}H_{38}O_{18}$	−0.02	Procyanidin trimer (cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.039; 107.0495	2a	[13]
1.99	433.1131	$C_{21}H_{21}O_{10}^+$	−0.56	Pelargonidin hexoside derivative	271.0601	MS2: 121.0287; 141.0700; 169.0648; 65.0392	2a	[13]
2.08	785.0830	$C_{34}H_{24}O_{22}$	0.10	Di-HHDP hexose	483.0764; 465.0655; 303.0141		2a	[13]
2.34	595.1657	$C_{27}H_{31}O_{15}$	1.06	Pelargonidin dihexoside		MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[30]
2.74	433.1130	$C_{21}H_{21}O_{10}^+$	−0.21	Pelargonidin hexoside derivative	271.0601	MS2: 121.0287; 141.0700; 169.0648; 65.0392	2a	[13]
2.76	851.2183	$C_{45}H_{38}O_{17}$	−0.18	Proanthocyanidin trimer (afz cat cat)		MS2: 107.0495; 123.0442; 135.0441; 126.9900; 68.9977; 139.039	2a	[13]
3.18	433.1130	$C_{21}H_{21}O_{10}^+$	−0.20	Pelargonidin hexoside derivative	271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
4.57	433.1130	$C_{21}H_{21}O_{10}^+$	−0.28	Pelargonidin hexoside derivative	271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]

(Continues)

TABLE 1 (Continued)

t_R	$[M+H]^+$	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
4.62	785.0831	$C_{34}H_{24}O_{22}$	0.00	Di-HHDP hexose	635.0881; 617.0775; 483.0772; 465.0667; 303.0136	Pseudo MS3: 201.0182; 145.02841; 83.0385; 173.02332; 117.03349; 229.01315	2a	[13]
5.27	447.0922	$C_{21}H_{19}O_{11}^+$	−0.22	Pelargonidin glucuronide	271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
5.70	579.1500	$C_{30}H_{26}O_{12}$	−0.60	Procyanidin B1 (cat cat)	271.0601	Pseudo MS3: 123.0442; 135.0441; 126.9900; 68.9977; 139.039; 107.0495	1	[13]
5.94	579.1500	$C_{30}H_{26}O_{12}$	−0.53	Procyanidin dimer (cat cat)	271.0601	Pseudo ms3: 123.0442; 135.0441; 126.9900; 68.9977; 139.039; 107.0495	2a	[13]
6.24	291.0863	$C_{15}H_{14}O_6$	−0.07	Catechin		MS2: 123.0443; 139.0391; 147.0441; 119.0495; 91.0547; 68.9977	1	[13]
6.28	741.2237	$C_{33}H_{41}O_{19}^+$	−0.14	Pelargonidin deoxyhexosyl dihexoside	595.1663; 271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
6.30	579.1489	$C_{30}H_{27}O_{12}^+$	2.31	Pelargonidin coumaroyl hexoside	433.1131	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2b	
6.59	1155.2775	$C_{60}H_{50}O_{25}$	−0.92	Procyanidin tetramer (cat cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.039; 107.0495	2a	[13]
7.15	867.2132	$C_{45}H_{38}O_{18}$	−0.15	Procyanidin trimer (cat cat cat)		MS2: 123.0442; 135.0441; 126.99; 68.9977; 139.039; 107.0495	2a	[13]
7.33	1155.2776	$C_{60}H_{50}O_{26}$	−1.06	Procyanidin tetramer (cat cat cat cat)		MS2: 123.0442; 135.0441; 126.99; 68.9977; 139.039; 107.0495	2a	[13]
7.62	433.1131	$C_{21}H_{21}O_{10}^+$	−0.46	Pelargonidin hexoside derivative	271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]

(Continues)

TABLE 1 (Continued)

t_R	$[M+H]^+$	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
8.00	433.1130	$C_{21}H_{21}O_{10}^+$	−0.23	Pelargonidin hexoside derivative	271.0601	MS2: 121.0287; 141.0700; 169.0648; 65.0392	2a	[13]
8.40	433.1129	$C_{21}H_{21}O_{10}^+$	0.00	Pelargonidin hexoside derivative		MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
8.53	867.2132	$C_{45}H_{38}O_{18}$	−0.17	Procyanidin trimer (cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.039; 107.0495	2a	[13]
8.58	635.0880	$C_{27}H_{22}O_{18}$	−0.11	Galloyl HHDP hexose	617.0770; 483.0768; 465.0671; 303.0138	Pseudo ms3: 201.0182; 145.0284; 83.0385; 173.0233; 117.0335; 229.0131	2a	[13]
8.72	465.1394	$C_{22}H_{24}O_{11}$	−0.57	Epicatechin acetyl pentoside	291.0863	MS2: 291.0863; 123.0443; 139.0391; 147.0441; 119.0495; 91.0547; 68.9977	2b	
8.75	741.2240	$C_{33}H_{41}O_{19}^+$	−0.52	Pelargonidin deoxyhexosyl dihexoside	433.113; 271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
9.12	433.1131	$C_{21}H_{21}O_{10}^+$	−0.38	Pelargonidin hexoside derivative	271.0601	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
9.30	1155.2771	$C_{60}H_{50}O_{24}$	−0.57	Procyanidin tetramer (cat cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.0390; 107.0495	2a	[13]
9.44	851.2182	$C_{45}H_{38}O_{17}$	−0.17	Proanthocyanidin trimer (afz cat cat)		MS2: 107.0495; 123.0442; 135.0441; 126.9900; 68.9977; 139.039;	2a	[13]
10.36	1155.2764	$C_{60}H_{50}O_{24}$	−1.20	Procyanidin tetramer (cat cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.0390; 107.0495	2a	[13]
10.58	851.2184	$C_{45}H_{38}O_{17}$	−0.41	Proanthocyanidin trimer (afz cat cat)		MS2: 107.0495; 123.0442; 135.0441; 126.9900; 68.9977; 139.0390	2a	[13]

(Continues)

TABLE 1 (Continued)

t_R	$[M+H]^+$	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
10.76	867.2147	$C_{45}H_{38}O_{18}$	−1.82	Procyanidin trimer (cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.0390; 107.0495	2a	[13]
10.94	291.0862	$C_{15}H_{14}O_6$	0.37	Epicatechin		MS2: 123.0443; 139.0391; 147.0441; 119.0495; 91.0547; 68.9977	1	[13]
11.21	851.2185	$C_{45}H_{38}O_{17}$	−0.42	Proanthocyanidin trimer (afz cat cat)		MS2: 107.0495; 123.0442; 135.0441; 126.9900; 68.9977; 139.039	2a	[13]
11.31	433.1131	$C_{21}H_{21}O_{10}^+$	−0.50	Pelargonidin hexoside	271.0597	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	1	[13]
11.90	451.1235	$C_{21}H_{22}O_{11}$	1.10	Phloretin glucuronide	275.0911	Pseudo MS3: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	2b	
11.95	403.1024	$C_{20}H_{19}O_9^+$	−0.16	Pelargonidin pentoside		MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
12.22	773.2135	$C_{33}H_{40}O_{21}$	−0.17	Quercetin deoxyhexosyl dihexoside		MS2: 303.0502; 153.0184; 71.0498; 85.0290; 137.0237; 229.0497; 201.0550	2a	[13]
12.81	835.2237	$C_{45}H_{38}O_{16}$	−0.62	Proanthocyanidin trimer (afz afz cat)		MS2: 107.0495; 123.0442; 135.0441; 126.9900; 68.9977; 139.0390	2a	[13]
13.41	303.0138	$C_{14}H_6O_8$	−1.05	Ellagic acid		Pseudo ms3: 201.0182; 145.02841; 83.0385; 173.02332; 117.03349; 229.0131	2a	[29]
14.03	1123.2872	$C_{60}H_{50}O_{22}$	−0.53	Proanthocyanidin tetramer (afz afz cat cat)		107.0495; 119.0494; 91.0547; 123.0444; 135.0443; 146.8300; 157.0650	2a	[13]
14.11	625.1400	$C_{27}H_{28}O_{17}$	0.74	Kaempferol hexosyl glucuronide	449.1078; 287.0548	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[32]

(Continues)

TABLE 1 (Continued)

t_R	$[M+H]^+$	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
14.36	611.1610	$C_{27}H_{30}O_{16}$	−0.67	Kaempferol dihexoside	287.0549	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
14.68	493.1342	$C_{23}H_{24}O_{12}$	0.76	Phloretin malonyl pentoside	275.0914	Pseudo MS3: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	2b	
14.82	1155.2777	$C_{60}H_{50}O_{27}$	−1.13	Procyanidin tetramer (cat cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.039; 107.0495	2a	[13]
15.47	465.0671	$C_{20}H_{16}O_{13}$	−1.65	Ellagic acid hexoside	303.0137	Pseudo ms3: 201.0182; 145.02841; 83.0385; 173.0233; 117.0335; 229.0131	2a	[25]
15.69	937.0942	$C_{41}H_{28}O_{26}$	−0.11	Galloyl-di-HHDP-hexose (casuarictin/potentillin)	919.0843; 617.0777; 483.0769; 465.0667; 303.0136	Pseudo MS3: 201.0182; 145.02841; 83.0385; 173.0233; 117.0335; 229.0131	2a	[13]
15.95	403.1024	$C_{20}H_{19}O_9^+$	−0.35	Pelargonidin pentoside	271.0599	MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[13]
17.05	785.0833	$C_{34}H_{24}O_{22}$	−0.25	Di-HHDP hexose	617.0780; 483.0774; 303.0137	Pseudo MS3: 201.0182; 145.0284; 83.0385; 173.0233; 117.0335; 229.0131	2a	[13]
17.22	1139.2800	$C_{60}H_{50}O_{23}$	0.85	Proanthocyanidin tetramer (afz cat cat cat)	867.2145		2a	[13]
17.37	303.0137	$C_{14}H_6O_8$	−0.57	Ellagic acid		Pseudo MS3: 201.0182; 145.0284; 83.0385; 173.0233; 117.0335; 229.0131	2a	[29]
17.38	489.1019	$C_{23}H_{21}O_{12}^+$	2.23	Pelargonidin malonyl pentoside	271.0600	MS2: 271.0602; 121.0286; 141.07; 115.0546; 65.0392; 169.0649; 197.0599	2b	

(Continues)

TABLE 1 (Continued)

t_R	[M+H]+	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
17.76	867.2137	$C_{45}H_{38}O_{18}$	−0.67	Procyanidin trimer (cat cat cat)		MS2: 123.0442; 135.0441; 126.9900; 68.9977; 139.0390; 107.0495	2a	[13]
17.83	935.0790	$C_{41}H_{26}O_{26}$	−0.50	Galloyl-HHDP-hexose-HHDP (castalagin, vescalagin)	303.0138	Pseudo MS3: 201.0182; 145.0284; 83.0385; 173.0233; 117.0335; 229.0131	2a	[13]
18.09	851.2181	$C_{45}H_{38}O_{17}$	0.05	Proanthocyanidin trimer (afz cat cat)		MS2: 107.0495; 123.0442; 135.0441; 126.9900; 68.9977; 139.0390	2a	[13]
18.49	611.1246	$C_{26}H_{26}O_{17}$	−0.61	Quercetin glucuronyl pentoside	479.0821; 303.0498	Pseudo MS3: 153.0184; 71.0498; 85.0290; 137.0237; 229.0497; 201.0550	2a	[13]
19.13	519.1137	$C_{24}H_{23}O_{13}$	−0.71	Pelargonidin malonyl hexoside		MS2: 271.0602; 121.0286; 141.0700; 115.0546; 65.0392; 169.0649; 197.0599	2a	[28]
19.75	303.0136	$C_{14}H_6O_8$	−0.48	Ellagic acid		Pseudo MS3: 201.0182; 145.0284; 83.0385; 173.0233; 117.0335; 229.0131	2a	[29]
20.85	449.0717	$C_{20}H_{16}O_{12}$	−0.73	Ellagic acid deoxyhexoside	303.0137	Pseudo MS3: 201.0182; 145.02841; 83.0385; 173.0233; 117.0335; 229.0131	2a	[25]
20.97	851.2185	$C_{45}H_{38}O_{17}$	−0.46	Proanthocyanidin trimer (afz cat cat)		MS2: 107.0495; 123.0442; 135.0441; 126.9900; 68.9977; 139.039	2a	[13]
21.14	479.0823	$C_{21}H_{18}O_{13}$	−0.72	Quercetin glucuronide	303.0499	MS2: 303.0499; 153.0184; 71.0498; 85.029; 137.0237; 229.0497; 201.0550	2a	[13]
21.26	611.1610	$C_{27}H_{30}O_{16}$	−0.67	Kaempferol dihexoside		MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
21.40	465.1030	$C_{21}H_{20}O_{12}$	−0.73	Quercetin hexoside	303.0499	MS2: 303.0499; 153.0184; 71.0498; 85.029; 137.0237; 229.0497; 201.0550	1	[13]

(Continues)

TABLE 1 (Continued)

t_R	$[M+H]^+$	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
21.80	937.0958	$C_{41}H_{28}O_{26}$	−1.85	Galloyl-di-HHDP-hexose (casuarictin/potentillin)	919.0841; 785.0833; 465.0675; 303.0138	Pseudo MS3: 201.0182; 145.02841; 83.0385; 173.02332; 117.03349; 229.01315	2a	[13]
21.96	581.1504	$C_{26}H_{28}O_{15}$	−0.72	Kaempferol pentosyl hexoside	449.1080; 287.0550	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
23.10	819.2286	$C_{45}H_{38}O_{15}$	−0.41	Proanthocyanidin trimer (afz afz afz)		MS2: 107.0494; 119.0494; 91.0547; 123.0439	2a	[13]
23.24	437.1444	$C_{21}H_{24}O_{10}$	−0.48	Phloretin hexoside		MS2: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	2a	[13]
24.06	449.1082	$C_{21}H_{20}O_{11}$	−0.81	Kaempferol hexoside	287.0551	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
24.10	551.1035	$C_{24}H_{22}O_{15}$	−0.72	Quercetin malonyl hexoside	303.0500	MS2: 303.0499; 153.0184; 71.0498; 85.029; 137.0237; 229.0497; 201.0550	2a	[29]
24.73	785.0834	$C_{34}H_{24}O_{22}$	−0.37	Di-HHDP hexose	486.077; 465.0662; 303.0137	Pseudo MS3: 201.0182; 145.02841; 83.0385; 173.02332; 117.03349; 229.01315	2a	[13]
25.01	595.1663	$C_{27}H_{30}O_{15}$	−0.07	Kaempferol deoxyhexosyl hexoside	433.1128	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[33]
25.04	463.0873	$C_{21}H_{18}O_{12}$	−0.39	Kaempferol glucuronide	287.0550	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
25.14	449.1081	$C_{21}H_{20}O_{11}$	−0.68	Kaempferol hexoside		MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
25.30	569.1867	$C_{26}H_{32}O_{14}$	0.44	Phloretin pentosyl hexoside	437.1443; 407.1341; 275.0914	MS2: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	2a	[31]
25.72	581.1871	$C_{27}H_{32}O_{14}$	−1.21	Naringenin deoxyhexosyl hexoside	273.0758	Pseudo MS3: 153.0185; 91.0548; 119.0495; 69.0341; 147.0443	2a	[13]

(Continues)

TABLE 1 (Continued)

t_R	[M+H]$^+$	Molecular formula	Δ ppm	Name	MS in-source fragments	MS2/Pseudo MS3 fragments[a]	Identification confidence level	References[b]
26.36	507.1138	$C_{23}H_{22}O_{13}$	−1.01	Quercetin acetyl hexoside	303.0500	MS2: 303.0499; 153.0184; 71.0498; 85.029; 137.0237; 229.0497; 201.0550	2a	[13]
26.73	595.1666	$C_{27}H_{30}O_{15}$	−0.56	Kaempferol deoxyhexosyl hexoside		MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[33]
27.31	437.1446	$C_{21}H_{24}O_{10}$	−0.87	Phloretin hexoside	275.0914	MS2: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	2a	[13]
28.25	535.1087	$C_{24}H_{22}O_{14}$	−0.87	Kaempferol malonyl hexoside	287.0549	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
28.70	535.1087	$C_{24}H_{22}O_{14}$	−0.94	Kaempferol malonyl hexoside	287.0551	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
29.10	435.1289	$C_{21}H_{22}O_{10}$	−1.03	Naringenin hexoside	273.0758	Pseudo MS3: 153.0185; 91.0548; 119.0495; 69.0341; 147.0443	2a	[13]
30.57	491.1186	$C_{23}H_{22}O_{12}$	−0.40	Kaempferol acetyl hexoside	287.0550	MS2: 287.0548; 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	2a	[13]
31.18	583.2032	$C_{27}H_{34}O_{14}$	−1.86	Phloretin deoxyhexosyl hexoside	275.0913	MS2: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	2a	[13]
31.89	303.0500	$C_{15}H_{10}O_7$	−0.32	Quercetin		MS2: 153.0184; 71.0498; 85.0290; 137.0237; 229.0497; 201.0550	1	[13]
35.60	567.1853	$C_{30}H_{30}O_{11}$	2.25	Phloretin p-coumaroyl deoxyhexoside		MS2: 107.0496; 58.9612; 95.0497; 84.9602; 79.0549	2b	
37.63	287.0550	C15H10O6	0.03	Kaempferol		MS2: 153.0185; 68.9978; 121.0288; 157.0652; 213.0546	1	[13]

[a]MS2 fragments are referred to fragments acquired by using HCD fragmentation from the molecular ion. Pseudo MS3 fragments are referred to fragments acquired by using HCD fragmentation on one of the "in source fragments" of the molecular ion.

[b]Identification confidence levels: level (1) confirmed structures where a reference standard is available; level (2a) evidence by spectrum matching from literature or library; and level (2b) diagnostic evidence where no other structure fits the experimental MS2 information.

HHDP, hexahydroxydiphenic acid.

column bed generates heat that dissipates radially and longitudinally through the entire chromatographic system [15,16]. The wall region of the column tends to be cooler than its center and the band velocity is not constant across the column resulting in band broadening. This phenomenon takes place in all columns, whatever the nature of the packing material, but the amplitude of the thermal gradient increases with decreasing the particle size. Band broadening could be accounted for the worse result observed for the Kinetex™ 1.7 µm sized particles packed columns in this study. However, a deeper investigation on the mechanism of retention of these chromatographic systems is needed to shed light on this issue.

3.3 | Identification

Compounds usually identified in strawberry extracts mainly belong to the chemical classes of flavonoids, dihydrochalcones, proanthocyanidins, and ellagitannins [25]. Flavonoids in strawberry are usually found as glycoconjugates, often acylated with aliphatic or aromatic acidic groups. The general trait of glycoconjugates mass spectra in full scan is the presence of signals corresponding both to the protonated glycosylated molecule $[M+H]^+$ and to the protonated aglycone formed during the "in-source" fragmentation processes before the ions reach the analyzer [19]. Based on this phenomenon, the in-source fragmentation was used to assess the glycosidic moiety of flavonoids, whereas the MS/MS spectra allowed to confirm the aglycone structure. The same analysis was used for dihydrochalcones.

Ellagitannins are esters of hexahydroxydiphenic acid and glucose (in some cases gallic acid). Typical ions found in MS/MS spectra of ellagitannins are due to the loss of galloyl units, hexahydroxydiphenic acid units and conjugates of these with glucose [22], gallic acid, and ellagic acid [26]. Some of these fragments were found as a result of the "in-source" fragmentation process, others in the MS/MS spectra.

Finally, the last class of compounds analyzed was composed by proanthocyanidins, polymers of flavanols such as catechin, afzelechin and gallocatechin, forming procyanidins, propelargonidins, and prodelphinidins, respectively. As reported by Lin et al., [27] the most important ms2 fragments of proanthocyanidins are formed by the break of the interflavan bond between the monomers, by the loss of the B-ring with C2–C3 part of the C-ring, and by the loss of the A-ring. Some of these fragments were found because of the "in-source" fragmentation process, others in the MS/MS spectra.

The best chromatographic condition, i.e. in series 2.6 µm particles sized packed columns (system c), provided the largest number of identified compounds including anthocyanins, dihydrochalcones, dihydroflavonols, dihydroflavanols, flavanones, flavan-3-ols, proanthocyanidins, and ellagitannins. The identification data of tentatively identified compounds for the 2.6 µm particles sized packed columns are summarized in Table 1 with the relative level of identification confidence. Based on the work of Schymanski et al. [24], level 2a refers to compounds tentatively identified by checking MS/MS spectra data in literature or in on-line spectral libraries, and level 2b refers to compounds tentatively identified by studying diagnostic fragments in MS/MS spectra but not supported by the literature data. Coherently, the compounds tentatively identified with identification levels 1 or 2a have been already found in the literature [13,25,28–33], whereas the compounds with identification level 2b have never been reported in the literature, to the best of our knowledge. The final confirmation of individual metabolites should be completed by comparison of retention times and fragmentation patterns with authentic standard compounds. For detailed information on the identification and on the compounds identified by means of the other chromatographic systems see the Supporting Information.

4 | CONCLUDING REMARKS

In this work, a systematic evaluation of the effect of column length and particle size on the analysis of complex phytochemical mixtures is proposed, first by a targeted analysis on a standard mixture and then by an untargeted analysis on a strawberry extract. The results showed that although column length did not significantly affect the targeted analysis of the standard mixture, still it did significantly affect the performance of a chromatographic system when dealing with a real complex phytochemical mixture, highlighting the importance of a chromatographic optimization before the untargeted profiling. More importantly, despite the generally acknowledged better performance of sub-2 µm chromatographic columns, the Kinetex™ core-shell 1.7 µm sized particles packed columns did not perform as well as the Kinetex™ core-shell 2.6 µm sized particles packed columns, allowing to detect and tentatively identify a lower number of compounds. A possible reason for this unexpected result could be the effect of high pressures and frictional heating on sub-2 µm sized particles packed columns. However, further deeper investigations are needed to justify the obtained results. The inconsistency among the theoretical assumptions and the applicative findings clearly shows the importance of an extensive chromatographic system investigation before proceeding with the untargeted profiling analysis of a complex matrix.

REFERENCES

1. Kohler, I., Giera, M., Recent advances in liquid-phase separations for clinical metabolomics. *J. Sep. Sci.* 2017, *40*, 93–108.
2. Zhang, H., Jiang, Y., Wu, J., Zheng, C., Ran, X., Li, D., Huang, M. B. H., Metabolic mapping of *Schisandra sphenanthera* extract and its active lignans using a metabolomic approach based on ultra

high performance liquid chromatography with high-resolution mass spectrometry. *J. Sep. Sci.* 2017, *40*, 574–586.

3. Yang, S., Synovec, R. E., Kalyuzhnaya, M. G., Lidstrom, M. E., Development of a solid phase extraction protocol coupled with liquid chromatography mass spectrometry to analyze central carbon metabolites in lake sediment microcosms. *J. Sep. Sci.* 2011, *34*, 3597–3605.
4. Kaufmann, A., Combining UHPLC and high-resolution ms: a viable approach for the analysis of complex samples? *TrAC Trends Anal. Chem.* 2014, *63*, 113–128.
5. Kim, S., Kim, J., Yun, E. J., Kim, K. H., Food metabolomics: from farm to human. *Curr. Opin. Biotechnol.* 2016, *37*, 16–23.
6. Werner, E., Heilier, J.-F., Ducruix, C., Ezan, E., Junot, C., Tabet, J.-C., Mass spectrometry for the identification of the discriminating signals from metabolomics: current status and future trends. *J. Chromatogr. B* 2008, *871*, 143–163.
7. Antignac, J.-P., Courant, F., Pinel, G., Bichon, E., Monteau, F., Elliott, C., Le Bizec, B., Mass spectrometry-based metabolomics applied to the chemical safety of food. *TrAC Trends Anal. Chem.* 2011, *30*, 292–301.
8. Cacciola, F., Farnetti, S., Dugo, P., Marriott, P. J., Mondello, L., Comprehensive two-dimensional liquid chromatography for polyphenol analysis in foodstuffs. *J. Sep. Sci.* 2017, *40*, 7–24.
9. Li, Q., Liu, Y., Han, L., Liu, J., Liu, W., Feng, F., Zhang, J., Xie, N., Chemical constituents and quality control of two Dracocephalum species based on high-performance liquid chromatographic fingerprints coupled with tandem mass spectrometry and chemometrics. *J. Sep. Sci.* 2016, *39*, 4071–4085.
10. Wolfender, J.-L., Marti, G., Thomas, A., Bertrand, S., Current approaches and challenges for the metabolite profiling of complex natural extracts. *J. Chromatogr. A* 2015, *1382*, 136–164.
11. Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M., Graham Cooks, R., The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* 2005, *40*, 430–443.
12. Stobiecki, M., Kachlicki, P., Wojakowska, A., Marczak, Ł., Application of LC/MS systems to structural characterization of flavonoid glycoconjugates. *Phytochem. Lett.* 2015, *11*, 358–367.
13. La Barbera, G., Capriotti, A. L., Cavaliere, C., Piovesana, S., Samperi, R., Zenezini Chiozzi, R., Laganà, A., Comprehensive polyphenol profiling of a strawberry extract (Fragaria × ananassa) by ultra-high-performance liquid chromatography coupled with high-resolution mass spectrometry. *Anal. Bioanal. Chem.* 2017, *409*, 2127–2142.
14. Available from: https://www.phenomenex.com/Info/Page/16knxconcertwb.
15. Gritti, F., Guiochon, G., Complete temperature profiles in ultra-high-pressure liquid chromatography columns. *Anal. Chem.* 2008, *80*, 5009–5020.
16. Kaczmarski, K., Gritti, F., Kostka, J., Guiochon, G., Modeling of thermal processes in high pressure liquid chromatography: II. Thermal heterogeneity at very high pressures. *J. Chromatogr. A* 2009, *1216*, 6575–6586.
17. Kajdžanoska, M., Petreska, J., Stefova, M., Comparison of different extraction solvent mixtures for characterization of phenolic compounds in strawberries. *J. Agric. Food Chem.* 2011, *59*, 5272–5278.
18. Pluskal, T., Castillo, S., Villar-Briones, A., Orešič, M., MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 2010, *11*, 395–405.
19. Nagy, Á., Abrankó, L., Profiling of hydroxycinnamoylquinic acids in plant extracts using in-source CID fragmentation. *J. Mass Spectrom.* 2016, *51*, 1130–1145.
20. Zhang, T., Creek, D. J., Barrett, M. P., Blackburn, G., Watson, D. G., Evaluation of coupling reversed phase, aqueous normal phase, and hydrophilic interaction liquid chromatography with Orbitrap mass spectrometry for metabolomic studies of human urine. *Anal. Chem.* 2012, *84*, 1994–2001.
21. Willemse, C. M., Stander, M. A., de Villiers, A., Hydrophilic interaction chromatographic analysis of anthocyanins. *J. Chromatogr. A* 2013, *1319*, 127–140.
22. de Villiers, A., Cabooter, D., Lynen, F., Desmet, G., Sandra, P., High performance liquid chromatography analysis of wine anthocyanins revisited: Effect of particle size and temperature. *J. Chromatogr. A* 2009, *1216*, 3270–3279.
23. Comandini, P., CZE separation of strawberry anthocyanins with acidic buffer and comparison with HPLC. *J. Sep. Sci.* 2008, *31*, 3257–3264.
24. Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., Hollender, J., Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environ. Sci. Technol.* 2014, *48*, 2097–2098.
25. Sun, J., Liu, X., Yang, T., Slovin, J., Chen, P., Profiling polyphenols of two diploid strawberry (Fragaria vesca) inbred lines using UHPLC-HRMSn. *Food Chem.* 2014, *146*, 289–298.
26. Hooi Poay, T., Sui Kiong, L., Cheng Hock, C., Characterisation of galloylated cyanogenic glucosides and hydrolysable tannins from leaves of *Phyllagathis rotundifolia* by LC-ESI-MS/MS: Characterisation of galloylated cyanogenic glucosides and tannins. *Phytochem. Anal.* 2011, *22*, 516–525.
27. Lin, L.-Z., Sun, J., Chen, P., Monagas, M. J., Harnly, J. M., UHPLC-PDA-ESI/HRMS^n profiling method to identify and quantify oligomeric proanthocyanidins in plant products. *J. Agric. Food Chem.* 2014, *62*, 9387–9400.
28. Kårlund, A., Hanhineva, K., Lehtonen, M., Karjalainen, R. O., Sandell, M., Nontargeted metabolite profiles and sensory properties of strawberry cultivars grown both organically and conventionally. *J. Agric. Food Chem.* 2015, *63*, 1010–1019.
29. Aaby, K., Mazur, S., Nes, A., Skrede, G., Phenolic compounds in strawberry (Fragaria x ananassa Duch.) fruits: composition in 27 cultivars and changes during ripening. *Food Chem.* 2012, *132*, 86–97.
30. D'Urso, G., d'Aquino, L., Pizza, C., Montoro, P., Integrated mass spectrometric and multivariate data analysis approaches for the discrimination of organic and conventional strawberry (Fragaria ananassa Duch.) crops. *Food Res.* 2015, Int. 77, Part 2, 264–272.
31. Álvarez-Fernández, M. A., Hornedo-Ortega, R., Cerezo, A. B., Troncoso, A. M., García-Parrilla, M. C., Determination of nonanthocyanin phenolic compounds using high-resolution mass spectrometry (UHPLC-Orbitrap-MS/MS) and impact of storage

conditions in a beverage made from strawberry by fermentation. *J. Agric. Food Chem.* 2016, *64*, 1367–1376.

32. Hanhineva, K., Rogachev, I., Kokko, H., Mintz-Oron, S., Venger, I., Kärenlampi, S., Aharoni, A., Non-targeted analysis of spatial metabolite composition in strawberry (Fragaria×ananassa) flowers. *Phytochemistry* 2008, *69*, 2463–2481.

33. Carazzone, C., Mascherpa, D., Gazzani, G., Papetti, A., Identification of phenolic constituents in red chicory salads (Cichorium intybus) by high-performance liquid chromatography with diode array detection and electrospray ionisation tandem mass spectrometry. *Food Chem.* 2013, *138*, 1062–1071.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Chiozzi RZ, Capriotti AL, Cavaliere C, et al. Evaluation of column length and particle size effect on the untargeted profiling of a phytochemical mixture by using UHPLC coupled to high-resolution mass spectrometry. *J Sep Sci.* 2017;40: 2541–2557. https://doi.org/10.1002/jssc.201700135

4 Paper V:

Chromatographic column evaluation for the untargeted profiling o f glucosinolates in cauliflower by means of ultra-high performance liquid chromatography coupled to high resolution mass spectrometry

Talanta xxx (2017) xxx-xxx

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com

Chromatographic column evaluation for the untargeted profiling of glucosinolates in cauliflower by means of ultra-high performance liquid chromatography coupled to high resolution mass spectrometry

Anna Laura Capriotti, Chiara Cavaliere, Giorgia La Barbera*, Carmela Maria Montone, Susy Piovesana, Riccardo Zenezini Chiozzi, Aldo Laganà

Department of Chemistry, University of Rome "La Sapienza", Piazzale Aldo Moro 5, 00185 Rome, Italy

ARTICLE INFO

Keywords:
High resolution mass spectrometry
Columns evaluation
Glucosinolates
Untargeted profiling
Cauliflower

ABSTRACT

The untargeted profiling is a promising approach for the characterization of secondary metabolites in biological matrices. Thanks to the recent rapid development of high-resolution mass spectrometry (HRMS) instrumentations, the number of applications by untargeted approaches for biological samples profiling has widely increased in the recent years. Despite the high potentialities of HRMS, however, a major issue in natural products analysis often arises in the upstream process of compounds separation. A separation technique is necessary to avoid phenomena such as signal suppression, and it is especially needed in the presence of isomeric metabolites, which are otherwise indistinguishable. Glucosinolates (GLSs), a group of secondary metabolites widely distributed among plants, resulted to be associated to the prevention of some serious diseases, such as cancer. This led to the development of several methods for the analysis of GLSs in vegetables tissues. The issue of GLSs chromatographic separation has been widely studied in the past because of the difficulty in the analysis of this highly polar and variable class of compounds. Several alternatives to reversed phase (RP) chromatography, sometimes not compatible with the coupling of liquid chromatography with mass spectrometry, have been tested for the analysis of intact GLSs. However, the availability of new stationary phases, in the last years, could allow the re-evaluation of RP chromatography for the analysis of intact GLSs.

In this work, a thorough evaluation of four RP chromatographic columns for the analysis of GLSs in cauliflower (*Brassica oleracea L.* var. *botrytis*) extracts by an ultra-high performance liquid chromatographic system coupled via electrospray source to a hybrid quadrupole-Orbitrap mass spectrometer is presented. The columns tested were the following: one column Luna Omega polar C_{18}, one column Kinetex Biphenyl, one column Kinetex core-shell XB-C_{18}, two columns Kinetex core-shell XB-C_{18}. After a previous optimization of the extraction method, cauliflower extracts were analyzed testing four different mobile phases onto the four columns for a total of sixteen different chromatographic conditions. The chromatographic systems were evaluated based on the number of detected and tentatively identified GLSs. Luna Polar stationary phase resulted to be the most suitable for the analysis of GLSs compared to Kinetex XB and Kinetex Biphenyl columns stationary phase. However, two in series Kinetex XB columns increased the number of tentatively identified GLSs compared to one Kinetex XB, showing the importance of column length in the analysis of complex mixtures. The data obtained with the best chromatographic system were deeply analyzed by MS/MS investigation for the final identification. Fifty-one GLSs were tentatively identified, 24 of which have never been identified in cauliflower. Finally the linearity of the analytes response over the analyzed range of concentration was checked, suggesting that the developed method is suitable for both qualitative and quantitative analysis of GLSs in phytochemical mixtures.

Abbreviations: AGC, Automatic gain control; CID, Collision-Induced Dissociation; ESI, Electrospray Ionization; FWHM, Full Width at Half Maximum; GSL, Glucosinolate; HCD, Higher-Energy Collisional Dissociation; HILIC, Hydrophilic Interaction Liquid Chromatography; HRMS, High Resolution- Mass Spectrometry; RP, Reversed Phase; RT, Room temperature; IRMPD, Infrared Multiphoton Dissociation; TFA, Trifluoroacetic Acid.

* Corresponding author.
Email addresses: annalaura.capriotti@uniroma1.it (A.L. Capriotti); chiara.cavaliere@uniroma1.it (C. Cavaliere); giorgia.labarbera@uniroma1.it (G. La Barbera); carmelamaria.montone@uniroma1.it (C.M. Montone); susy.piovesana@uniroma1.it (S. Piovesana); riccardo.zenezini@uniroma1.it (R.Z. Chiozzi); aldo.lagana@uniroma1.it (A. Laganà)

https://doi.org/10.1016/j.talanta.2017.12.019
Received 10 November 2017; Received in revised form 5 December 2017; Accepted 6 December 2017
Available online xxx
0039-9140/

1. Introduction

Over the past few years, metabolomics has gained an important role in food analysis and food quality assessment [1–3] due to the close relationship between secondary metabolites and biological activity [3]. Among the several methods that have been used for the characterization of plant secondary metabolites, the untargeted profiling approach is thought to be the one allowing the widest metabolic coverage [4]. Indeed, in contrast to targeted methods, where the identification is usually limited to metabolites for which authentic reference standards are available, the untargeted approach allows to detect, and possibly identify, thousands of analytical features. The recent rapid development of high-resolution mass spectrometers coupled to ultra-high performance liquid chromatography (UHPLC) systems had an important role in the increase of the untargeted profiling applications on complex biological samples [5,6]. High-resolution mass spectrometry (HRMS), indeed, allows to distinguish compounds by their accurate *m/z* signals and to study their fragmentation pattern in tandem HRMS mode for the final identification. However, despite the high potentialities of HRMS, a major issue in natural products analysis arises in the upstream process of compounds separation. An efficient separation technique is necessary to provide compounds individual retention times, which are later used to support their identification, and it is especially needed in the presence of isomeric metabolites which are otherwise indistinguishable [7]. In complex phytochemical mixtures there is a wide variety of positional isomers that generally elute together, making them indistinguishable. The coelution of isomeric metabolites, together with the coelution of others high abundant compounds present in the matrix, could generate ionization suppression of the less abundant compounds. Therefore, optimizing the chromatographic conditions could be crucial for maximizing the number of detected and consequently identified compounds [6,8].

Glucosinolates (GLSs) are a group of secondary metabolites widely distributed in the plants of the order of Capparales, which includes the family of *Brassicaceae* (syn. *Cruciferae*). Several epidemiologic studies showed an association between GLSs intake and the prevention of some serious diseases such as cancer [9–12]. The antitumor properties of cruciferous vegetables is mostly due to the breakdown products of GLSs [13]. During storage, food processing and cooking procedures, GLSs are decomposed into isothiocyanates, thiocyanates, indole derivatives, and nitriles in the presence of myrosinase enzyme [14–16]. The evidence of an association between antitumor properties and GLSs is engendering growing interest of researchers in their investigation and characterization in vegetables tissues. GLSs exhibit a common chemical structure of a β-thioglucose moiety, a sulfated oxime group, and a variable –R aliphatic, aromatic or indolic group depending upon whether they originate from the metabolism of aliphatic amino acids, aromatic amino acids, or tryptophan, respectively [9,17,18]. The presence of the sulfate group, together with the high variability of the –R side chain, represent an issue in the GLSs chromatographic separation. In the past years, the most common used approach was the conversion of the intact GLSs into desulfo-derivatives, in order to decrease the polarity of GLSs allowing an increased retention and a better chromatographic resolution. However, the desulfation step makes the sample preparation more time-consuming [19], therefore, studying the development of new methods for the analysis of intact GLSs is still an issue of interest. Several alternatives to reversed phase (RP) chromatography such as hydrophilic interaction chromatography (HILIC) [20,21] or ion pairing chromatography [22] have been tested for the analysis of intact GLSs. However, these methods often require some additives that can lead to strong ion suppression when coupling chromatography with mass spectrometry (MS). The availability of new stationary phases in the market could allow the re-evaluation of RP chromatography for the analysis of intact GLSs. Few recent studies have been published on RP chromatographic column evaluation for the analysis of GLSs [23–26]. The coupling of an optimized chromatographic system with HRMS, that is considered the best technology for the structural elucidation of unknown GLSs [23,27], could lead to the identification of a larger number of compounds and, in turn, to a widest metabolic coverage.

For this purpose, in this work a thorough evaluation of four RP chromatographic columns for the analysis of GLSs in cauliflower (*Brassica oleracea L.* var. *botrytis*) extracts was accomplished by an UHPLC system coupled via electrospray (ESI) source to a hybrid quadrupole-Orbitrap mass spectrometer. After a previous optimization of the extraction method, cauliflower extracts were analyzed testing four different mobile phases onto four columns for a total of sixteen different chromatographic conditions. The chromatographic systems were evaluated based on the number of detected and putatively identified GLSs. The data obtained with the best chromatographic system were deeply analyzed for the final identification and for checking the linear response of GLSs in cauliflower extracts.

2. Materials and methods

2.1. Chemicals and reagents

LC-MS grade water was purchased from Fisher Scientific. LC-MS grade methanol and acetonitrile were purchased by VWR International (Milan, IT). Formic acid, ammonium formate and acetone were purchased from Sigma Aldrich (St. Louis, MO, US).

2.2. Sample preparation

Cauliflowers were bought in a local farm, washed with distilled water and cut into quarters. The samples were freeze-dried, then ground using liquid nitrogen. GLSs were extracted following the protocol by Sun et al. [23] with some modifications. The protocol was operated at two different temperatures, namely room temperature (RT) and 70 °C, to evaluate the effect of heating on the extraction. Briefly, 3 g-sample was extracted in 30 mL of MeOH/H_2O 70:30 (*v/v*). In one case the extract was stirred for 10 min and sonicated for 10 min at RT, in the second case the extract was stirred for 10 min and sonicated for 10 min at 70 °C. In both cases, the sample was centrifuged for 10 min at 6000×*g* at 4 °C. The supernatant was collected. The extraction was repeated once with 12 mL of MeOH/H_2O 70:30 (*v/v*). The supernatants were mixed and dried under a gentle nitrogen flow in a water bath at 37 °C. The residue was reconstituted with 3 mL of H_2O/ACN 90:10 (*v/v*). The sample was filtered through 13 mm Acrodisc syringe filter with 0.2 μm GHP membrane (Pall Co., Ann Arbor, MI, US). The extract was divided in aliquots and stored at −20 °C for further analysis. Three experimental replicates for each extraction protocol were done.

2.3. UHPLC-ESI–MS/MS analysis

The UHPLC system was a Vanquish binary pump H (Thermo Fisher Scientific, Bremen, Germany), equipped with a thermostatted autosampler and a thermostatted column compartment. The separation was carried out onto four different chromatographic systems: (1) one column Luna Omega polar C_{18} (100 Å pore size, Phenomenex, Torrance, CA, US) 100 mm × 2.1 mm i.d. with 1.6 μm particles size; (2) one column Kinetex core-shell Biphenyl (100 Å pore size, Phenomenex, Torrance, CA, US) 100 mm × 2.1 mm i.d. with 1.7 μm particles size; (3) one column Kinetex core-shell XB-C_{18} (100 Å pore size, Phenomenex, Torrance, CA, US) 100 mm × 2.1 mm i.d) with 2.6 μm particles size; (4) two columns Kinetex core-shell XB-C_{18} (100 Å pore size, Phenomenex,

Torrance, CA, US) 100 mm × 2.1 mm i.d. with 2.6 μm particles size connected by means of a virtually zero-dead-volume UHPLC fingertight fitting (Thermo Scientific). All columns were operated at a flow rate of 400 μL min^{-1} except for the two columns in series that were operated at 600 μL min^{-1}.

Several different phases were explored for all the columns: a) H_2O:HCOOH 99.9:0.1 (*v/v*) as phase A and ACN:HCOOH 99.9:0.1 (*v/v*) as phase B; b) H_2O:HCOOH 99.8:0.2 (*v/v*) as phase A and ACN:HCOOH 99.8:0.2 (*v/v*) as phase B; c) H_2O 5 mmol L^{-1} $HCOONH_4$ as phase A and ACN 5 mM $HCOONH_4$ as phase B; d) H_2O 10 mmol L^{-1} $HCOONH_4$ as phase A and ACN as phase B. The several chromatographic systems will be referred in the text addressing to the number and the letter corresponding to the used column and phases, respectively. For a summary of the tested conditions see Table S1 in the Supplementary information.

The chromatographic gradient for systems (1–2–3) was the following: 5% B for 2 min, 5–15% B in 10 min, 15–95% B in 3 min, a washing step at 95% B for 5 min and a re-equilibration step at 5% B for 13 min. The chromatographic gradient for system (4) was the following: 5% B for 4 min, 5–25% B in 40 min, 25–95% B in 10 min, a washing step at 95% B for 5 min and a re-equilibration step at 5% B for 15 min. In all gradients, the columns were maintained at 40 °C using the still air option. The injection volume was 5 μL.

The chromatographic system previously described was coupled using a heated ESI source to a hybrid quadrupole-Orbitrap mass spectrometer QExactive (Thermo Fisher Scientific). ESI source parameters in negative mode were as follows: capillary temperature 275 °C; sheath gas 50 (arbitrary units); auxiliary gas 15 (arbitrary units); sweep gas 3 (arbitrary units); spray voltage 3500 V; auxiliary gas heater temperature 450 °C; S-Lens RF level 50 (%).

For all the investigated chromatographic conditions, the detection was conducted in HRMS data dependent acquisition mode, setting the appropriate parameters according to La Barbera et al. [6]. MS data were acquired in the range *m/z* 150–1000 in negative ionization mode with a resolution (full width at half maximum, FWHM, @*m/z* 200) of 70000. Automatic gain control (AGC) target value was 200,000 in full scan, with a max ion injection time set at 100 ms. Isolation window width was 2 *m/z*. MS/MS fragmentation was performed on the 5 most intense ions detected in full scan with a resolution (FWHM, @*m/z* 200) of 17500. AGC target value was 10,000. Dynamic exclusion was set to 2 s. An exclusion list containing the most abundant ions commonly detected in the blank was set. Fragmentation was achieved in the higher-energy collisional-dissociation cell at 40% normalized collision energy.

The comparison of the extracts obtained at RT and at 70 °C was carried out, prior to the columns comparison, with the chromatographic system (1a). Once the RT extraction protocol was chosen as the best one, the chromatographic evaluation was carried out. The source was cleaned at the beginning of the day before running the sample batch. The samples were acquired in continuous during the day and within one week to assure reproducible instrumental performance. On each day, one phase was used for the four columns conditioning and chromatographic runs. Before the injection of the extracts on each column, a pre-conditioning step of 1 h was carried out. The cauliflower extracts were run in triplicate followed by the blank injection (H_2O:ACN, 90:10 (*v/v*)) for each chromatographic condition.

2.4. Data analysis

2.4.1. Features extraction

Accurate mass ion chromatograms obtained from three consecutive injections of cauliflower extract and blank for each condition were processed by the software MZmine v2.26, an open source software (http://mzmine.sourceforge.net/) usually employed in the pre-processing of metabolomics data [28]. The pre-processing workflow was optimized for each column. Briefly, the accurate mass LC–MS data were imported, and a list of ions for each scan was generated using a mass detection module (Exact mass algorithm with noise 1E4). A filter was applied to correct the false signals around *m/z* peaks, called "shoulder peaks" that are residues of the Fourier Transform function (Gaussian peak model function). Then, mass lists generated and filtered for each MS scan were used to build a chromatogram for each mass that could be detected continuously over the scans by means of the chromatogram builder tool (minimum time span 0.05 min; minimum height 5E4; *m/z* tolerance 0.0015 or 5 ppm). Later, the deconvolution module deconvoluted these chromatograms into individual peaks (local minimum search algorithm with chromatographic threshold 30; minimum retention time range 0.02 min for Luna, 0.04 min for one Kinetex XB, 0.03 min Kinetex Biphenyl and two Kinetex XB columns); minimum relative height 0%; minimum absolute height 5E4; minimum top/edge 1.3 for Luna, 1.1 for Kinetex Biphenyl, one Kinetex XB and two Kinetex XB columns; peak duration 2.5 min). With the isotopic peak grouper algorithms, the isotopic peaks were removed from the peak list keeping only the highest peak (*m/z* tolerance 0.0015 or 5 ppm; retention time tolerance 0; maximum charge 2; most intense representative isotope). The peak lists obtained for the replicates and the blanks were aligned with the Join alignment module (*m/z* tolerance 0.0015 or 5 ppm; retention time tolerance 0.1 min; weight for *m/z* 90; weight for retention time 10). Peaks coming from the blank were subtracted. Only peaks common to all three replicates were maintained (repeatable features).

2.4.2. Identification

The first identification step consisted in the application of the Fragment search module (Retention time tolerance 0.0 min; *m/z* tolerance of MS2 data 0.001 or 5 ppm; max fragment peak height 100%; min MS2 peak height 5E3). This module identifies fragment peaks present in the MS spectra due to the in-source fragmentation process. In-source fragments are identified by two conditions: 1) same retention time for the precursor and the in-source fragment ions; 2) presence in the precursor ion tandem mass spectrum of a peak with the same *m/z* as the in-source fragment ion. The second identification step consisted in the application of the Adduct search module ($[M\text{-}H\text{-}H_2O]^-$, $[M+Na\text{-}2H]$, $[M+Cl]^-$, $[M+FA\text{-}H]^-$, retention time tolerance 0.0 min; *m/z* tolerance 0.0015 or 5 ppm; max adduct peak height 100%). This module identifies adducts if the retention time of the original ion and the adduct ion is the same. In-source fragment ions together with the detected adducts were removed from the list of the detected features because they do not represent a compound but they were used as additional information to identify their precursor. The third identification step consisted in searching the obtained features against a custom database (*m/z* tolerance 0.0015 or 5 ppm, retention time not set) created by looking through the typical reported compounds in vegetable matrices and combining all the possible conjugated compounds which might be expected [27]. The compounds that matched with the in-house database were further investigated by the inspection of the MS/MS spectra and classified according to four levels of identification confidence. Based on the work of Schymanski et al. [29], level 1 is associated to compounds identified by comparison with an authentic reference standard, level 2 is associated to compounds tentatively identified by typical MS/MS fragments, level 3 is associated to compounds that show some informative fragments, not sufficient to tentatively identify one exact structure only, level 4 is associated to compounds to which a molecular formula can be unambiguously assigned. Based on these indications, we classified as level 2 compounds that showed R1 side chain specific fragments; as level 3 compounds that showed a fragmentation pattern in common with all GLSs; as level 4 compounds associated to a specific molecular formula by means of accurate mass. For more details on the identification see subsection 3.1 in Section 3.

2.4.3. Relative quantification for the comparison of the extraction methods

The chromatograms resulting from the analysis of the samples obtained with the two extraction protocols were processed with MZmine, setting the parameters previously described for Luna column. The areas of the compounds identified at level 3 were compared between the two groups of samples by the MZmine tools logratio and CV (coefficient of variation). If $Area_{Rt}$ and $Area_{70\ °C}$ are the average of the areas obtained for each metabolite over the 3 experimental replicates of the RT extraction and the 70 °C extraction, respectively, the log ratio is defined as:

$$logratio = \frac{\ln\ (Area_{70°C}/Area_{Rt})}{\ln\ (2)}$$

Moreover, CV, defined as the ratio of the standard deviation to the mean, has been calculated for each metabolite. In Fig. 1, logratio and CV are shown for each metabolite.

2.5. Dynamic linear range

To assess the response linearity of GLSs in cauliflower extracts, the sample was diluted 5, 10, 100, 1000, 10000 times with H_2O/ACN:90/10 (*v/v*) and analyzed with the best chromatographic system (4a). GLSs identified at level 3 have been considered for the evaluation. The correlation of GLSs levels with LC-MS peak areas was examined by calculating the Pearson correlation coefficient (R^2) by means of Excel 2007.

3. Results and discussion

In this work, an evaluation of several chromatographic conditions for the analysis of GLSs in a complex matrix is proposed by means of an untargeted approach on a cauliflower extract.

At first, a rapid comparison between two different extraction protocols was carried out using the column Luna under the chromatographic condition (1a) prior to the columns evaluation. It is known that GLSs are hydrolyzed by myrosinase when heated at high temperatures [15,30,31]. Therefore, several precautions should be taken when analyzing GLSs, such as storing them at −20 °C until analysis and choosing an extraction protocol that could reduce myrosinase activity. The myrosinase is denatured in organic solvent and heating, so most of the extractions are performed in ethanol/water (50/50) or methanol/water (70/30) at 70 °C [24,32,33]. However, several other extractions performed at RT have been published [23,34]. To clarify this point, the same extraction protocol was tested on our cauliflower samples at 70 °C and RT, respectively. The comparison of the two extraction methods was accomplished by an untargeted approach, looking for the compounds identified at confidence level 3 as explained in the Material and Method section. Briefly, the compounds that matched against the database (level 4) were further investigated checking the most intense fragments in the MS/MS spectra. It is reported that GLSs fragmentation pattern, beside the fragments depending on the –R variable group, presents some ions common to all the GLSs compounds [35]. When the compounds of interest showed in the MS/MS spectra all three fragments $[HSO_4]^-$, $[C_2H_3OS]^-$, $[SO_3]^{\cdot-}$, corresponding to *m/z* 96.9596, 74.9904 and 79.9568, respectively, they were classified as level 3 identified GLSs and selected for the comparison of the two extraction methods. The log function of ratio of metabolites area in the extraction at 70 °C over metabolites area over the extraction at 25 °C is shown in Fig. 1a. The CV plot, highlighting the points having more significance in the log graph, is shown in Fig. 1b. Between two variables the one with the smaller CV is less dispersed than the variable with the larger CV. Although several metabolites are not significantly different in the two extraction protocols (red dots in Fig. 1b), it is evident that most of the significantly different metabolites (green dots in Fig. 1b) are more abundant in the extraction carried out at 25 °C than in the extraction carried out at 70 °C (green dots in Fig. 1a).

This result, in contrast with most papers in which a comparison of several extraction methods has been carried out, is in agreement with the work by Doheny-Adams *et al.* [36], showing the better performance of the extraction at RT in 80% MeOH compared to the boiling 80% MeOH. Based on these results, cauliflower samples were extracted at RT and used for the following chromatographic evaluation.

Chromatographic separation of intact GLSs was historically a technical challenge that required several techniques such as ion pairing HPLC [22] or HILIC chromatography [20,21], because of the poor retention of GLSs with hydrophilic side chains. With the discovery that moderately acidified eluents allow easy separation of intact GLSs on RP columns [37], their separation is thought to be no longer a technical challenge. Intuitively, it might be expected that acids with lower pKa than that of the GLSs sulfate group (pKa ~ 2) would be required for sufficient retention, but it has been shown that both trifluoroacetic acid (TFA) (pKa 0.52) [37] and formic acid (pKa 3.74) [33] induce sufficient retention of GLSs to RP HPLC columns. Ammonium formate has also been used giving efficient ion pairing [38]. However, when cou-

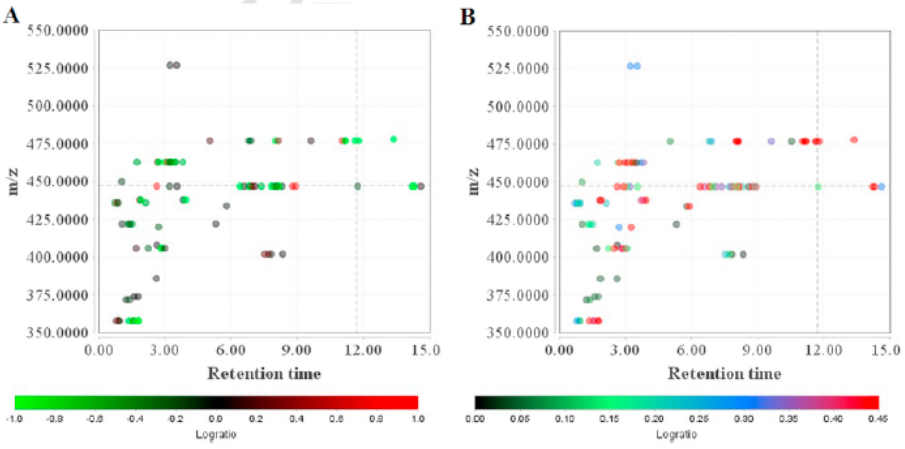


Fig. 1. a) Plot of log ratio of metabolites area obtained from the room temperature extraction and the 70 °C extraction and relative CV plot b).

pling HPLC with MS, besides a good chromatographic separation, effects such as ion suppression should be taken into account when using these compounds as additive in the mobile phase. While it is known that TFA can give ion suppression during the ionization process, formic acid and ammonium formate, within certain concentration limits, are compatible with MS. Therefore, we wanted to test formic acid and ammonium formate at different concentrations in the chromatographic phases to obtain the best compromise between an efficient separation and a good ionization. Four different columns were evaluated with four different phases for a total of 16 different chromatographic conditions (Table S1). The chromatographic and mass spectrometric method was kept unchanged for columns (1) (2) and (3) and was adapted to system (4) by doubling the gradient time. The temperature was set to 40 °C and the flow-rate was set to 400 μL min^{-1} for columns (1) (2) and (3) and 600 μL min^{-1} for column (4) to achieve the best efficiency, as suggested by the specification sheet.

By a quick visual inspection of the chromatograms it was seen that the difference in terms of retention time and peak shape was larger when comparing different columns than when comparing different phases on the same column. A large chromatographic retention time shift and a big difference in peaks shape can severely affect the peak picking and alignment steps in data analysis, sometimes giving distorted results. Based on these considerations, we decided to analyze the obtained chromatograms comparing the four phases for each column in a single data analysis workflow and, once chosen the best phase for each column, we proceeded with a visual inspection of the peaks to compare the four columns. The comparison of the different phases on each column was accomplished by the number of detected features that matched against a home-made database (i.e., level 4 identification). Although many false-positive features can be caused both by poor chromatographic signals and by non-molecular ions (i.e., dimers $[2M - H]^-$, salt cluster ions and in-source fragments), in this study, adducts and in-source fragments were identified using MZmine (see Material and Method section). Such non-molecular ion features were removed and the remained features were searched against a home-made database. Looking at the table shown in Fig. 2a, it is possible to see that, in all cases, the phases with a percentage of formic acid give better results compared to the phases with ammonium formate. The obtained result, repeated for all the columns, suggests that the improved number of features is mostly due to an improvement in the ionization process rather than a real gain in chromatographic separation.

When comparing different concentration of formic acid, in all cases it was shown an increase of detected GLSs when using 0.1% compared to 0.2% of formic acid. Apparently, whereas the peak shape was slightly better with 0.2% formic acid, the intensity of the peaks was sometimes decreased, losing the detection of the less abundant compounds. Sensitivity in LC-MS is dependent mainly on the single compound ionization efficiency, and the magnitude of ionization in the ESI source varies greatly with the chemical properties of the mobile phase composition. Based on these results, we decided that 0.1% formic acid in the mobile phases was the good compromise between separation improvement and effective ionization.

Once chosen the best phase for the four columns, we proceeded with the evaluation of the chromatographic performances among the columns. Four different RP columns were evaluated. Indeed, although several studies have been published in the past on RP columns for the analysis of GLSs, few recent studies have been published [23–26]. Therefore, the availability of new stationary phases in the market could provide better performances compared to the past. For this purpose, we wanted to explore different particle morphologies, column lengths, stationary phases that, based on the specification sheet of column vendors, could be suitable for GLSs analysis. The Kinetex XB-C_{18} has some protective butyl side chains that allow better peak shape and enhance separation of basic compounds compared to normal C18 phases. However, also the retention of acidic compound results to be increased [39]. Based on these characteristics, this stationary phase was considered to be suitable for the analysis of highly polar and acidic sulphated compounds, such as GLSs. For the same reasons, Luna Omega Polar C18 has been selected for the analysis of GLSs. Indeed, while the C18 chains provide hydrophobic interactions, a polar modified particle surface provides enhanced retention of the highly polar compounds. Luna Polar C18, allowing a more balanced retention of polar and hydrophobic compounds, was considered to be suitable for a class of compounds such as GLSs, showing both high polarity and high heterogeneity in terms of hydrophobicity due to the variability of the –R side chain and the possible acylation of the thioglucose. The Kinetex Biphenyl instead, is suitable for aromatic compounds separation. Therefore, an ameliorant on the separation of aromatic and the indolic GLSs, representing one of the most abundant classes of GLSs in cauliflower, was expected. Finally, aside from the different surface chemistry, the columns under investigation also present a different particle morphology. Indeed, whereas the Luna column is packed with fully porous silica particles, Kinetex XM-C_{18} and Kinetex Biphenyl are core-shell particles packed

	Luna	Biphenyl	1 Kinetex	2 Kinetex
HCOOH 0.1%	173	242	335	306
HCOOH 0.2%	155	239	290	280
$HCOONH_4$ 5mM	62	82	95	84
$HCOONH_4$ 10mM	50	69	57	57

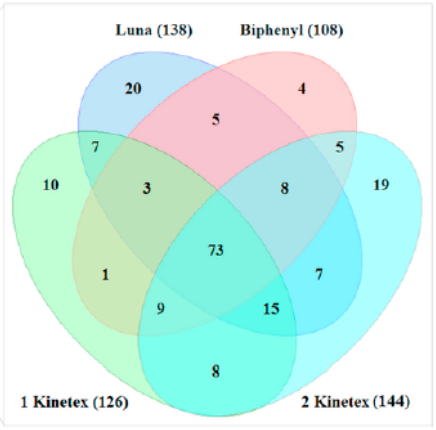


Fig. 2. a) Number of repeatable tentatively identified features at confidence level 4 for the 16 chromatographic conditions; b) Venn diagram of the tentatively identified features at confidence level 3 for the four columns under the best chromatographic conditions showed in Fig. 2a).

columns. The novel core-shell RP columns show minor band broadening compared to fully porous particles, leading to an important efficiency improvement. Moreover, the much lower back pressure of core-shell materials allows to evaluate longer columns systems. Therefore, also the coupling of two core-shell columns in series was evaluated to observe if an improvement in the analysis of GLSs in a rich and complex extract, such as the cauliflower extract, could be observed.

Whereas the four phases were compared throughout each column by the compounds identified at level 4, the four columns, in their best condition (a), were compared after the visual inspection of the MS/MS spectra, thus based on the number of compounds identified at level 3.

As shown in the Venn diagram, the Kinetex Biphenyl column gave the worst results among the 4 columns. The Kinetex Biphenyl column, that was expected to show good results at least for indolic and phenyl GLSs, resulted to show the worst peak shape and worst isomers separation compared to the other columns, for all the classes of GLSs. As shown in the Fig. 2b, the number of tentatively identified GLSs was higher for the two Kinetex XB in series compared to the other systems. By a visual inspection of the chromatograms it could also be observed that the Luna Polar column and the two Kinetex XB showed the best results in term of peak shape and repeatability (Fig. S1 and Fig. S2, Supplementary information). The similar peak shape together with the slightly difference in the number of tentatively identified GLSs obtained with the Luna column and the two Kinetex XB columns pose a question about the choice between these two chromatographic systems for the analysis of GLSs. The obtained better result for the 2 Kinetex XB columns is likely due to the higher retention of two coupled columns in series and to a minor suppression of less abundant GLSs by coeluting compounds present in the matrix rather than a more favourable chemical interaction of the analytes and the Kinetex XB stationary phase. Indeed, when comparing one column Kinetex XB with one column Luna Polar, the number of the identified compounds is higher and the shape of the peaks is better for the latter. In Table S2, in the Supplementary information, is shown the list of the GLSs located in each intersection of the Venn diagram (Fig. 2), together with a classification of GLSs detected by each column in order to find a correlation between column chemistry and compound class. In general, some classes of GLSs resulted to be better separated by Kinetex XB column such as saturated aliphatic GLSs, methylthio substituted GLSs and indolic GLSs, whereas other classes such as unsaturated aliphatic GLSs, phenyl substituted GLSs, hydroxylated, sulfonyl and sulfinyl substituted GLSs resulted to be better separated by Luna column. Finally, non acylated GLSs are better retained by Luna column, likely because of their higher polarity compared to acylated GLSs. In general, it is clear that Luna column performs better than Kinetex XB column in separation of more polar compounds such as sulfonyl, sulfinyl, hydroxylated GLSs. This is probably due to the polar modifier bonded on the surface of Luna Polar together with the C18 chains that allow to retain a wide variety of polar compounds. This would suggest that the coupling of two Luna columns could give much better performances than the coupling of two Kinetex XB columns. In general, coupling two columns could be advantageous for a complete and comprehensive profiling of a matrix [8]. However, when dealing with a large number of samples, and, when the differences between column performances are so slight, a system implying the use of a single column should be preferred, in order to reduce the time and costs of the analysis. However, in order to present a comprehensive profiling of GLSs in only one variety of cauliflower, in this work the two columns Kinetex XB in series were chosen as the best system for the final identification and linearity investigation. The extracted ion chromatograms of some of the tentatively identified GLSs are reported in Fig. 3 for Luna Polar column and two in series Kinetex XB.

3.1. Identification

GLSs consist of a common glycone group and a variable aglycone side chain (R1). The glycone is characterized by a sulfonated oxime group and a thioglucose that can be conjugated to some acyl groups (R2), namely phenolic acids such as cinnamic acid, coumaric acid, ferulic acid and sinapic acid. Due to the presence of a sulfonate moiety, GLSs naturally occur in anionic form as $[M\text{-}H]^-$, thus being identified in negative mode. Typical MS fragmentation patterns of GLSs have been widely investigated [27,33–35], showing the presence of both product ions common to all GLSs and highly specific fragments due to the variability of the side chain (R1) and of the acyl group (R2). The major common fragmentation pathway [35] includes the following ions: $[HSO_4]^-$ at *m/z* 97, $[C_2H_3O_5S]^-$ at *m/z* 139, $[HO_4S_2]^-$ at *m/z* 129, $[C_2H_2O_4NS]^-$ at *m/z* 136, $[SO3]^{-\cdot}$ at *m/z* 80, $[C_2H_3OS]^-$ at *m/z* 75 and $[C_2H_4O_5NS]^-$ at *m/z* 154. Although the most intense signal is at *m/z* 97, the bisulfate anion $[HSO_4]^-$ is not specific for GLSs because most organic sulphates form this anion [35]. Moreover, also the formation of an *m/z* 80 $[SO_3]^{-\cdot}$ product ion from *m/z* 97 by loss of a hydroxyl group is not specific. Therefore, the detection of these two ions was not considered enough to classify the detected features as tentatively identified GLSs. The three most intense ions (*m/z* 97, 75, 80), typical of the GLSs fragmentation pattern, were used instead for a rapid screening of features in the step of columns comparison. The features that showed in

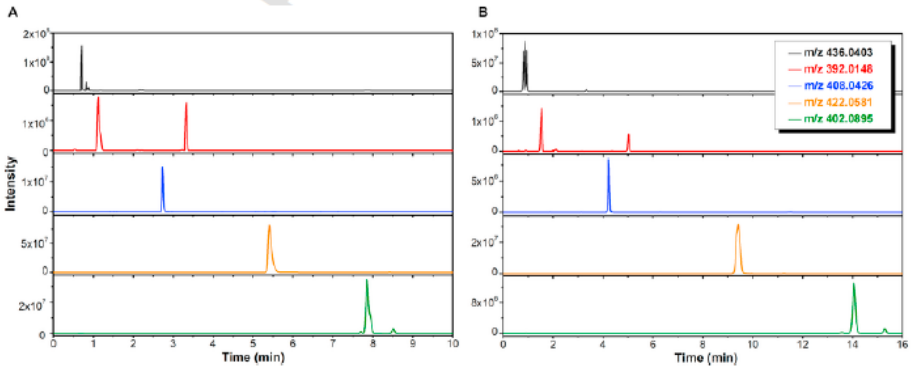


Fig. 3. Extracted ion chromatograms of methylsulfinylbutyl GLS (Glucoraphanin, *m/z* 436.0403), methylthioethyl GLS (*m/z* 392.0148), benzyl GLS (Glucotropaeolin, *m/z* 408.0426), phenylethyl GLS (*m/z* 422.0581) and hexyl GLS (*m/z* 402.0895), acquired with Luna Polar Omega (a) and two in series Kinetex XB columns (b) in the best chromatographic conditions, i.e phases with HCOOH 0.1%.

their MS/MS spectra the typical common GLSs fragmentation pattern were classified as tentatively identified with confidence level 3.

Once chosen the best system, consisting of two Kinetex XB-C_{18} in series, a deeper investigation on GLSs structure was carried out: fragments resulting specific of the R1 or R2 side chains of GLSs were searched in the MS/MS spectra.

If the detected fragments were diagnostic of non-acylated or acylated GLSs, the compounds were assigned to level 3a and 3b, respectively. Indeed, non acylated compounds show some typical fragments due to the thioglucose moiety: the ion at *m/z* 195 corresponds to the D-thioglucose group $[C_6H_{11}O_5S]^-$; the ion *m/z* 259 derives from the loss of R1-N=C=S; the loss of a water molecule from the ion 259 gives the ion at *m/z* 241; the loss of R1-N=C=O corresponds to the ion at *m/z* 275. The last three ions are probably formed through an intramolecular rearrangements in which the sulfate group is transferred to the thioglucose moiety. When these fragments were found in the MS/MS spectra, GLSs were considered belonging to the class of non-acylated GLSs and thus classified as level 3a. On the contrary, in the acylated GLSs spectra, fragments due to the thioglucose moiety exhibit *m/z* value increase due to the acyl group mass. For instance, in sinapoyl GLSs a typical fragmentation pattern could show the ions at 481 = 275 + 206, 465 = 259 + 206, 447 = 241 + 206, and 401 = 195 + 206. Moreover, ions at *m/z* 193, 147, 163, 223, and 207, corresponding to isoferuloylate ($[C_{10}H_9O_4]^-$), cinnamoylate ($[C_9H_7O_2]^-$), coumaroylate ($[C_9H_7O_3]^-$), sinapoylate ($[C_{11}H_{11}O_5]^-$), and dimethoxy-cinnamoylate ($[C_{11}H_{11}O_4]^-$) can also be detected. Therefore, if either fragments typical of the acylthioglucose moiety or fragments due to the acyl groups were found in the MS/MS spectra, GLSs were considered belonging to the class of acylated GLSs and thus classified as level 3b.

Finally, some specific fragments due to the variable aglycone side chain R1 can be detected and can be used to assign an unequivocal structure to the specific GLS. Indeed, the loss of the R1 group by cleavage of the bond on either sides of the sulphur atom and with charge retention on the sugar aglycone can be observed. Due to this phenomenon, in both acylated and non acylated GLSs constant neutral losses of the *m/z* values $[M\text{-}196\text{-}H]^-$, $[M\text{-}178\text{-}H]^-$, and $[M\text{-}162\text{-}H]^-$, $[M\text{-}162\text{-}SO_3\text{-}H]^-$, can be detected. Therefore, either acylated or non acylated GLSs that showed the typical fragments due to their side chain R1 were classified as level 2. Of course, some of the accurate masses associated to the R1 side chains, correspond to different isomeric structures. Therefore, a comparison with reference standard would be requested for their unequivocal identification.

In Table 1 the compounds identified at level 2, 3a and 3b have been reported together with the detected diagnostic fragments. Whereas the diagnostic fragments allowed the identification of the accurate mass of the R side chain of some GLSs, it was not possible to distinguish among positional isomers. Therefore, in case in which several isomers have been detected, the general name has been associated to the detected isomers. Instead, whether one only isomer was detected it was associated to the GLSs isomer usually found in Brassicacee. For example, in the case of the ion at *m/z* 434.0616, tentatively identified as methylthiopentyl GLS, only the peak at retention time 10.18 has been detected and then associated to the most common isomer, namely, glucoberteroin. Whereas in the case of the ions at *m/z* 388.0373, because of the detection of three isomers, none of them could be certainly associated to the most commonly found one, namely (epi)progoitrin, but they were referred to as hydroxyl-butenyl GLSs.

Several GLSs showed different isomers because of the different position of hydroxyl groups (i.e hydroxy-butenyl GLSs, hydroxyglucobrassicin), because of the different arrangement of the aliphatic chain (i.e. pentyl GLS, isobutyl GLS) or because of the position of methoxyl groups. For instance, the two isomers methoxyglucobrassicin and neoglucobrassicin have been associated to the compounds at retention time 11.6 and 17.4 respectively, based on literature data where it is reported that neoglucobrassicin elutes later than methoxyglucobrassicin [32,40]. Moreover, in the MS spectrum of neoglucobrassicin, the in source fragment at *m/z* 447.0532 corresponding to glucobrassicin was detected with high intensity (Fig. S4, Supplementary information). The weakest bond between the methoxyl group and the nitrogen in neoglucobrassicin could be responsible of the MS behaviour and be a further confirmation of the assignment of these identities to the two compounds. A similar behaviour was found for N-sulfoglucobrassicin. Indeed, a very low intensity peak with accurate mass corresponding to N-sulfoglucobrassicin was found in the MS spectra, whereas a very abundant peak due to the in source fragment at *m/z* 447 corresponding to glucobrassicin was detected (Fig. S4, Supplementary information). It is noteworthy that glucobrassicin was not found in the MS spectra of cauliflower extracts, in contrast with several previous works on cauliflower GLSs profiling [15,40–45]. Nevertheless, only one of these works confirmed the identity of glucobrassicin with an authentic reference standard [43]. On the contrary, on the best of our knowledge, it is the first time that N-sulfoglucobrassicin has been detected in cauliflower extracts. It is likely that, due to the ease of sulfate loss, less sensitive techniques allowed only the detection of the more abundant in source fragment ion instead of the less abundant precursor ion, leading to misinterpretation of the data and to the identification of glucobrassicin over N-sulfoglucobrassicin. However, further identity confirmation should be accomplished by means of standard compounds.

N-sulfoglucobrassicin was not the only compound to be detected for the first time in cauliflower extracts. Indeed, the development of the used method allowed the tentative identification (confidence level 2, 3a, 3b) of 51 GLSs, 24 of which have never been identified in cauliflower. However, among the identified compounds just few acylated compounds have been identified at level 2. Indeed, just for few of the detected compounds the typical fragmentation behaviour of acylated GLSs was found. This could be due to the fact that previous works reporting the typical fragmentation pattern of acylated GLSs used different fragmentation techniques. Moreover, few acylated GLSs have been reported in the literature so far [24]. Compounds already found in cauliflower are highlighted by the reported reference in Table 1, whereas for the compounds never found in cauliflower any reference has been reported.

3.2. Linear response for GLSs in cauliflower extracts

Observing the obtained chromatograms, some of the most intense peaks resulted to be wide, broadening or, sometimes, splitting. This chromatographic behaviour suggested a possible saturation effect. Indeed, if the functional C18 groups on column particles are saturated, the remaining molecules interact with the not conjugated Si-OH, causing peak broadening or splitting. As well as a limited columns loading capacity can be observed, a limited amount of compounds can be effectively ionized by the ESI source for compounds present as major components. Indeed, in many cases, a saturation effect, owing to the fact that the total number of ions per time unit formed during ESI is about constant, could also be observed. In this case, a sample dilution, decreasing compound concentrations, decreases the competition for ion evaporation from the droplet surface in the ESI step, reducing remarkably the matrix effect on co-eluting analytes response [46]. Therefore, for a suitable quantitative analysis, it is necessary to avoid, as much as possible, any chromatographic as well as any ionization saturation phenomenon. To check the linearity response of GLSs in cauliflower extracts in the range of concentration used in this work, the sample was diluted 5, 10, 100, 1000, 10000 times and analyzed with the best chromatographic system 4a. The correlation of GLSs levels with LC-MS peak areas was examined by calculating the Pearson correlation coefficient (R^2) for GLSs identified at level 3. Whereas the original sample allowed the detection of 51 compounds, the following dilutions led to

Table 1
Compounds identified (confidence identification level 2, 3a ,3b) with two in series Kinetex XB columns in the best chromatographic conditions, i.e phases with HCOOH 0.1%.

t_R (min)	Identity	[M-H]$^-$	Molecular formula	Δ (ppm)	MS/MS fragments	Identification confidence level	Reference
0.80	methylsulfinylpropyl GLS	422.0249	$C_{11}H_{21}O_{10}NS_3$	−1.35	180.0148 (R1CNHOS); 274.9897; 259.0123; 241.0023; 195.0323; 96.9588; 74.9897; 79.9560; 138.9696; 135.9699; 128.9309	2	[40–42,47]
0.81	methylsulfinylbutyl GLS	436.0403	$C_{12}H_{23}O_{10}NS_3$	−1.88	194.0307 (R1CNHOS); 259.0123; 274.9989; 195.0325; 96.9588; 74.9897; 79.9560; 138.9697	2	[40–42,47]
0.88	methylsulfinylbutyl GLS	436.0405	$C_{12}H_{23}O_{10}NS_3$	−1.42	194.0308 (R1CNHOS); 259.0124; 274.9990; 195.0327; 96.9588; 74.9897; 79.9560; 138.9697	2	
0.93	methylsulfinylbutyl GLS	436.0403	$C_{12}H_{23}O_{10}NS_3$	−1.88	194.0306 (R1CNHOS); 259.0125; 274.9988; 195.0330; 96.9588; 74.9897; 79.9560; 138.9697	2	
1.01	Glucoconringiin (hydroxy-methylpropyl GLS)	390.0531	$C_{11}H_{21}O_{10}NS_2$	−0.76	148.0425 (R1CNHOS); 274.9895; 259.0131; 195.0319; 194.0124 ($R1CNO_4S$-H); 96.9588; 74.9897; 79.9560; 138.9691; 135.9698	2	
1.02	Sinigrin (allyl GLS)	358.0265	$C_{10}H_{17}O_9NS_2$	−1.91	116.0163 (R1CNHOS); 274.9888; 259.0136; 241.0026; 195.0323; 161.9858 ($R1CNO_4S$-H); 96.9588; 74.9896; 138.9694; 79.9560	2	
1.32	methylsulfinylpropyl GLS	422.0257	$C_{11}H_{21}O_{10}NS_3$	0.54	180.0151 (R1CNHOS); 274.9897; 259.0132; 195.0318; 96.9588; 74.9897; 79.9560; 138.9697; 135.9699	2	
1.34	Glucoputranjivin (methylethyl GLS)	360.0428	$C_{10}H_{19}O_9NS_2$	−0.09	118.0317 (R1CNHOS); 195.0309; 96.9588; 74.9896; 139.0025; 79.9560; 129.0178; 79.9560	2	
1.37	hydroxybenzyl GLS	424.0378	$C_{14}H_{19}O_{10}NS_2$	0.12	259.0121; 96.9588; 74.9897; 138.9691	3a	[43]
1.47	methylsulfinylpropyl GLS	422.0255	$C_{11}H_{21}O_{10}NS_3$	0.07	180.0151 (R1CNHOS); 274.9897; 259.0132; 195.0318; 96.9588; 74.9897; 79.9560; 138.9696; 135.9697	2	
1.49	methylthioethyl GLS	392.0148	$C_{10}H_{19}O_9NS_3$	−0.27	150.0041 (R1CNHOS); 274.9899; 259.0134; 195.0316; 96.9587; 74.9897; 79.9559; 138.9693	2	[40–42]
1.75	Glucomatronalin (dihydroxybenzyl GLS)	440.0297	$C_{14}H_{19}O_{11}NS_2$	−6.73	274.9902; 241.0007; 96.9588; 74.9897; 79.9560; 138.9697	3a	[23,40–42,47,48]
1.75	Gluconapin (Butenyl GLS)	372.0421	$C_{11}H_{19}O_9NS_2$	−1.97	130.0319 (R1CNHOS); 196.0293 ($R1CNHO_5S$); 259.0136; 274.9888; 241.0031; 195.0320; 96.9588; 74.9896; 79.9560; 138.9694; 128.9309	2	[23,40–42,47,48]
1.80	Hydroxyglucobrassicin (hydroxy-indolylmethyl GLS)	463.0485	$C_{16}H_{20}O_{10}N_2S_2$	−0.32	274.9886; 96.9588; 74.9897; 79.9560; 138.9701	3a	[23,40–42,47,48]
1.81	coumaroyl glucocapparin (methyl GLS)	478.0478	$C_{17}H_{21}O_{11}NS_2$	−1.07	96.9588; 163.0388; 74.9898	3b	[23,40–42,47,48]
1.85	methylsulfinylpropyl GLS	422.0252	$C_{11}H_{21}O_{10}NS_3$	−0.64	180.0148 (R1CNHOS); 274.9905; 259.0119; 195.0316; 96.9588; 74.9897; 79.9560; 138.9698; 128.9308	2	[40–42]

Table 1 (*Continued*)

t_R (min)	Identity	[M-H]$^-$	Molecular formula	Δ (ppm)	MS/MS fragments	Identification confidence level	Reference
1.85	sinapoyl hydroxyglucobrassicin (hydroxy-indolylmethyl GLS)	669.1081	$C_{27}H_{30}O_{14}N_2S_2$	2.31	96.9588; 223.0604; 74.9899	3b	[40–42]
1.95	Hydroxyglucobrassicin (hydroxy-indolylmethyl GLS)	463.0482	$C_{16}H_{20}O_{10}N_2S_2$	−0.97	259.0146; 96.9588; 74.9897; 79.9560	3a	
2.17	pentyl GLS	374.0578	$C_{11}H_{21}O_9NS_2$	−1.83	132.0476 (R1CNHOS); 274.9894; 259.0125; 195.0326; 178.0163 (R1CNO$_4$S-H); 96.9587; 74.9896; 79.9559	2	[41,42,47]
2.19	methylthiopropyl GLS	406.0296	$C_{11}H_{21}O_9NS_3$	−2.35	164.0198 (R1CNHOS); 274.9898; 259.0129; 241.0018; 195.0322; 209.9891 (R1CNO4S-H); 96.9588; 74.9897; 79.9560; 138.9693; 128.9309	2	[40,42,47]
2.27	pentyl GLS	374.0578	$C_{11}H_{21}O_9NS_2$	−1.83	132.0477 (R1CNHOS); 274.9893; 259.0123; 241.0020; 195.0321; 96.9588; 74.9896; 79.9560; 138.9695	2	[40,42,47]
2.27	Hydroxyglucobrassicin (hydroxy-indolylmethyl GLS)	463.0485	$C_{16}H_{20}O_{10}N_2S_2$	−0.32	221.0382 (R1CNHOS); 267.0087 (R1CNO$_4$S-H); 274.9908; 259.0124; 241.0020; 195.0326; 96.9588; 74.9897; 79.9560; 128.9316; 138.9701; 135.9699	2	
2.34	coumaroyl benzyloxyethyl GLS	612.0834	$C_{25}H_{27}O_{13}NS_2$	−2.77	288.0171 (R1CNHO$_5$S); 74.9896; 79.9560; 96.9587	2	[23,40,41,47,48]
2.54	glucosibarin (hydroxyphenylethyl GLS)	438.0534	$C_{15}H_{21}O_{10}NS_2$	0.00	241.0016; 96.9588; 74.9897; 79.9560; 138.9697	3a	[23,40,41,47,48]
2.63	pentyl GLS	374.058	$C_{11}H_{21}O_9NS_2$	−1.29	132.0476 (R1CNHOS); 274.9906; 259.0139; 195.0319; 96.9588; 74.9896; 79.9560; 138.9696	2	[23,40,41,47,48]
2.65	Methylthiopropyl GLS	406.0297	$C_{11}H_{21}O_9NS_3$	−2.11	164.0201 (R1CNHOS); 274.9899; 259.0131; 241.0017; 195.0320; 96.9588; 74.9897; 79.9560; 138.9696	2	[23,40,41,47,48]
2.66	cinnamoyl eptyl GLS	546.1496	$C_{25}H_{33}O_{10}NS_2$	4.21	96.9589; 147.0439; 74.9897	3b	
2.70	Hydroxyglucobrassicin (hydroxy-indolylmethyl GLS)	463.0482	$C_{16}H_{20}O_{10}N_2S_2$	−0.97	221.0382 (R1CNHOS); 267.0087 (R1CNO$_4$S-H); 274.9886; 259.0132; 241.0013; 195.0330; 96.9588; 74.9897; 79.9560; 138.9701	2	
2.84	pentyl GLS	374.0581	$C_{11}H_{21}O_9NS_2$	−1.02	132.0476 (R1CNHOS); 274.9906; 259.0139; 195.0319; 96.9588; 74.9896; 79.9560; 138.9699	2	
3.33	Glucoraphanin (methylsulfinylbutyl GLS)	436.0408	$C_{12}H_{23}O_{10}NS_3$	−0.73	194.0307 (R1CNHOS); 259.0124; 275.0001; 195.0332; 96.9588; 74.9897; 79.9560; 138.9697	2	[23,40–42,44,48]
4.12	Hydroxyglucobrassicin (hydroxy-indolylmethyl GLS)	463.0485	$C_{16}H_{20}O_{10}N_2S_2$	−0.32	221.0379 (R1CNHOS); 274.9886; 259.0127; 96.9588; 74.9897; 79.9560; 138.9701	2	[23,40–42,44,48]
4.22	Glucotropaeolin (benzyl GLS)	408.0426	$C_{14}H_{19}O_9NS_2$	−0.57	166.0327 (R1CNHOS); 230.0123 (R1CNHO$_5$S); 274.9903; 259.0125; 241.0027; 195.0324; 212.0009 (R1CNO4S-H); 96.9588; 74.9897; 79.9560; 138.9696; 128.9310	2	[23,40–42,44,48]

Table 1 (Continued)

t_R (min)	Identity	[M-H]$^-$	Molecular formula	Δ (ppm)	MS/MS fragments	Identification confidence level	Reference
4.24	Glucobrassicanapin (pentenyl GLS)	386.0586	$C_{12}H_{21}O_9NS_2$	0.30	144.0484 (R1CNHOS); 259.0142; 96.9587; 74.9896; 79.9560; 138.9701	2	[23,40–42,44,48]
4.46	Glucoerucin (methylthiobutyl GLS)	420.0457	$C_{12}H_{23}O_9NS_3$	−1.20	178.0355 (R1CNHOS); 274.9899; 259.0123; 241.0017; 195.0325; 224.006 (R1CNO$_4$S-H); 96.9588; 74.9897; 79.9560; 138.9695	2	[23,40–42,44,48]
4.49	hydroxybenzyl GLS	424.0358	$C_{14}H_{19}O_{10}NS_2$	−4.59	182.0264 (R1CNHOS); 259.0123; 195.0312; 96.9588; 74.9897	2	[23,40,41,48]
5.09	N-sulfoglucobrassicin (N-sulfoindolylmethyl GLS)[a]	527.0096	$C_{16}H_{20}O_{12}N_2S_3$	−1.80	447.0532 in source; 205.0433 (R1CNHOS); 251.0132 (R1CNO$_4$S-H); 259.0127; 274.9903; 241.0020; 195.0327; 96.9588; 74.9897; 79.9560; 138.9694	2	[41]
7.88	coumaroyl benzyloxynonyl GLS	710.1939	$C_{32}H_{41}O_{13}NS_2$	−1.05	96.9587; 163.0389; 74.9897	3b	[41]
8.30	Hydroxyglucobrassicin (hydroxy-indolylmethyl GLS)	463.0485	$C_{16}H_{20}O_{10}N_2S_2$	−0.32	274.9886; 259.0127; 195.0330; 96.9588; 74.9897; 79.9560; 128.9316; 138.9701; 135.9699	3a	
9.40	phenylethyl GLS	422.0583	$C_{15}H_{21}O_9NS_2$	−0.43	180.0483 (R1CNHOS); 274.9898; 259.0123; 241.0019; 195.0325; 96.9588; 74.9897; 79.9560; 138.9697	2	[40–42,44,48]
9.55	sinapoyl glucoerucin (methylthiobutyl GLS)	626.1055	$C_{23}H_{33}O_{13}NS_3$	2.21	74.99; 96.9588; 223.0604	3b	
10.18	Glucoberteroin (methylthiopentyl GLS)	434.0616	$C_{13}H_{25}O_9NS_3$	−0.59	259.0124; 96.9588; 74.9897; 79.9560	3a	
11.58	Methoxyglucobrassicin (methoxy-indolylmethyl GLS)	477.0635	$C_{17}H_{22}O_{10}N_2S_2$	−1.67	235.0541 (R1CNHOS); 274.9898; 259.0127; 241.0015; 195.0325; 96.9588; 74.9897; 79.9560; 128.9308; 138.9694; 135.9699	2	
14.03	hexyl GLS	402.0895	$C_{13}H_{25}O_9NS_2$	−0.70	160.0789 (R1CNHOS); 274.9905; 259.0125; 241.0021; 195.0324; 96.9588; 74.9897; 79.9560; 138.9694; 128.9311	2	
15.25	hexyl GLS	402.0895	$C_{13}H_{25}O_9NS_2$	−0.70	160.0785 (R1CNHOS); 259.0126; 96.9587; 74.9896; 79.9560; 138.9694; 128.9312	2	
17.36	Neoglucobrassicin (N-methoxyindolylmethyl GLS)[a]	477.0636	$C_{17}H_{22}O_{10}N_2S_2$	−1.46	447.0532 in source; 205.0433 (R1CNHOS); 274.9905; 259.0127; 241.0022; 96.9588; 74.9897; 79.9560; 128.9309; 138.9694;	2	
19.28	cinnamoyl octyl GLS	560.1615	$C_{24}H_{35}O_{10}NS_2$	−2.58	389.0674 (R2C$_6$H$_{10}$O$_9$S); 96.9589; 74.9898	3b	
23.31	eptyl GLS	416.1056	$C_{14}H_{27}O_9NS_2$	0.40	274.9909; 259.0125; 195.0330; 96.9588; 74.9897; 79.9560	3a	
34.79	octyl GLS	430.1207	$C_{15}H_{29}O_9NS_2$	−0.89	188.126 (R1CNHOS); 96.9587; 74.9897; 79.9560	2	
35.60	sinapoyl glucocheirolin (methylsufonylpropyl GLS)	644.0771	$C_{22}H_{31}O_{13}NS_3$	−1.86	96.9587; 223.0606; 74.9899	3b	
47.07	cinnamoyl indolyl GLS	563.0796	$C_{24}H_{24}O_{10}N_2S_2$	−0.62	96.9587; 128.9318; 74.9897	3b	
48.08	cinnamoyl indolyl GLS	563.0798	$C_{24}H_{24}O_{10}N_2S_2$	−0.26	96.9587; 128.9318; 74.9897	3b	[23,41,42]

[a] The reported fragments have been observed in the MS/MS spectra of the ion produced by means of in source fragmentation.

the decrease of the number of detected GLSs (51, 51, 40, 13, 12, 5 for 5, 10, 100, 1000, 10000 diluted samples, respectively). Therefore, it was possible to calculate the correlation coefficients for the first 13 compounds only. All the calculated R^2 resulted to be greater than 0.95 (Table S3, Supplementary information). The good linearity obtained across the dynamic range, together with the evidence that further dilu-

tion implied the loss of detection of the less abundant compounds, led to the conclusion that the dilution used in this work is sufficient for satisfying a metabolomics quantification studies. However, it has to be taken into account that, although any saturation effect was shown regarding the ionization phenomenon, saturation at the chromatographic level was observed for few peaks. For instance, the highly abundant allyl GLS, i.e. sinigrin, showed broadening or splitting peaks at the highest concentration level (Fig. S3, Supplementary information). Wide, broadening or splitting peaks could lead to misinterpretation of the spectra when a pre-processing software is used for data analysis, requiring a mandatory visual inspection of the data. Therefore, in case of complex mixtures analysis, in which a wide variability in terms of abundance is observed, two distinct chromatographic runs for minor and major component quantification is recommended.

4. Concluding remarks

In this work, an evaluation of a suitable extraction, chromatographic and mass spectrometric method for the analysis of GLSs in cauliflower extracts was carried out by means of an untargeted profiling approach. Firstly, a RT extraction method was chosen because of the better results compared to the commonly used method. Secondly, a systematic evaluation of different chromatographic systems was carried out testing different phases and columns. The condition deriving from the addition of 0.1% of formic acid in both phases resulted the best for all the tested columns showing the superiority of formic acid on ammonium formate for the ionization of GLSs. Once chosen formic acid modified phases, a comparison of the four tested columns in the above mentioned condition showed that Luna Omega Polar column is the best stationary phase for the analysis of GLSs. The polar modifier linked to the surface of Luna column particles together with the C18 chains allow a good separation for a wide range of polar compounds, such as GLSs. However, when coupling 2 Kinetex XB columns in series a larger number of compounds was detected, probably due to the higher retention of a longer column setup and to a minor suppression of less abundant GLSs by coeluting compounds present in the matrix. Therefore, although Luna Polar column was considered to be the most suitable column among Luna, Kinetex XB and Kinetex Biphenyl, the coupling of two Kinetex XB columns was chosen for the final identification. The identification of 51 compounds, 24of which have never been found in cauliflower, was accomplished. Moreover, the linearity of the analytes response, in the range of concentration analyzed in this work, also demonstrated that the method is suitable for a quantitative analysis. However, in the cases of complex matrices in which a high compounds variability in terms of abundance is observed, different chromatographic runs at different dilution levels are suggested in order to avoid peak broadening due to saturation effects and consequently misinterpretation of results.

Finally, it has to be considered that, whereas the two in series column can be suitable for a comprehensive qualitative profiling of a complex mixture, a single column setup can be preferred when a quantitative analysis has to be carried out on a larger number of samples. Although the two Kinetex XB in series showed better results compared to one Luna Polar column, the slightly higher number of identified GLSs does not justify the choice of a longer column setup in terms of time and costs (consumed solvents) when performing a comparative quantitative analysis on a large number of samples. In conclusion, several aspects, such as sample preparation, chromatographic separation and mass spectrometric method have to be taken into account when performing an untargeted profiling on complex mixtures. The obtained better results on GLSs identification, compared to previous works on cauliflower analysis, showed the developed method to be optimal for both qualitative and quantitative analysis of GLSs in complex phytochemical mixtures.

Acknowledgment

This work was supported by PRIN 2015, project number 2015FFY97L_002.

Conflict of interest

The authors have declared no conflict of interest

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.talanta.2017.12.019.

References

[1] J. Riedl, S. Esslinger, C. Fauhl-Hassek, Review of validation and reporting of non-targeted fingerprinting approaches for food authentication, Anal. Chim. Acta 885 (2015) 17–32, https://doi.org/10.1016/j.aca.2015.06.003.
[2] S. Kim, J. Kim, E.J. Yun, K.H. Kim, Food metabolomics: from farm to human, Curr. Opin. Biotechnol. 37 (2016) 16–23, https://doi.org/10.1016/j.copbio.2015.09.004.
[3] G. La Barbera, A.L. Capriotti, C. Cavaliere, C.M. Montone, S. Piovesana, R. Samperi, R. Zenezini Chiozzi, A. Laganà, Liquid chromatography-high resolution mass spectrometry for the analysis of phytochemicals in vegetal-derived food and beverages, Food Res. Int. 100 (2017) 28–52, https://doi.org/10.1016/j.foodres.2017.07.080.
[4] R. Schuhmacher, R. Krska, W. Weckwerth, R. Goodacre, Metabolomics and metabolite profiling, Anal. Bioanal. Chem. 405 (2013) 5003–5004, https://doi.org/10.1007/s00216-013-6939-5.
[5] A. Kaufmann, Combining UHPLC and high-resolution MS: a viable approach for the analysis of complex samples?, TrAC Trends Anal. Chem. 63 (2014) 113–128, https://doi.org/10.1016/j.trac.2014.06.025.
[6] G.L. Barbera, A.L. Capriotti, C. Cavaliere, S. Piovesana, R. Samperi, R.Z. Chiozzi, A. Laganà, Comprehensive polyphenol profiling of a strawberry extract (Fragaria × ananassa) by ultra-high-performance liquid chromatography coupled with high-resolution mass spectrometry, Anal. Bioanal. Chem. (2017) 1–16, https://doi.org/10.1007/s00216-016-0159-8.
[7] R. Zhang, D.G. Watson, L. Wang, G.D. Westrop, G.H. Coombs, T. Zhang, Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based untargeted metabolite profiling of Leishmania parasites, J. Chromatogr. A 1362 (2014) 168–179, https://doi.org/10.1016/j.chroma.2014.08.039.
[8] R. Zenezini Chiozzi, A.L. Capriotti, C. Cavaliere, F. Ferraris, G. La Barbera, S. Piovesana, A. Laganà, Evaluation of column length and particle size effect on the untargeted profiling of a phytochemical mixture by using UHPLC coupled to high-resolution mass spectrometry, J. Sep. Sci. 40 (2017) 2541–2557, https://doi.org/10.1002/jssc.201700135.
[9] J.W. Fahey, A.T. Zalcmann, P. Talalay, The chemical diversity and distribution of glucosinolates and isothiocyanates among plants, Phytochemistry 56 (2001) 5–51.
[10] A.-S. Keck, J.W. Finley, Cruciferous vegetables: cancer protective mechanisms of glucosinolate hydrolysis products and selenium, Integr. Cancer Ther. 3 (2004) 5–12, https://doi.org/10.1177/1534735403261831.
[11] A.R. Kristal, J.W. Lampe, Brassica vegetables and prostate cancer risk: a review of the epidemiological evidence, Nutr. Cancer 42 (2002) 1–9, https://doi.org/10.1207/S15327914NC421_1.
[12] A. Steinbrecher, K. Nimptsch, A. Hüsing, S. Rohrmann, J. Linseisen, Dietary glucosinolate intake and risk of prostate cancer in the EPIC-Heidelberg cohort study, Int. J. Cancer 125 (2009) 2179–2186, https://doi.org/10.1002/ijc.24555.
[13] R.F. Mithen, M. Dekker, R. Verkerk, S. Rabot, I.T. Johnson, The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods, J. Sci. Food Agric. 80 (2000) 967–984, https://doi.org/10.1002/(SICI)1097-0010(20000515)80:7<967::AID-JSFA597>3.0.CO;2-V.
[14] L. Tang, Y. Zhang, Dietary isothiocyanates inhibit the growth of human bladder carcinoma cells, J. Nutr. 134 (2004) 2004–2010.
[15] G. Yuan, B. Sun, J. Yuan, Q. Wang, Effects of different cooking methods on health-promoting compounds of broccoli*, J. Zhejiang Univ. Sci. B 10 (2009) 580–588, https://doi.org/10.1631/jzus.B0920051.
[16] I. Radojčić Redovniković, T. Glivetić, K. Delonga, J. Vorkapić-Furač, Glucosinolates and their potential role in plant, Period. Biol. 110 (2008) 297–309.
[17] D.B. Clarke, Glucosinolates, structures and analysis in food, Anal. Methods 2 (2010) 310–325, https://doi.org/10.1039/B9AY00280D.
[18] N. Agerbirk, C.E. Olsen, Glucosinolate structures in evolution, Phytochemistry 77 (2012) 16–45, https://doi.org/10.1016/j.phytochem.2012.02.005.
[19] G. Glauser, F. Schweizer, T.C.J. Turlings, P. Reymond, Rapid profiling of intact glucosinolates in Arabidopsis leaves by UHPLC-QTOFMS using a charged surface hybrid column, Phytochem. Anal. PCA 23 (2012) 520–528, https://doi.org/10.1002/pca.2350.
[20] J.K. Troyer, K.K. Stephenson, J.W. Fahey, Analysis of glucosinolates from broccoli and other cruciferous vegetables by hydrophilic interaction liquid chromatogra-

phy, J. Chromatogr. A 919 (2001) 299–304, https://doi.org/10.1016/S0021-9673(01)00842-1.
[21] K.L. Wade, L.J. Garrard, J.W. Fahey, Improved hydrophilic interaction chromatography method for the identification and quantification of glucosinolates, J. Chromatogr. A 1154 (2007) 469–472, https://doi.org/10.1016/j.chroma.2007.04.034.
[22] C.L. Zrybko, E.K. Fukuda, R.T. Rosen, Determination of glucosinolates in domestic and wild mustard by high-performance liquid chromatography with confirmation by electrospray mass spectrometry and photodiode-array detection, J. Chromatogr. A 767 (1997) 43–52, https://doi.org/10.1016/S0021-9673(96)01068-0.
[23] J. Sun, M. Zhang, P. Chen, GLS-finder: a platform for fast profiling of glucosinolates in Brassica vegetables, J. Agric. Food Chem. 64 (2016) 4407–4415, https://doi.org/10.1021/acs.jafc.6b01277.
[24] G. Bianco, N. Agerbirk, I. Losito, T.R.I. Cataldi, Acylated glucosinolates with diverse acyl groups investigated by high resolution mass spectrometry and infrared multiphoton dissociation, Phytochemistry 100 (2014) 92–102, https://doi.org/10.1016/j.phytochem.2014.01.010.
[25] P. Franco, S. Spinozzi, E. Pagnotta, L. Lazzeri, L. Ugolini, C. Camborata, A. Roda, Development of a liquid chromatography-electrospray ionization-tandem mass spectrometry method for the simultaneous analysis of intact glucosinolates and isothiocyanates in Brassicaceae seeds and functional foods, J. Chromatogr. A 1428 (2016) 154–161, https://doi.org/10.1016/j.chroma.2015.09.001.
[26] A.M. Ares, M.J. Nozal, J. Bernal, Development and validation of a liquid chromatography-tandem mass spectrometry method to determine intact glucosinolates in bee pollen, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 1000 (2015) 49–56, https://doi.org/10.1016/j.jchromb.2015.07.017.
[27] G. Bianco, R. Pascale, F. Lelario, S.A. Bufo, T.R.I. Cataldi, The investigation of glucosinolates by mass spectrometry, in: J.-M. Mérillon, K.G. Ramawat (Eds.), Glucosinolates, Springer International Publishing, 2016, pp. 1–32, https://doi.org/10.1007/978-3-319-26479-0_12-1.
[28] T. Pluskal, S. Castillo, A. Villar-Briones, M. Orešič, MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, BMC Bioinforma. 11 (2010) 395, https://doi.org/10.1186/1471-2105-11-395.
[29] E.L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H.P. Singer, J. Hollender, Identifying small molecules via high resolution mass spectrometry: communicating confidence, Environ. Sci. Technol. 48 (2014) 2097–2098, https://doi.org/10.1021/es5002105.
[30] K. Oerlemans, D.M. Barrett, C.B. Suades, R. Verkerk, M. Dekker, Thermal degradation of glucosinolates in red cabbage, Food Chem. 95 (2006) 19–29, https://doi.org/10.1016/j.foodchem.2004.12.013.
[31] V. Rungapamestry, A.J. Duncan, Z. Fuller, B. Ratcliffe, Effect of cooking brassica vegetables on the subsequent hydrolysis and metabolic fate of glucosinolates, Proc. Nutr. Soc. 66 (2007) 69–81, https://doi.org/10.1017/S0029665107005319.
[32] T.R.I. Cataldi, F. Lelario, D. Orlando, S.A. Bufo, Collision-induced dissociation of the A + 2 isotope ion facilitates glucosinolates structure elucidation by electrospray ionization-tandem mass spectrometry with a linear quadrupole ion trap, Anal. Chem. 82 (2010) 5686–5696, https://doi.org/10.1021/ac100703w.
[33] T.R.I. Cataldi, A. Rubino, F. Lelario, S.A. Bufo, Naturally occurring glucosinolates in plant extracts of rocket salad (Eruca sativa L.) identified by liquid chromatography coupled with negative ion electrospray ionization and quadrupole ion-trap mass spectrometry, Rapid Commun. Mass Spectrom. 21 (2007) 2374–2388, https://doi.org/10.1002/rcm.3101.
[34] L.-Z. Lin, J. Sun, P. Chen, R.-W. Zhang, X.-E. Fan, L.-W. Li, J.M. Harnly, Profiling of Glucosinolates and Flavonoids in Rorippa indica (Linn.) Hiern. (Cruciferae) by UHPLC-PDA-ESI/HRMSn, J. Agric. Food Chem. 62 (2014) 6118–6129, https://doi.org/10.1021/jf405538d.
[35] J.B. Bialecki, J. Ruzicka, C.S. Weisbecker, M. Haribal, A.B. Attygalle, Collision-induced dissociation mass spectra of glucosinolate anions, J. Mass Spectrom. Jms. 45 (2010) 272–283, https://doi.org/10.1002/jms.1711.
[36] T. Doheny-Adams, K. Redeker, V. Kittipol, I. Bancroft, S.E. Hartley, Development of an efficient glucosinolate extraction method, Plant Methods 13 (2017) 17, https://doi.org/10.1186/s13007-017-0164-8.
[37] F.A. Mellon, R.N. Bennett, B. Holst, G. Williamson, Intact glucosinolate analysis in plant extracts by programmed cone voltage electrospray LC/MS: performance and comparison with LC/MS/MS methods, Anal. Biochem. 306 (2002) 83–91, https://doi.org/10.1006/abio.2002.5677.
[38] T. Mohn, B. Cutting, B. Ernst, M. Hamburger, Extraction and analysis of intact glucosinolates—a validated pressurized liquid extraction/liquid chromatography–mass spectrometry protocol for Isatis tinctoria, and qualitative analysis of other cruciferous plants, J. Chromatogr. A 1166 (2007) 142–151, https://doi.org/10.1016/j.chroma.2007.08.028.
[39] C18, C8 HPLC Column and UHPLC Column Application Search using Kinetex Columns from Phenomenex. ⟨https://www.phenomenex.com/Kinetex/ApplicationSearch?Phase=XB-C18⟩ (Accessed 11 October 2017), n.d.
[40] F. Lelario, G. Bianco, S.A. Bufo, T.R. Cataldi, Establishing the occurrence of major and minor glucosinolates in Brassicaceae by LC-ESI-hybrid linear ion-trap and Fourier-transform ion cyclotron resonance mass spectrometry, Phytochemistry 73 (2012) 74–83, https://doi.org/10.1016/j.phytochem.2011.09.010.
[41] M. Gratacós-Cubarsí, A. Ribas-Agustí, J.A. García-Regueiro, M. Castellari, Simultaneous evaluation of intact glucosinolates and phenolic compounds by UPLC-DAD-MS/MS in Brassica oleracea L. var. botrytis, Food Chem. 121 (2010) 257–263, https://doi.org/10.1016/j.foodchem.2009.11.081.
[42] J. Nilsson, K. Olsson, G. Engqvist, J. Ekvall, M. Olsson, M. Nyman, B. Åkesson, Variation in the content of glucosinolates, hydroxycinnamic acids, carotenoids, total antioxidant capacity and low-molecular-weight carbohydrates in Brassica vegetables, J. Sci. Food Agric. 86 (2006) 528–538, https://doi.org/10.1002/jsfa.2355.
[43] S.R. Bhandari, J.S. Jo, J.G. Lee, Comparison of glucosinolate profiles in different tissues of nine brassica crops, Molecules 20 (2015) 15827–15841, https://doi.org/10.3390/molecules200915827.
[44] J. Volden, G.B. Bengtsson, T. Wicklund, Glucosinolates, L-ascorbic acid, total phenols, anthocyanins, antioxidant capacities and colour in cauliflower (Brassica oleracea l. ssp. botrytis); effects of long-term freezer storage, Food Chem. 112 (2009) 967–976, https://doi.org/10.1016/j.foodchem.2008.07.018.
[45] S.I. Vicas, A.C. Teusdea, M. Carbunar, S.A. Socaci, C. Socaciu, Glucosinolates profile and antioxidant capacity of Romanian Brassica vegetables obtained by organic and conventional agricultural practices, Plant Foods Hum. Nutr. 68 (2013) 313–321, https://doi.org/10.1007/s11130-013-0367-8.
[46] C. Cavaliere, P. Foglia, R. Gubbiotti, P. Sacchetti, R. Samperi, A. Laganà, Rapid-resolution liquid chromatography/mass spectrometry for determination and quantitation of polyphenols in grape berries, Rapid Commun. Mass Spectrom. 22 (2008) 3089–3099, https://doi.org/10.1002/rcm.3705.
[47] L. Song, J.J. Morrison, N.P. Botting, P.J. Thornalley, Analysis of glucosinolates, isothiocyanates, and amine degradation products in vegetable extracts and blood plasma by LC-MS/MS, Anal. Biochem. 347 (2005) 234–243, https://doi.org/10.1016/j.ab.2005.09.040.
[48] N. Pellegrini, E. Chiavaro, C. Gardana, T. Mazzeo, D. Contino, M. Gallo, P. Riso, V. Fogliano, M. Porrini, Effect of different cooking methods on color, phytochemical concentration, and antioxidant capacity of raw and frozen brassica vegetables, J. Agric. Food Chem. 58 (2010) 4310–4321, https://doi.org/10.1021/jf904306r.

Chapter 5:

Untargeted approach and nutrition

1 Untargeted analysis for nutritional biomarkers discovery

The third section of the thesis concerns the application of a typical untargeted approach for the identification of nutritional biomarkers. In the next paragraphs, a work about the identification of intake biomarkers after meat and dairy products consumption will be presented. The issue under investigation has been treated by means of an untargeted analysis, focusing in particular on the development of a reliable data analysis approach.

1.1 Background

As already discussed, several epidemiological studies have reported that diet is strictly related to certain diseases (Chapter 2, paragraph 5.3). However, proving a causal relationship diet-disease is particularly difficult. Indeed, quantitative assessment of food intake is usually accomplished by the use of food frequency questionnaires, 24-h recalls, and weighed food diaries [118, 119]. These methods resulted to be very imprecise due to their subjective nature[118–120], providing only moderately reliable information and not allowing to certainly prove diet-disease relationships. Dietary biomarkers have been shown to be a powerful tool for the assessment of dietary exposure in clinical and observational studies [68, 121]. However, the detection and identification of a reliable biomarker is a particularly difficult issue requiring that, study design, time of sampling and analytical method are correctly set and, overall, that data analysis allows the selection of a biomarker fulfilling some specific validation criteria. On the presented work, a data analysis workflow was developed to identify valid and reliable meat and dairy intake biomarkers.

The interest in assessing meat and dairy intake biomarkers is due to the evidence that red meat intake has been associated to increased risk of colorectal cancer [122], cardiovascular diseases[123], type 2 diabetes[124] and mortality[123]. However, a causal relation between red meat consumption and negative health outcomes is still controversial[125] because of the difficulty in the identification of reliable meat intake biomarkers [126]. Indeed, meat source and cooking methods strictly affects meat composition, moreover, the composition of meat is very similar to that of human proteins as well as that of other foods rich of proteins. As consequence, it is particularly difficult to distinguish between metabolites derived from

endogenous proteins or other protein-rich foods and metabolites derived from meat. In this work, an untargeted metabolomics workflow was developed in order to identify and distinguish specific and reliable meat and dairy intake biomarkers.

1.2 Study design

Firstly, in order to identify reliable dairy and meat intake biomarkers, a proper study design should be set. The cross-over intervention study is the most robust intervention study, because in the data analysis step each subject is the control of himself, thus avoiding any eventual effects due to the individual variability. Therefore, 17 men and women were recruited and the study was performed as a randomized, controlled, cross-over single meal intervention with 2 study periods of 1 day and a wash-out period of 7 days between periods. The two periods consisted of a test day where subjects were served with either a test meal with high content of milk protein or a test meal with high content of meat protein in randomized order. Twenty-four hours before and after the intervention visits subjects were asked to restrain from ingesting similar dinner meals.

Prior to and after the test meal, four urine samples were collected: (1) from 24 hours prior the intervention visit (T 0), (2) from 0 until 2 hours following intake of the test meal (T 0–2), (3) from 2 until 4 hours after the test meal (T 2–4), and (4) during the rest of the day for 24 hours (T 4–24).

1.3 Chromatographic and mass spectrometric method

Urine samples were centrifuged and diluted in a volume ratio 1:1 with aqueous 5% 30:70 (v/v) ACN:MeOH containing a solution with 7 Internal Standard as previously described [127]. Attention has to be posed when performing an untargeted analysis, in order to control the quality of the analytical platform and consequently the reliability of the obtained data. Therefore, an external metabolite standard mixture with 44 different compounds [127] and pooled samples, containing equal amounts of all urine samples analyzed, were prepared and added to separate wells on all plates for subsequent control of batch drift in the data analysis.

Moreover, to minimize intra-individual variation due to plate differences, samples from the same person were randomized within one plate.

The urine samples were analyzed on an Acquity UHPLC coupled through ESI to a orthogonal acceleration Premier qTOF mass-spectrometer. The chromatographic separation was accomplished in reversed phase on a HSS T3 C18 column as described previously [127]. A full scan acquisition in both positive and negative polarity modes was done. Blanks and external metabolomics standard mixtures were injected after every 30 samples, the pooled samples were injected every 50 samples.

1.4 Data analysis

The obtained data needed a pre-treatment step as explained before (Chapter 4, paragraph 1.3). The features obtained with MZmine software were imported into Matlab® and further filtered. Features present in the blank, features early or late eluting, features with low intensity were excluded [128]. A further correction to remove inter-batch variation was carried out. Finally, features were assigned as one unique feature group if they had good correlation coefficients. The pre-treatment step was necessary in order to obtain a final number of possibly reliable metabolites.

Once selected the compounds, the metabolites discriminating between the two diets had to be selected. In order to achieve this goal, two different statistical approaches were applied. The univariate analysis on the urine samples collected before the test meal (T 0), during the first 2 (T 0-2), 4 (T 2-4) and 24 hours (T4-24) after the test meal was applied in order to identify the features significantly different between the two meals and with a proper excretion profile over the 24 hours[129]. The multivariate data analysis on the pooled 24-hours urine samples was applied because twenty-four hours urine are usually the preferred samples to monitor food intake in observational studies. Moreover, two different multivariate data analysis approaches were compared on the 24 hours urine: Partial Least Square-Discriminant Analysis (PLS-DA)[130] and Multilevel Partial Least-Squares Discriminant Analysis (ML-PLSDA) [131]. The use of these two different multivariate methods allows to distinguish those markers that are not affected by inter-individual variation (PLS-DA) from those that are affected from the different individual response of each subjects (ML-PLSDA). Compounds resulting significant for discriminating the two

diets from univariate data analysis and at least one of the two multivariate approaches have been selected as intake biomarkers. However, since the purpose of the study was the achievement of a list of reliable biomarkers, a further and deeper investigation was carried out to select those metabolites resulting even more robust.

Once the features were selected as possible biomarkers, their excretion profiles were visually inspected and metabolites classified based on their kinetics: 1)early and fast biomarkers, those markers that are excreted right after the consumption and quickly within the 24 hours after the consumption; 2) early and slow biomarkers, those excreted after the consumption but reaching the baseline in more than 24 hours; 3) late and slow biomarkers, those excreted at least after 2 hours after the consumption and reaching the baseline in more than 24 hours. Moreover, only the features which clearly increased after the intake of one specific meal were kept for further investigation. The rest of the features were removed because the excretion of the metabolites increased over the time in both groups that means they were present in other food than the test meal. Figure 5.1 shows some typical excretion kinetics profiles.

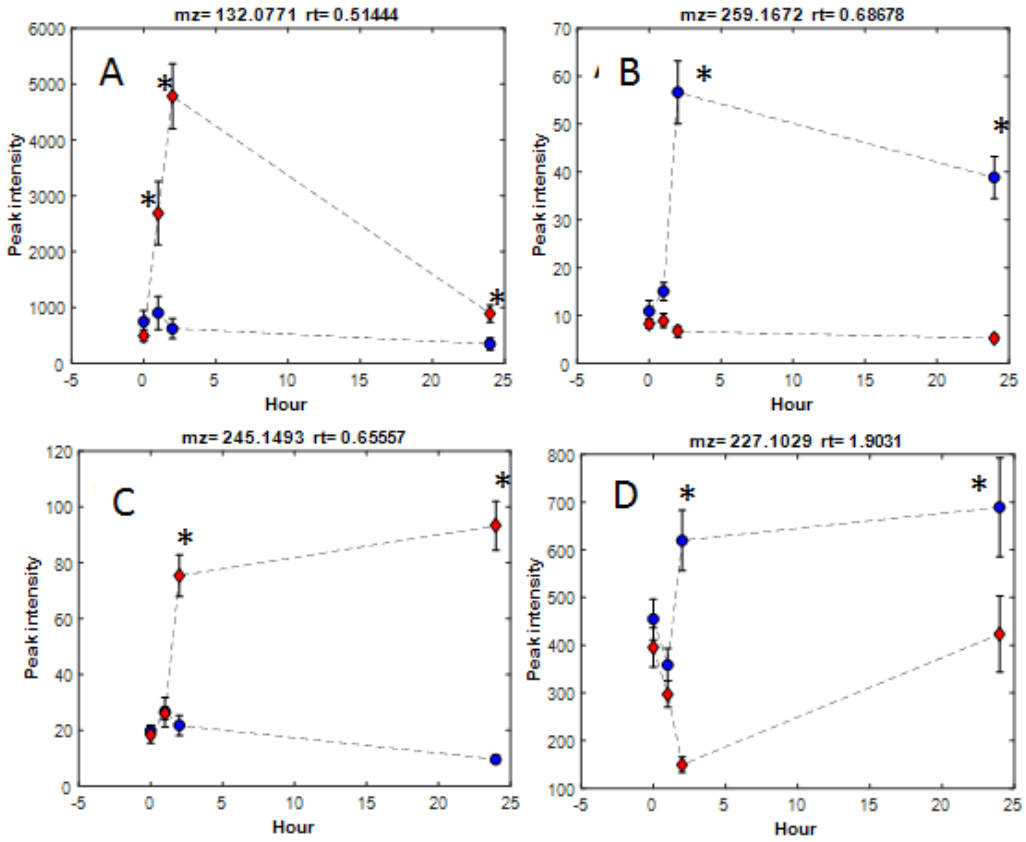


Figure 5.1: excretion kinetics of A) a level 1 biomarker, B) a level 2 biomarker, C) a level 3 biomarker, D) an excluded biomarker.

Whereas the kinetics classification is useful from a practical point of view (knowing the collection time of the markers), the last constrain is essential in order to select only specific biomarkers, i.e biomarkers not deriving from the intake of food other than that ingested during the single meal test. The kinetic levels are reported in Table 5.1 and 5.2 for each selected biomarker.

1.5 Identification

Once selected the candidate biomarkers, the final investigation was carried out. Firstly, a search within an in-house database, containing retention time information and MS spectra of reference substances, was carried out. Secondly, for unknown compounds MS/MS fragmentation experiments were conducted in product-ion scan for their structural characterization. The CID was set to 10, 20, and 30 eV to obtain a comprehensive fragmentation pattern. Thirdly, searches of the fragments and parent ions were performed on different databases such as Human Metabolome Database, Metlin, Chembank, Chemspider, mzCloud and similar websites. Lastly, some of the metabolites were confirmed with authentic standard compounds. Whether the standards were not commercially available, they were chemically or biologically synthesized. The intake biomarkers were classified as level I-IV in accordance with the criteria established by Sumner et al. as a standard for metabolomics investigations[53]. Briefly, level I was assigned to compounds matched to standards: level II was assigned to compounds matched to a MS/MS spectrum published in the literature; level III was assigned to compounds identified at the level of compound class and level IV was assigned to unknown compounds. The resulting biomarkers are reported in Table 5.1 and Table 5.2 for meat and dairy biomarkers, respectively, together with the suggested identity, the identity confidence level, the significance at the multivariate data analysis and the kinetics level.

The found biomarkers were searched in the literature to verify their plausibility as meat or dairy biomarkers. Three detected metabolites, in particular Carnosine, Taurine and Creatine, have been already found as meat intake biomarkers[126]. Tyramine-O-sulfate, and phenyl lactic acid instead, has been already found as dairy biomarker[132]. Moreover, several di or tripeptides have been found such as Hydroxyprolyl-Proline, Prolyl-hydroxyproline, Valyl-Proline, Pyroglutamyl-Proline, Leucyl-proline, Leucyl-

Hydroxyproline. Di/Tripepetides are an incomplete breakdown product of protein digestion or protein catabolism, therefore, they can likely be related to the consumption of protein-rich foods. These dipeptides have not yet been identified in human tissues or biofluids. Other found metabolites can be correlated to other metabolic pathways such as 3-Hydroxy-2-methyl-butanoic acid (HMBA), that is a normal urinary metabolite involved in the isoleucine catabolism[133]. Other found metabolites have not been correlated to any particular metabolic pathway such as Methoxysalicylic acid or p-hydroxyphenylacetic acid that can be found throughout all human tissues and biofluids.

Table 5.1. Markers of meat-based meal

Metabolite (Level identification)	Molecular formula	MLPLSDA	PLSDA	Excretion kinetics
carnosine[I]	$C_9H_{14}N_4O_3$	*	*	1
taurine[I]	$C_2H_7NO_3S$	*		2
creatine[I]	$C_4H_9N_3O_2$	*	*	2
Hydroxyprolyl-Proline[III]	$C_{10}H_{16}N_2O_4$	*	*	3
Prolyl-Hydroxyproline[II]	$C_{10}H_{16}N_2O_4$	*	*	3
thiacremonone[IV]	$C_6H_8O_3S$	*		3
Leucyl/Isoleucyl-Hydroxyproline[III]	$C_{11}H_{20}N_2O_4$	*	*	3
Hydroxyprolyl-Leucine/Isoleucine[II]	$C_{11}H_{20}N_2O_4$	*		1
Pro-Phe-Gly[II]	$C_{16}H_{21}N_3O_4$	*	*	3
acetyl leucine[IV]	$C_8H_{15}NO_3$		*	2
unknown[IV]	$C_{10}H_9NO_5$		*	2
unknown [IV]		*		1
methoxysalicylic acid sulfate [II]	$C_8H_8O_7S$	*	*	2
Unknown sulfate[IV]	$C_{18}H_{24}SNO_7$	*	*	2
unknown[IV]	$C_{12}H_{19}NO_3$	*	*	2
Unknown sulfate[IV]	$C_{18}H_{23}SNO_7$	*	*	2
Erythronilic acid [IV]	$C_5H_{10}O_3$	*	*	2

Whereas some of the found metabolites showed some metabolic pathways that could, possibly, be linked to the metabolism of meat and dairy intake, other found metabolites have been correlated to other sources. Thiacremonone, for instance, is formed in the Maillard reaction of 4-Hydroxy-2,5-dimethyl-3(2H)-furanone, that is a compound found in strawberry. P-hydroxyphenylacetic acid can be found throughout numerous foods such as Olives, Cocoa beans, Oats, and Mushrooms. Phenylacetylglycine, that is an acyl glycine, could be a minor metabolites of fatty acids but also of fruits and vegetables[133].

Cyclohexane carboxilic acid glycine and glutamine hydroxyphenylacetic acid have already been found as biomarkers of strawberry[134] and plums respectively, suggesting they are not specific of the dairy intake. Trimethylphenol is used as a food additive like flavouring ingredient[133].

Table 5.2. Markers of dairy-based meal

Metabolite (Level identification)	Molecular formula	MLPLSDA	PLSDA	Excretion kinetics
unknown[IV]		*	*	3
tyramine sulfate[I]	$C_8H_{11}NO_4S$	*	*	2
Valyl-Proline[II]	$C_{10}H_{18}N_2O_3$	*		1
unknown[IV]	$C_{14}H_{17}N_3O_4S$		*	3
unknown[IV]	$C_{10}H_{19}NO_3$	*	*	2
Pyroglutamine proline[II]	$C_{10}H_{14}N_2O_4$	*		3
glutamine hydroxyphenylacetic acid [II]	$C_{13}H_{16}N_2O_5$	*		3
Ile/Leu-Proline[II]	$C_{11}H_{20}N_2O_3$	*	*	2
4-hydroxyphenylacetic acid sulfate[II]	$C_8H_8O_6S$	*	*	3
unknown[IV]	$C_{13}H_{17}NO_3$	*	*	2
unknown[IV]	$C_{13}H_{17}NO_3$	*	*	2
unknown[IV]		*	*	2
4-hydroxyphenylacetic acid[I]	$C_8H_8O_3$	*		1
Phenylacetylglycine[I]	$C_{10}H_{11}NO_3$	*	*	1
unknown[IV]	$C_{18}H_{26}N_2O_4$	*	*	2
phenyllactic acid[I]	$C_9H_{10}O_3$	*	*	2
Cyclohexane carboxilic acid glycine[I]	$C_9H_{15}NO_3$	*		1
unknown[IV]		*	*	2
3,5-dimethylphenol sulfate[I]	$C_8H_{10}O$	*	*	3
Trimethylphenol sulfate[III]	$C_9H_{12}O$	*	*	3

The literature research is important to evaluate the plausibility of the tentatively identified biomarkers. Some of the found information suggest that some biomarkers are correlated to other kinds of food, and thus they cannot be considered as specific and robust biomarkers. However, before stating such conclusions, a definitive identification by means of the synthesis of the proper standards and the confirmation of retention times and fragmentation should be accomplished. Moreover, more reliable studies should be carried out to observe the dairy and meat intake effect over a longer period of test.

1.6 Conclusion

As it has been shown, several aspects of the untargeted workflow should be considered to allow a reliable biomarker detection and identification. Firstly, a robust study design together with a suitable sample collection is needed. Secondly, a proper sample analysis should be carried out, checking the method quality and stability over the time. Thirdly, a strong data analysis should be carried out. In this section, all the aspects related to the development of an untargeted approach have been treated but, particular attention has been posed on the data analysis and the metabolites identification. The combination of the univariate and multivariate data analysis allowed to obtain a list of reliable biomarkers. However, by a deeper investigation on the kinetics profiles, it was possible to find some markers deriving from the consumption of other kinds of food. Therefore, only those markers being specific of the dairy either meat based diet were selected. Moreover, the identification step allowed to select the markers that showed a reasonable biological meaning. Finally, the comparison between PLSDA and ML-PLSDA allowed to distinguish those markers that are not affected or affected from the different individual response.

Information like the plausibility of a biomarker, its specificity, its dependence on the individual response, its kinetics of excretion, are necessary for the choice of the suitable biomarker in the observational and epidemiological studies on diet-disease relationship investigation. In this thesis, it was shown how data analysis and identification are essential to obtain these information and then to identify a reliable biomarker. However, further improvements in data analysis tools and in the availability of complete databases are needed in order to simplify the identification step that is, by far, the major bottleneck of untargeted workflows.

1.7 Acknowledgments

The study on meat and dairy intake biomarkers identification (MEDA study) has been carried out during six months of visit at the University of Copenhagen (Denmark) under the supervision of Prof. Lars Ove Dragsted. MEDA study was supported by grants from the Danish Dairy Foundation and from the Danish Meat Research Institute.

Conclusions

As shown in the previous sections, several aspects should be taken into account when analyzing complex matrices like food by means of different approaches. Sample preparation, chromatographic and mass spectrometric methods and data analysis have to be treated in different ways based on the purpose of the analysis.

Whereas in targeted methods sample preparation should be as selective and specific as possible towards the few analytes of interest, in suspected and untargeted methods it should be comprehensive enough to extract the whole class under investigation or the whole metabolome, respectively. For this reason, targeted approaches mainly focus on the development of a proper sample preparation, clean-up and enrichment, in order to identify and quantify a specific class of analytes. Of course, several aspects related to sample preparation have to be considered: recoveries, time, costs, and applicability to different matrices or to a wide range of diverse compounds. The overall process efficiency of the developed method has to be a compromise between performance and application easiness and rapidity.

The chromatographic and mass spectrometric method optimization is of course the key step in both the targeted and suspected/ untargeted approaches. However, whereas in the first case the method should be optimized for the few analytes of interest, in the suspected and untargeted approaches should be, again, suitable for the analysis of a wide class of compounds. The method optimization for a small class of compounds is of course easier compared to a large class of compounds, that is the reason why few chromatographic methods evaluations in suspected and untargeted approaches analysis have been published so far. However, a chromatographic evaluation on the real system is always preferable compared to an evaluation on a simulated systems such as a mixture of standards. The chromatographic behaviour strictly affects the detection of compounds, therefore, phases, gradient, temperature should be optimized in order to achieve the best separation of the compounds present in a complex mixture. Also, the coupling of the chromatographic system to the ESI source and to the mass spectrometer have to be considered. The right parameters in order to have an effective ionization, detection and fragmentation should be chosen. Even in this case, optimizing mass spectrometric parameters for a few analytes is different

compared to a huge range of analytes. Therefore, some completely different methods are applied in targeted and suspected/untargeted approaches. For instance, considering the mass spectrometric acquisition, MRM and DDA are usually carried out in targeted and suspected analysis, respectively. In untargeted analysis, instead, DDA acquisition or two steps workflows, i.e. full scan acquisition and then selected ions fragmentation, are carried out. Therefore, also in this case, the different purposes of the analysis imply some differences in the way the analytical method is developed.

Finally, the data analysis and identification step is a particularly though issue. Luckily, targeted approaches do not suffer from the difficulties of the time consuming data analysis step that suspected and untargeted approaches have to face. Indeed, whereas in targeted analysis the compounds to analyze are already known, in the case of suspected and, mostly, untargeted approaches the data analysis and identification result the major bottlenecks. Whereas in the suspected approach the only difficulty is in the identification of a wide range of "expected" compounds, in untargeted approach both the step of the selection of the markers of interest and the step of their identification require efforts and time. Further improvements in data analysis tools and in the availability of complete databases are needed in order to simplify the suspected and untargeted workflows and making them more accessible to a wider researchers community.

In conclusion, food science finds in mass spectrometry a powerful tool for food safety and quality assessment and for studies on nutrition. Several approaches are available and can be exploited to accomplish different aims. The advantages of each approach can be exploited based on the purpose of the study like, for instance, targeted for the quantification of traces of contaminants in foods, suspected for metabolic profiling of food bioactive compounds, untargeted for nutritional biomarkers discovery. Of course, whatever approach is used, a previous method development and optimization is necessary to obtain the best results.

References

1. Cifuentes A (2013) Foodomics: Principles and Applications. In: Cifuentes A (ed) Foodomics. John Wiley & Sons, Inc., pp 1–13

2. La Barbera G, Capriotti AL, Cavaliere C, Montone CM, Piovesana S, Samperi R, Zenezini Chiozzi R, Laganà A (2017) Liquid chromatography-high resolution mass spectrometry for the analysis of phytochemicals in vegetal-derived food and beverages. Food Res Int 100:28–52. doi: 10.1016/j.foodres.2017.07.080

3. Rome Declaration and Plan of Action. http://www.fao.org/docrep/003/w3613e/w3613e00.HTM. Accessed 18 Sep 2017

4. Deblonde T, Cossu-Leguille C, Hartemann P (2011) Emerging pollutants in wastewater: A review of the literature. Int J Hyg Environ Health 214:442–448. doi: 10.1016/j.ijheh.2011.08.002

5. Daughton CG (2004) Non-regulated water contaminants: emerging research. Environ Impact Assess Rev 24:711–732. doi: 10.1016/j.eiar.2004.06.003

6. Picó Y (2012) Chemical Analysis of Food: Techniques and Applications. Academic Press

7. Mulder PPJ, Sánchez PL, These A, Preiss-Weigert A, Castellari M (2015) Occurrence of Pyrrolizidine Alkaloids in food. EFSA Support Publ 12:n/a-n/a. doi: 10.2903/sp.efsa.2015.EN-859

8. Capriotti AL, Caruso G, Cavaliere C, Foglia P, Samperi R, Laganà A (2012) Multiclass mycotoxin analysis in food, environmental and biological matrices with chromatography/mass spectrometry. Mass Spectrom Rev 31:466–503. doi: 10.1002/mas.20351

9. EUR-Lex - 32006R1881 - EN - EUR-Lex. http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32006R1881. Accessed 30 Sep 2017

10. (2017) Mycotoxins. http://www.biomin.net/en/species/ruminants/mycotoxins/. Accessed 18 Sep 2017

11. Biesalski HK, Dragsted LO, Elmadfa I, Grossklaus R, Müller M, Schrenk D, Walter P, Weber P (2009) Bioactive compounds: Definition and assessment of activity. Nutrition 25:1202–1205. doi: 10.1016/j.nut.2009.04.023

12. DeFelice SL (1995) The nutraceutical revolution: its impact on food industry R&D. Trends Food Sci Technol 6:59–61. doi: 10.1016/S0924-2244(00)88944-X

13. Buchanan BB, Gruissem W, Jones RL (2015) Biochemistry and Molecular Biology of Plants. John Wiley & Sons

14. Dewick PM (2001) Alkaloids. In: Med. Nat. Prod. John Wiley & Sons, Ltd, pp 291–403

15. Crozier A, Jaganath IB, Clifford MN (2006) Phenols, Polyphenols and Tannins: An Overview. In: Crozier A, Clifford MN, Ashihara H (eds) Plant Second. Metab. Blackwell Publishing Ltd, pp 1–24

16. Herrmann KM, Weaver LM (1999) THE SHIKIMATE PATHWAY. Annu Rev Plant Physiol Plant Mol Biol 50:473–503. doi: 10.1146/annurev.arplant.50.1.473

17. Robbins RJ (2003) Phenolic acids in foods: an overview of analytical methodology. J Agric Food Chem 51:2866–2887. doi: 10.1021/jf026182t

18. Winkel-Shirley B (1999) Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. Physiol Plant 107:142–149. doi: 10.1034/j.1399-3054.1999.100119.x

19. Clifford MN (1999) Chlorogenic acids and other cinnamates – nature, occurrence and dietary burden. J Sci Food Agric 79:362–372. doi: 10.1002/(SICI)1097-0010(19990301)79:3<362::AID-JSFA256>3.0.CO;2-D

20. Dixon RA, Paiva NL (1995) Stress-Induced Phenylpropanoid Metabolism. Plant Cell Online 7:1085–1097. doi: 10.1105/tpc.7.7.1085

21. Tsao R (2010) Chemistry and Biochemistry of Dietary Polyphenols. Nutrients 2:1231–1246. doi: 10.3390/nu2121231

22. Crozier A, Jaganath IB, Clifford MN (2009) Dietary phenolics: chemistry, bioavailability and effects on health. Nat Prod Rep 26:1001–1043. doi: 10.1039/b802662a

23. Hemingway RW (1998) Practical polyphenolics: from structure to molecular recognition and physiological action, by Edwin Haslam.[Book review]. J. Nat. Prod. 1454-1455

24. Santos-Buelga C, Scalbert A (2000) Proanthocyanidins and tannin-like compounds – nature, occurrence, dietary intake and effects on nutrition and health. J Sci Food Agric 80:1094–1117. doi: 10.1002/(SICI)1097-0010(20000515)80:7<1094::AID-JSFA569>3.0.CO;2-1

25. Landete JM (2011) Ellagitannins, ellagic acid and their derived metabolites: A review about source, metabolism, functions and health. Food Res Int 44:1150–1160. doi: 10.1016/j.foodres.2011.04.027

26. Ascacio-Valdes JA, Buenrostro-Figueroa JJ, Aguilera-Carbo A, Prado-Barragan A, Rodriguez-Herrera R, Aguilar CN (2011) Ellagitannins: Biosynthesis, biodegradation and biological properties. J Med Plants Res 5:4696–4703.

27. Hellström JK, Törrönen AR, Mattila PH (2009) Proanthocyanidins in Common Food Products of Plant Origin. J Agric Food Chem 57:7899–7906. doi: 10.1021/jf901434d

28. Cataldi TRI, Rubino A, Lelario F, Bufo SA (2007) Naturally occurring glucosinolates in plant extracts of rocket salad (Eruca sativa L.) identified by liquid chromatography coupled with negative ion electrospray ionization and quadrupole ion-trap mass spectrometry. Rapid Commun Mass Spectrom 21:2374–2388. doi: 10.1002/rcm.3101

29. Fahey JW, Zalcmann AT, Talalay P (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56:5–51.

30. Graser G, Schneider B, Oldham NJ, Gershenzon J (2000) The methionine chain elongation pathway in the biosynthesis of glucosinolates in Eruca sativa (Brassicaceae). Arch Biochem Biophys 378:411–419. doi: 10.1006/abbi.2000.1812

31. Halkier BA, Gershenzon J (2006) Biology and biochemistry of glucosinolates. Annu Rev Plant Biol 57:303–333. doi: 10.1146/annurev.arplant.57.032905.105228

32. Gibson TM, Ferrucci LM, Tangrea JA, Schatzkin A (2010) Epidemiological and Clinical Studies of Nutrition. Semin Oncol 37:282–296. doi: 10.1053/j.seminoncol.2010.05.011

33. European Food Safety Authority | Trusted science for safe food. http://www.efsa.europa.eu/. Accessed 18 Sep 2017

34. Gilsenan MB (2011) Nutrition & health claims in the European Union: A regulatory overview. Trends Food Sci Technol 22:536–542. doi: 10.1016/j.tifs.2011.03.004

35. Senyuva HZ, Gökmen V, Sarikaya EA (2015) Future perspectives in Orbitrap™-high-resolution mass spectrometry in food analysis: a review. Food Addit Contam Part Chem Anal Control Expo Risk Assess 32:1568–1606. doi: 10.1080/19440049.2015.1057240

36. Holčapek M, Jirásko R, Lísa M (2012) Recent developments in liquid chromatography-mass spectrometry and related techniques. J Chromatogr A 1259:3–15. doi: 10.1016/j.chroma.2012.08.072

37. Zubarev RA, Makarov A (2013) Orbitrap mass spectrometry. Anal Chem 85:5288–5296. doi: 10.1021/ac4001223

38. Wang X, Wang S, Cai Z (2013) The latest developments and applications of mass spectrometry in food-safety and quality analysis. TrAC Trends Anal Chem 52:170–185. doi: 10.1016/j.trac.2013.08.005

39. (1998) Handbook of Instrumental Techniques for Analytical Chemistry. J Liq Chromatogr Relat Technol 21:3072–3076. doi: 10.1080/10826079808006889

40. Lough WJ, Wainer IW (1995) High Performance Liquid Chromatography: Fundamental Principles and Practice. CRC Press

41. Hoffmann E de, Stroobant V (2007) Mass Spectrometry: Principles and Applications. John Wiley & Sons

42. Mass Spectrometry :: Introduction, Principle of Mass Spectrometry, Components of Mass Spectrometer, Applications. http://www.premierbiosoft.com/tech_notes/mass-spectrometry.html. Accessed 19 Sep 2017

43. Gaskell SJ (1997) Electrospray: Principles and Practice. J Mass Spectrom 32:677–688. doi: 10.1002/(SICI)1096-9888(199707)32:7<677::AID-JMS536>3.0.CO;2-G

44. Bristol University - High Performance Liquid Chromatography Mass Spectrometry (HPLC/MS). http://www.bris.ac.uk/nerclsmsf/techniques/hplcms.html. Accessed 19 Sep 2017

45. Cech NB, Krone JR, Enke CG (2001) Predicting electrospray response from chromatographic retention time. Anal Chem 73:208–213.

46. Breakthrough Technology for Bottom-Up Proteomics and Small Molecule Research | American Laboratory. http://www.americanlaboratory.com/914-Application-Notes/18815-Breakthrough-Technology-for-Bottom-Up-Proteomics-and-Small-Molecule-Research/. Accessed 19 Sep 2017

47. Michalski A, Damoc E, Hauschild J-P, Lange O, Wieghaus A, Makarov A, Nagaraj N, Cox J, Mann M, Horning S (2011) Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. Mol Cell Proteomics MCP. doi: 10.1074/mcp.M111.011015

48. Pozo ÓJ, Sancho JV, Ibáñez M, Hernández F, Niessen WMA (2006) Confirmation of organic micropollutants detected in environmental samples by liquid chromatography tandem mass spectrometry: Achievements and pitfalls. TrAC Trends Anal Chem 25:1030–1042. doi: 10.1016/j.trac.2006.06.012

49. Krauss M, Singer H, Hollender J (2010) LC–high resolution MS in environmental analysis: from target screening to the identification of unknowns. Anal Bioanal Chem 397:943–951. doi: 10.1007/s00216-010-3608-9

50. Núñez O, Gallart-Ayala H, Martins CPB, Lucci P (2015) Fast Liquid Chromatography–Mass Spectrometry Methods in Food and Environmental Analysis. World Scientific

51. Kaufmann A, Butcher P, Maden K, Walker S, Widmer M (2011) Quantitative and confirmative performance of liquid chromatography coupled to high-resolution mass spectrometry compared to tandem mass spectrometry. Rapid Commun Mass Spectrom 25:979–992. doi: 10.1002/rcm.4952

52. Calderón-Santiago M, López-Bascón MA, Peralbo-Molina Á, Priego-Capote F (2017) MetaboQC: A tool for correcting untargeted metabolomics data with mass spectrometry detection using quality controls. Talanta 174:29–37. doi: 10.1016/j.talanta.2017.05.076

53. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW-M, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reily MD, Thaden JJ, Viant MR (2007) Proposed minimum reporting standards for chemical analysis Chemical

Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics Off J Metabolomic Soc 3:211–221. doi: 10.1007/s11306-007-0082-2

54. Gago-Ferrero P, Schymanski EL, Hollender J, Thomaidis NS (2016) Chapter 13 - Nontarget Analysis of Environmental Samples Based on Liquid Chromatography Coupled to High Resolution Mass Spectrometry (LC-HRMS). In: Pérez S, Eichhorn P, Barceló D (eds) Compr. Anal. Chem. Elsevier, pp 381–403

55. Wolf S, Schmidt S, Müller-Hannemann M, Neumann S (2010) In silico fragmentation for computer assisted identification of metabolite mass spectra. BMC Bioinformatics 11:148. doi: 10.1186/1471-2105-11-148

56. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J (2014) Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. Environ Sci Technol 48:2097–2098. doi: 10.1021/es5002105

57. Creek DJ, Dunn WB, Fiehn O, Griffin JL, Hall RD, Lei Z, Mistrik R, Neumann S, Schymanski EL, Sumner LW, Trengove R, Wolfender JL (2014) Metabolite identification. Metabolomics 10:350–353. doi: 10.1007/s11306-014-0656-8

58. Stein S (2012) Mass Spectral Reference Libraries: An Ever-Expanding Resource for Chemical Identification. Anal Chem 84:7274–7282. doi: 10.1021/ac301205z

59. Kern S, Fenner K, Singer HP, Schwarzenbach RP, Hollender J (2009) Identification of transformation products of organic contaminants in natural waters by computer-aided prediction and high-resolution mass spectrometry. Environ Sci Technol 43:7039–7046.

60. Sugimoto M, Kawakami M, Robert M, Soga T, Tomita M (2012) Bioinformatics Tools for Mass Spectroscopy-Based Metabolomic Data Processing and Analysis. Curr Bioinforma 7:96–108. doi: 10.2174/157489312799304431

61. Castro-Puyana M, Pérez-Míguez R, Montero L, Herrero M (2017) Application of mass spectrometry-based metabolomics approaches for food safety, quality and traceability. TrAC Trends Anal Chem 93:102–118. doi: 10.1016/j.trac.2017.05.004

62. Theodoridis GA, Gika HG, Want EJ, Wilson ID (2012) Liquid chromatography-mass spectrometry based global metabolite profiling: A review. Anal Chim Acta 711:7–16. doi: 10.1016/j.aca.2011.09.042

63. Wolfender J-L, Marti G, Thomas A, Bertrand S (2015) Current approaches and challenges for the metabolite profiling of complex natural extracts. J Chromatogr A 1382:136–164. doi: 10.1016/j.chroma.2014.10.091

64. Souza-Silva ÉA, Gionfriddo E, Pawliszyn J (2015) A critical review of the state of the art of solid-phase microextraction of complex matrices II. Food analysis. TrAC Trends Anal Chem 71:236–248. doi: 10.1016/j.trac.2015.04.018

65. González-Curbelo MÁ, Socas-Rodríguez B, Herrera-Herrera AV, González-Sálamo J, Hernández-Borges J, Rodríguez-Delgado MÁ (2015) Evolution and applications of

the QuEChERS method. TrAC Trends Anal Chem 71:169–185. doi: 10.1016/j.trac.2015.04.012

66. Knolhoff AM, Croley TR (2016) Non-targeted screening approaches for contaminants and adulterants in food using liquid chromatography hyphenated to high resolution mass spectrometry. J Chromatogr A 1428:86–96. doi: 10.1016/j.chroma.2015.08.059

67. Scalbert A, Andres-Lacueva C, Arita M, Kroon P, Manach C, Urpi-Sarda M, Wishart D (2011) Databases on Food Phytochemicals and Their Health-Promoting Effects. J Agric Food Chem 59:4331–4348. doi: 10.1021/jf200591d

68. Andersen M-BS, Rinnan Å, Manach C, Poulsen SK, Pujos-Guillot E, Larsen TM, Astrup A, Dragsted LO (2014) Untargeted metabolomics as a screening tool for estimating compliance to a dietary pattern. J Proteome Res 13:1405–1418. doi: 10.1021/pr400964s

69. Kim S, Kim J, Yun EJ, Kim KH (2016) Food metabolomics: from farm to human. Curr Opin Biotechnol 37:16–23. doi: 10.1016/j.copbio.2015.09.004

70. Fallaize R, Forster H, Macready AL, Walsh MC, Mathers JC, Brennan L, Gibney ER, Gibney MJ, Lovegrove JA (2014) Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. J Med Internet Res 16:e190. doi: 10.2196/jmir.3355

71. Kristal AR, Peters U, Potter JD (2005) Is it time to abandon the food frequency questionnaire? Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol 14:2826–2828. doi: 10.1158/1055-9965.EPI-12-ED1

72. Pujos-Guillot E, Hubert J, Martin J-F, Lyan B, Quintana M, Claude S, Chabanas B, Rothwell JA, Bennetau-Pelissero C, Scalbert A, Comte B, Hercberg S, Morand C, Galan P, Manach C (2013) Mass spectrometry-based metabolomics for the discovery of biomarkers of fruit and vegetable intake: citrus fruit as a case study. J Proteome Res 12:1645–1659. doi: 10.1021/pr300997c

73. Hanhineva K, Lankinen MA, Pedret A, Schwab U, Kolehmainen M, Paananen J, de Mello V, Sola R, Lehtonen M, Poutanen K, Uusitupa M, Mykkänen H (2015) Nontargeted metabolite profiling discriminates diet-specific biomarkers for consumption of whole grains, fatty fish, and bilberries in a randomized controlled trial. J Nutr 145:7–17. doi: 10.3945/jn.114.196840

74. O'Sullivan A, Gibney MJ, Brennan L (2011) Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. Am J Clin Nutr 93:314–321. doi: 10.3945/ajcn.110.000950

75. Lee L-S, Choi JH, Sung MJ, Hur J-Y, Hur HJ, Park J-D, Kim Y-C, Gu E-J, Min B, Kim H-J (2015) Green tea changes serum and liver metabolomic profiles in mice with high-fat diet-induced obesity. Mol Nutr Food Res 59:784–794. doi: 10.1002/mnfr.201400470

76. Guertin KA, Loftfield E, Boca SM, Sampson JN, Moore SC, Xiao Q, Huang W-Y, Xiong X, Freedman ND, Cross AJ, Sinha R (2015) Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer. Am J Clin Nutr 101:1000–1011. doi: 10.3945/ajcn.114.096099

77. Li Z-Y, Ding L-L, Li J-M, Xu B-L, Yang L, Bi K-S, Wang Z-T (2015) ^{1}H-NMR and MS based metabolomics study of the intervention effect of curcumin on hyperlipidemia mice induced by high-fat diet. PloS One 10:e0120950. doi: 10.1371/journal.pone.0120950

78. Kim S, Hwang J, Xuan J, Jung YH, Cha H-S, Kim KH (2014) Global metabolite profiling of synovial fluid for the specific diagnosis of rheumatoid arthritis from other inflammatory arthritis. PloS One 9:e97501. doi: 10.1371/journal.pone.0097501

79. Corona G, Polesel J, Fratino L, Miolo G, Rizzolio F, Crivellari D, Addobbati R, Cervo S, Toffoli G (2014) Metabolomics biomarkers of frailty in elderly breast cancer patients. J Cell Physiol 229:898–902. doi: 10.1002/jcp.24520

80. Doré J, Blottière H (2015) The influence of diet on the gut microbiota and its consequences for health. Curr Opin Biotechnol 32:195–199. doi: 10.1016/j.copbio.2015.01.002

81. Metzler M, Pfeiffer E, Hildebrand A (2010) Zearalenone and its metabolites as endocrine disrupting chemicals. World Mycotoxin J 3:385–401. doi: 10.3920/WMJ2010.1244

82. Flores-Flores ME, Lizarraga E, López de Cerain A, González-Peñas E (2015) Presence of mycotoxins in animal milk: A review. Food Control 53:163–176. doi: 10.1016/j.foodcont.2015.01.020

83. Ríos A, Zougagh M, Bouri M (2013) Magnetic (nano)materials as an useful tool for sample preparation in analytical methods. A review. Anal Methods 5:4558–4573. doi: 10.1039/C3AY40306H

84. Wierucka M, Biziuk M Application of magnetic nanoparticles for magnetic solid-phase extraction in preparing biological, environmental and food sampless. TrAC Trends Anal Chem 59:50–58.

85. Capriotti AL, Cavaliere C, Piovesana S, Stampachiacchiere S, Samperi R, Ventura S, Laganà A (2015) Simultaneous Determination of Naturally Occurring Estrogens and Mycoestrogens in Milk by Ultrahigh-Performance Liquid Chromatography-Tandem Mass Spectrometry Analysis. J Agric Food Chem 63:8940–8946. doi: 10.1021/acs.jafc.5b02815

86. EUR-Lex - 32002D0657 - EN - EUR-Lex. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32002D0657. Accessed 30 Sep 2017

87. Socas-Rodríguez B, Asensio-Ramos M, Hernández-Borges J, Herrera-Herrera AV, Rodríguez-Delgado MÁ (2013) Chromatographic analysis of natural and synthetic

estrogens in milk and dairy products. TrAC Trends Anal Chem 44:58–77. doi: 10.1016/j.trac.2012.10.013

88. Gao Q, Luo D, Bai M, Chen Z-W, Feng Y-Q (2011) Rapid Determination of Estrogens in Milk Samples Based on Magnetite Nanoparticles/Polypyrrole Magnetic Solid-Phase Extraction Coupled with Liquid Chromatography–Tandem Mass Spectrometry. J Agric Food Chem 59:8543–8549. doi: 10.1021/jf201372r

89. Capriotti AL, Cavaliere C, La Barbera G, Piovesana S, Samperi R, Zenezini Chiozzi R, Laganà A (2016) Polydopamine-coated magnetic nanoparticles for isolation and enrichment of estrogenic compounds from surface water samples followed by liquid chromatography-tandem mass spectrometry determination. Anal Bioanal Chem 408:4011–4020. doi: 10.1007/s00216-016-9489-9

90. Cavaliere C, D'Ascenzo G, Foglia P, Pastorini E, Samperi R, Laganà A (2005) Determination of type B trichothecenes and macrocyclic lactone mycotoxins in field contaminated maize. Food Chem 92:559–568. doi: 10.1016/j.foodchem.2004.10.008

91. Hashemi M, Taherimaslak Z, Rashidi S (2014) Application of magnetic solid phase extraction for separation and determination of aflatoxins B1 and B2 in cereal products by high performance liquid chromatography-fluorescence detection. J Chromatogr B 960:200–208. doi: 10.1016/j.jchromb.2014.03.035

92. Mashhadizadeh MH, Amoli-Diva M, Pourghazi K (2013) Magnetic nanoparticles solid phase extraction for determination of ochratoxin A in cereals using high-performance liquid chromatography with fluorescence detection. J Chromatogr A 1320:17–26. doi: 10.1016/j.chroma.2013.10.062

93. Manafi MH, Allahyari M, Pourghazi K, Amoli-Diva M, Taherimaslak Z (2015) Surfactant-enhanced spectrofluorimetric determination of total aflatoxins from wheat samples after magnetic solid-phase extraction using modified Fe3O4 nanoparticles. Spectrochim Acta A Mol Biomol Spectrosc 146:43–49. doi: 10.1016/j.saa.2015.03.050

94. Jouany J-P (2007) Methods of preventing, decontaminating and minimizing the toxicity of mycotoxins in feeds. Anim Feed Sci Technol 137:342–362. doi: 10.1016/j.anifeedsci.2007.06.009

95. Schuhmacher R, Krska R, Weckwerth W, Goodacre R (2013) Metabolomics and metabolite profiling. Anal Bioanal Chem 405:5003–5004. doi: 10.1007/s00216-013-6939-5

96. Zhang R, Watson DG, Wang L, Westrop GD, Coombs GH, Zhang T (2014) Evaluation of mobile phase characteristics on three zwitterionic columns in hydrophilic interaction liquid chromatography mode for liquid chromatography-high resolution mass spectrometry based untargeted metabolite profiling of Leishmania parasites. J Chromatogr A 1362:168–179. doi: 10.1016/j.chroma.2014.08.039

97. Stobiecki M, Kachlicki P, Wojakowska A, Marczak Ł (2015) Application of LC/MS systems to structural characterization of flavonoid glycoconjugates. Phytochem Lett 11:358–367. doi: 10.1016/j.phytol.2014.10.018

98. Mullard G, Allwood JW, Weber R, Brown M, Begley P, Hollywood KA, Jones M, Unwin RD, Bishop PN, Cooper GJS, Dunn WB (2015) A new strategy for MS/MS data acquisition applying multiple data dependent experiments on Orbitrap mass spectrometers in non-targeted metabolomic applications. Metabolomics 11:1068–1080. doi: 10.1007/s11306-014-0763-6

99. Kajdžanoska M, Petreska J, Stefova M (2011) Comparison of Different Extraction Solvent Mixtures for Characterization of Phenolic Compounds in Strawberries. J Agric Food Chem 59:5272–5278. doi: 10.1021/jf2007826

100. Blažević I, Mastelić J (2009) Glucosinolate degradation products and other bound and free volatiles in the leaves and roots of radish (Raphanus sativus L.). Food Chem 113:96–102. doi: 10.1016/j.foodchem.2008.07.029

101. Bianco G, Agerbirk N, Losito I, Cataldi TRI (2014) Acylated glucosinolates with diverse acyl groups investigated by high resolution mass spectrometry and infrared multiphoton dissociation. Phytochemistry 100:92–102. doi: 10.1016/j.phytochem.2014.01.010

102. Bianco G, Pascale R, Lelario F, Bufo SA, Cataldi TRI (2016) The Investigation of Glucosinolates by Mass Spectrometry. In: Mérillon J-M, Ramawat KG (eds) Glucosinolates. Springer International Publishing, pp 1–32

103. Lin L-Z, Sun J, Chen P, Zhang R-W, Fan X-E, Li L-W, Harnly JM (2014) Profiling of Glucosinolates and Flavonoids in Rorippa indica (Linn.) Hiern. (Cruciferae) by UHPLC-PDA-ESI/HRMSn. J Agric Food Chem 62:6118–6129. doi: 10.1021/jf405538d

104. Sun J, Zhang M, Chen P (2016) GLS-Finder: A Platform for Fast Profiling of Glucosinolates in Brassica Vegetables. J Agric Food Chem 64:4407–4415. doi: 10.1021/acs.jafc.6b01277

105. Doheny-Adams T, Redeker K, Kittipol V, Bancroft I, Hartley SE (2017) Development of an efficient glucosinolate extraction method. Plant Methods 13:17. doi: 10.1186/s13007-017-0164-8

106. Benton HP, Ivanisevic J, Mahieu NG, Kurczy ME, Johnson CH, Franco L, Rinehart D, Valentine E, Gowda H, Ubhi BK, Tautenhahn R, Gieschen A, Fields MW, Patti GJ, Siuzdak G (2015) Autonomous Metabolomics for Rapid Metabolite Identification in Global Profiling. Anal Chem 87:884–891. doi: 10.1021/ac5025649

107. Willemse CM, Stander MA, de Villiers A (2013) Hydrophilic interaction chromatographic analysis of anthocyanins. J Chromatogr A 1319:127–140. doi: 10.1016/j.chroma.2013.10.045

108. de Villiers A, Cabooter D, Lynen F, Desmet G, Sandra P (2009) High performance liquid chromatography analysis of wine anthocyanins revisited: Effect of particle size and temperature. J Chromatogr A 1216:3270–3279. doi: 10.1016/j.chroma.2009.02.038

109. Gritti F, Guiochon G (2008) Complete Temperature Profiles in Ultra-High-Pressure Liquid Chromatography Columns. Anal Chem 80:5009–5020. doi: 10.1021/ac800280c

110. Kaczmarski K, Gritti F, Kostka J, Guiochon G (2009) Modeling of thermal processes in high pressure liquid chromatography: II. Thermal heterogeneity at very high pressures. J Chromatogr A 1216:6575–6586. doi: 10.1016/j.chroma.2009.07.049

111. Zrybko CL, Fukuda EK, Rosen RT (1997) Determination of glucosinolates in domestic and wild mustard by high-performance liquid chromatography with confirmation by electrospray mass spectrometry and photodiode-array detection. J Chromatogr A 767:43–52. doi: 10.1016/S0021-9673(96)01068-0

112. Troyer JK, Stephenson KK, Fahey JW (2001) Analysis of glucosinolates from broccoli and other cruciferous vegetables by hydrophilic interaction liquid chromatography. J Chromatogr A 919:299–304. doi: 10.1016/S0021-9673(01)00842-1

113. Wade KL, Garrard IJ, Fahey JW (2007) Improved hydrophilic interaction chromatography method for the identification and quantification of glucosinolates. J Chromatogr A 1154:469–472. doi: 10.1016/j.chroma.2007.04.034

114. Mellon FA, Bennett RN, Holst B, Williamson G (2002) Intact Glucosinolate Analysis in Plant Extracts by Programmed Cone Voltage Electrospray LC/MS: Performance and Comparison with LC/MS/MS Methods. Anal Biochem 306:83–91. doi: 10.1006/abio.2002.5677

115. Sun J, Liu X, Yang T, Slovin J, Chen P (2014) Profiling polyphenols of two diploid strawberry (Fragaria vesca) inbred lines using UHPLC-HRMSn. Food Chem 146:289–298. doi: 10.1016/j.foodchem.2013.08.089

116. Kårlund A, Hanhineva K, Lehtonen M, Karjalainen RO, Sandell M (2015) Nontargeted Metabolite Profiles and Sensory Properties of Strawberry Cultivars Grown both Organically and Conventionally. J Agric Food Chem 63:1010–1019. doi: 10.1021/jf505183j

117. Bialecki JB, Ruzicka J, Weisbecker CS, Haribal M, Attygalle AB (2010) Collision-induced dissociation mass spectra of glucosinolate anions. J Mass Spectrom JMS 45:272–283. doi: 10.1002/jms.1711

118. Bingham S., Gill C, Welch A, Day K, Cassidy A, Al. E (1994) Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. Br J Nutr 72:619–643. doi: 10.1079/BJN19940064

119. Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, Ferrari P, Carroll RJ (2002) Bias in dietary-report instruments and its implications for nutritional epidemiology. Public Health Nutr 5:915–23. doi: 10.1079/PHN2002383

120. Neuhouser ML, Tinker L, Shaw PA, Schoeller D, Bingham SA, Horn LV, Beresford SAA, Caan B, Thomson C, Satterfield S, Kuller L, Heiss G, Smit E, Sarto G, Ockene

J, Stefanick ML, Assaf A, Runswick S, Prentice RL (2008) Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. Am J Epidemiol 167:1247–1259. doi: 10.1093/aje/kwn026

121. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hooft JJJ, Wishart DS (2014) The food metabolome: a window over dietary exposure. Am J Clin Nutr 99:1286–1308. doi: 10.3945/ajcn.113.076133

122. Bouvard V, Loomis D, Guyton KZ, Grosse Y, Ghissassi FE, Benbrahim-Tallaa L, Guha N, Mattock H, Straif K (2015) Carcinogenicity of consumption of red and processed meat. Lancet Oncol 16:1599–1600. doi: 10.1016/S1470-2045(15)00444-1

123. Abete I, Romaguera D, Vieira AR, Lopez de Munain A, Norat T (2014) Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. Br J Nutr 112:762–775. doi: 10.1017/S000711451400124X

124. Aune D, Ursin G, Veierød MB (2009) Meat consumption and the risk of type 2 diabetes: a systematic review and meta-analysis of cohort studies. Diabetologia 52:2277–2287. doi: 10.1007/s00125-009-1481-x

125. Klurfeld DM (2015) Research gaps in evaluating the relationship of meat and health. Meat Sci 109:86–95. doi: 10.1016/j.meatsci.2015.05.022

126. Dragsted LO (2010) Biomarkers of meat intake and the application of nutrigenomics. Meat Sci 84:301–307. doi: 10.1016/j.meatsci.2009.08.028

127. T B, J H-J, K H, Lo D (2012) Metabolic fingerprinting of high-fat plasma samples processed by centrifugation- and filtration-based protein precipitation delineates significant differences in metabolite information coverage. Anal Chim Acta 718:47–57. doi: 10.1016/j.aca.2011.12.065

128. Bijlsma S, Bobeldijk I, Verheij ER, Ramaker R, Kochhar S, Macdonald IA, van Ommen B, Smilde AK (2006) Large-Scale Human Metabolomics Studies: A Strategy for Data (Pre-) Processing and Validation. Anal Chem 78:567–574. doi: 10.1021/ac051495j

129. Castillo S, Gopalacharyulu P, Yetukuri L, Orešič M (2011) Algorithms and tools for the preprocessing of LC–MS metabolomics data. Chemom Intell Lab Syst 1:23–32. doi: 10.1016/j.chemolab.2011.03.010

130. Li B, Tang J, Yang Q, Cui X, Li S, Chen S, Cao Q, Xue W, Chen N, Zhu F (2016) Performance Evaluation and Online Realization of Data-driven Normalization Methods Used in LC/MS based Untargeted Metabolomics Analysis. Sci Rep 6:38881. doi: 10.1038/srep38881

131. Westerhuis JA, Velzen EJJ van, Hoefsloot HCJ, Smilde AK (2010) Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. Metabolomics 6:119–128. doi: 10.1007/s11306-009-0185-z

132. Münger LH, Trimigno A, Picone G, Freiburghaus C, Pimentel G, Burton KJ, Pralong FP, Vionnet N, Capozzi F, Badertscher R, Vergères G (2017) Identification of Urinary Food Intake Biomarkers for Milk, Cheese, and Soy-Based Drink by Untargeted GC-MS and NMR in Healthy Humans. J Proteome Res 16:3321–3335. doi: 10.1021/acs.jproteome.7b00319

133. Human Metabolome Database. http://www.hmdb.ca/. Accessed 30 Sep 2017

134. Cuparencu CS, Andersen M-BS, Gürdeniz G, Schou SS, Mortensen MW, Raben A, Astrup A, Dragsted LO (2016) Identification of urinary biomarkers after consumption of sea buckthorn and strawberry, by untargeted LC–MS metabolomics: a meal study in adult men. Metabolomics 12:31. doi: 10.1007/s11306-015-0934-0

Appendix

Liquid chromatography-high resolution mass spectrometry for the analysis of phytochemicals in vegetal-derived food and beverages

Food Research International 100 (2017) 28–52

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Review

Liquid chromatography-high resolution mass spectrometry for the analysis of phytochemicals in vegetal-derived food and beverages

Giorgia La Barbera, Anna Laura Capriotti, Chiara Cavaliere, Carmela Maria Montone, Susy Piovesana*, Roberto Samperi, Riccardo Zenezini Chiozzi, Aldo Laganà

Dipartimento di Chimica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy

ARTICLE INFO

Keywords:
Phytochemicals
Food analysis
High resolution mass spectrometry
High resolution liquid chromatography
Phenolics
Polyphenols
Metabolomics

ABSTRACT

The recent years witnessed a change in the perception of nutrition. Diet does not only provide nutrients to meet the metabolic requirements of the body, but it also constitutes an active way for the consumption of compounds beneficial for human health. Fruit and vegetables are an excellent source of such compounds, thus the growing interest in characterizing phytochemical sources, structures and activities. Given the interest for phytochemicals in food, the development of advanced and suitable analytical techniques for their identification is fundamental for the advancement of food research. In this review, the state of the art of phytochemical research in food plants is described, starting from sample preparation, throughout extract clean-up and compound separation techniques, to the final analysis, considering both qualitative and quantitative investigations. In this regard, from an analytical point of view, fruit and vegetable extracts are complex matrices, which greatly benefit from the use of modern hyphenated techniques, in particular from the combination of high performance liquid chromatography separation and high resolution mass spectrometry, powerful tools which are being increasingly used in the recent years. Therefore, selected applications to real samples are presented and discussed, in particular for the analysis of phenols, polyphenols and phenolic acids. Finally, some hot points are discussed, such as waste characterization for high value-compounds recovery and the untargeted metabolomics approach.

1. Introduction

The increasing knowledge about the impact of diet on health is changing the opinion on the role of nutrition. Diet does not only provide adequate nutrients to meet the metabolic requirements of the body, but it can be a source of bioactive compounds able to contribute to the improvement of human health. Consequently, the identification of such compounds is crucial for both the food and the dietary integrators market, to provide customers with a healthy, balanced diet. Bioactive compounds are being studied in the prevention of cancer, heart disease, and other diseases (https://www.cancer.gov/publications/dictionaries/cancer-terms?cdrid=703278). In this context, the vegetable kingdom is a primary source of such phytochemicals with potential health benefits. At present, researchers have identified hundreds of compounds with health-promoting, disease-preventing, or curative properties, and new discoveries concerning the complex interactions between secondary metabolites in plants and health are continually made (Weidner et al., 2012). Table 1, adapted from (Sbrana, Avio, & Giovannetti, 2014), shows some of the phytochemicals occurring in edible plants with health-promoting, disease-preventing, or medicinal properties. A considerable amount of data on dietary phytochemicals accumulated over the past years has been collected and stored in electronic databases (Scalbert et al., 2011). Additionally, the advent of metabolomics, which is a powerful approach enabling the comprehensive, qualitative and quantitative analysis of all metabolites (Fiehn, 2002), prompted research in this field; high resolution mass spectrometry (HRMS) coupled to (ultra)high performance liquid chromatography ((U)HPLC) is increasingly used in metabolomics. The recent innovations in instrumentation and informatics currently allow performing a comprehensive analysis of metabolites and reliably compare samples in a semi-automated manner. However, rigorous methodologies are required. Plants have a very large metabolome: about 200,000 phytochemicals are known so far and 20,000 of them have been identified as originating from fruits, vegetables and grains; moreover, they differ in molecular weight, physico-chemical characteristics and dynamic range (Oz & Kafkas, 2017). Thus, despite the potential, a fully exhaustive characterization in plant metabolomics remains challenging.

Metabolomics is subdivided into targeted (Roberts, Souza, Gerszten, & Clish, 2012) and untargeted (Vinayavekhin & Saghatelian,

* Corresponding author at: Dipartimento di Chimica, Sapienza Università di Roma, Box no 34 - Roma 62, Piazzale Aldo Moro 5, 00185 Rome, Italy.
E-mail address: susy.piovesana@uniroma1.it (S. Piovesana).

http://dx.doi.org/10.1016/j.foodres.2017.07.080
Received 19 May 2017; Received in revised form 30 July 2017; Accepted 31 July 2017
Available online 02 August 2017
0963-9969/

Table 1
Health effect of some phytochemicals (adapted from (Sbrana et al., 2014) with the addition of new items and references).

Phytochemicals	Plant	Health promoting activities	References
Activity on metabolic diseases			
Curcumin, capsaicin, catechins	Curcuma, chili peppers, cocoa	Antidiabetic	(Leiherer, Mündlein, & Drexel, 2013)
Amorfrutins	*Glycyrrhiza foetida* root	Antidiabetic, lipid-lowering	(Weidner et al., 2012)
Proanthocyanidins	Seed shells	Antiobesity	(Yokota, Kimura, Ogawa, & Akihiro, 2013)
Luteolin, phytosterols, isoflavones	Artichoke, soy	Hypocholesterolemic activity	(Gebhardt, 1997; Mannarino, Ministrini, & Pirro, 2014)
Anti-carcinogenic activity			
Theaflavins, thearubigins	Black teas	Antiproliferative activity	(Bhattacharya, Halder, Mukhopadhyay, & Giri, 2009)
Baicalein, baicalin, curcumin, quercetin, lutein	Garlic, lemon, carrot, basil, grape	Antiproliferative action on leukemia, hepatoma, melanoma, breast, bladder, pancreas, ovary, brain, kidney, lung, colon, and stomach carcinoma cells	(Cherng, Shieh, Chiang, Chang, & Chiang, 2007; Rai, Kaur, Jacobs, & Singh, 2010; Russo, Nigro, Rosiello, D'Arienzo, & Russo, 2007; Shehzad, Lee, & Lee, 2013)
Erythrodiol, uvaol, oleanolic, maslinic acid	Olive fruits	Cytotoxic effects on human breast cancer cells	(Allouche et al., 2011)
Isothiocyanates, dithiolethiones, sulforaphanes	Brassicaceae	Inhibition of carcinogenesis by modulating carcinogens metabolism and detoxification	(Ares, Nozal, & Bernal, 2013; Dillard & Bruce German, 2000; Tang et al., 2010)
Gamma-carotene, lycopene, lutein	Cruciferous vegetables	Protection against uterine, prostate, breast, colorectal, lung, and digestive tract cancers	(Jian, Du, Lee, & Binns, 2005; Toniolo et al., 2001)
Other activities			
S-allyl cysteine, allicin	Garlic	Antihypertensive	(Shouk, Abdou, Shetty, Sarkar, & Eid, 2014)
Luteolin, kaempferol, apigenin, myricetin	Teas, onions, apples	Anti-inflammatory and antibacterial activities	(Dillard & Bruce German, 2000)
Beta-sitosterol and its glycosides	Legumes, cereals, corn oil, nuts	Antioxidant activity, hypocholesterolemic activity, anti-inflammatory, antineoplastic, antipyretic, and immune system-modulating activity	(Loizou, Lekakis, Chrousos, & Moutsatsou, 2010; Woyengo, Ramprasath, & Jones, 2009)
Tocotrienols, tocopherols	Palm oil	Antioxidant, antiproliferative and apoptotic activities, preventing or reducing the risk of breast cancer	(McIntyre, Briski, Gapor, & Sylvester, 2000)
Phenolics in general, especially polyphenolics	Figs and grape seeds	Antioxidant, amyloid disease prevention and therapy, apoptotic on cancer cells	(Antunes-Ricardo et al., 2014; Ares et al., 2013; Lourenço, Gago, Barbosa, De Freitas, & Laranjinha, 2008; Stefani & Rigacci, 2014)
Epicatechin, epigallocatechin, epicatechin-gallate, epigallocatechin-gallate	Grapes, berries, cocoa, green tea, nuts	Reduction of dysregulations and degenerative phenomena, anticarcinogenic effects	(Lambert & Elias, 2010; Yang, Lambert, & Sang, 2009)
Carotenoids	Carrots, tomatoes	Antioxidants that might aid in the prevention of several human chronic degenerative diseases, such as cancer, cardiovascular diseases and age-related eye diseases	(Bijttebier, D'Hondt et al., 2014)

2010) approaches. In the targeted approach a set of selected, known metabolites are identified and quantified by using authentic standards or available databases (in the latter case a specific list of compounds is screened but no reference compounds are available for confirmation). The resulting data can be used as input for statistical analysis to distinguish sample classes with accuracy. In contrast, untargeted metabolomics aims to obtain information about all detectable compounds (no hypothesis is made on which compounds are present in the sample and the aim is to analyze everything, to achieve the widest metabolic coverage). Although this aim has not been reached yet, clear progress in this direction is ongoing. Suspect screening is an intermediate approach, which relies on specific information on some compounds, such as molecular formula and structure, suspected to be present in the sample (Flamini et al., 2013). This approach is relatively much more employed in the environmental analysis (Krauss, Singer, & Hollender, 2010; Scheibner, Squibb, Greco, & Steiner, 2016) than in phytochemical research.

Sometimes only metabolites related to either specific metabolic pathways or a class of compounds are investigated (Theodoridis, Gika, Want, & Wilson, 2012; Wolfender, Marti, Thomas, & Bertrand, 2015).

Until recently, most of the plant metabolite studies relied on analysis/quantitation of selected bioactive compounds. The development of powerful analytical techniques is essential for the advancement in phytochemical research. Due to the complexity of natural matrices, hyphenated HPLC-HRMS techniques appear a powerful tool and have been increasingly used in the recent years. Despite the very large number of reviews dealing with phytochemical analysis by chromatography-MS (Ignat, Volf, & Popa, 2011; Khoddami, Wilkes, & Roberts, 2013; Valls, Millán, Martí, Borràs, & Arola, 2009; Wang, Zhang, Yan, Han, & Sun, 2014; Wu et al., 2013), none of them was devoted to phytochemicals in food, neither were they focused on methodological aspects peculiar to HRMS, nor on issues related to the chemical classes. Considering all these aspects, the purpose of this review is to present and discuss the modern analytical approaches, based on HPLC and UHPLC coupled with HRMS and tandem mass spectrometry (MS/MS), employed to characterize and/or identify phytochemicals in food of plant origin. Decoctions, infusions and medicinal herbs in general will not be included in this review. The advantages, as well as the limitations of the various methods, are reported in order to provide an overview of the most suitable approaches for class-specific or multi-class analysis.

2. Sample preparation

2.1. General remarks

Sample preparation is a critical step, which determines the compounds isolated and the accuracy of both qualitative and quantitative results; therefore, all the aspects of the procedures should be considered from the beginning. In principle, sample preparation methods for HRMS do not differ significantly from traditional methods for low resolution MS analysis (Kim & Verpoorte, 2010). However, in HRMS experiments large sets of analytes are usually investigated, then the extraction should be a general one able to extract them all. Due to the presence of a wide variety of phytochemicals in plants, with very different polarities, no single solvent or solvent mixture is suitable for this purpose. Several extractions with solvents having different polarity (at least a polar one and a non-polar one) are needed to cover the entire range of

Table 2
Schematic summary of the most important papers analyzing phytochemicals in food plant used for the review.

Matrix	Extraction	Chromatography	MS	Identification/validation/quantitation	Compounds	Ref.
Apple peel (*Malus* × *domestica Barkh cv. Kanzi®*)	Lyophilized, ground. Extraction: MeOH/water 20:80, + MeOH, ultrasounds. No clean-up	UPLC: BEH Shield RP18, 3.0 × 150 mm, 1.7 μm Water/ACN + 0.1% formic acid. 23 min	ESI-Orbitrap Exactive DAD	Targeted quantitative. Validation: LOQs 200–900 ng g^{-1}; recoveries 74–100%	39 phenolics	(De Paepe et al., 2013)
Pulp and peel 9 plantain cultivar (*Musa* sp.)	Lyophilized, ground. Extraction: acetone/water/CH_3COOH 50:49:1; under agitation at 40 °C × 3 times. Clean-up by C18 SPE	UHPLC: XSelect CSH RP C18, 3 × 100 mm; 2.5 μm Water/ACN + 0.1% formic acid. 50 min	ESI-LTQ-Orbitrap-XL MS/MS DAD	Untargeted. No validation. Quantification: DAD	18 phenolics, mainly glycosylated	(Passo Tsamo et al., 2015)
12 commercial sour cherry cultivars (*Prunus cerasus* L.)	Lyophilized, pulverized Extraction: MeOH/water/formic acid 60:39:1 + ultrasound. No clean-up	UHPLC: Kinetex C18, 4.6 × 150 mm, 2.6 μm. Water/ACN + 0.5% formic acid; ~50 min	ESI + Q-TOF MS/MS DAD QqQ	Untargeted, no validation. Quantification: DAD	4 genistein glycosides	(Abrankó et al., 2015)
Bottle gourd fruit (*Lagenaria siceraria stand.*)	Lyophilized Extraction: MeOH/water 70:30 homogenized + ultrasound. No clean-up	HPLC: RP diphenyl 3 × 250 mm, 5 μm. Water + 0.5% formic acid/MeOH. 100 min	ESI-Q-TOF MS DAD, ionTRAP	Untargeted, identification + confirmation. No validation.	22 phenolic glycosides	(Jaiswal & Kuhnert, 2014)
13 varieties of avocado (*Persea americana*)	Lyophilized, homogenized. Extraction: MeOH and vortexing. No clean-up	UHPLC: Acclaim RSLC 120 C18, 2.1 × 100 mm 1.8 μm. Water/ACN + 0.1% formic acid. ~45 min	ESI +/− TOF MS UV	Untargeted + targeted (32 comp.). Identification. Validation: LOQs 9–1700 ng g^{-1}; recoveries 97–102%	59 metabolites (about 35 phenolics)	(Hurtado-Fernández et al., 2011)
Peel and flesh of fruits from four different mango varieties (*Mangifera indica* L.)	Lyophilized, pulverized. Extraction: Soxhlet, 1° fraction exane: discard. 2° fraction MeOH, cleaned up by LLE. 5 fraction by preparative RP HPLC	HPLC: Gemini NX C18, 2.0 × 150 mm, 3 μm Water/ACN + 0.1% formic acid. ~40 min	ESI +/− Q-TOF MS/MS UV	Untargeted. Identification and confirmation. Identification, confirmation. No validation.	77 compounds	(Pierson et al., 2014)
Red chili peppers (*Capsicum frutescens* L.)	Lyophilized, milled and homogenized. Extraction: acetone/MeOH 70:30 ASE. Clean-up by LLE	UPLC: Acquity HSS C18 SB 2.1 × 100 mm, 1.8 μm. 5 mM NH_4 Ac./MeOH/ACN/EtAc.// ACN/EtAc. ~25 min	APCI + Exactive Orbitrap MS DAD	Untargeted. Identification and semiquantification of some compounds. No fully validation.	123 non-polar compounds	(Bijttebier, Zhani et al., 2014)
Broccoli (*Brassica oleracea* L. *var. italica*), cauliflower (*B. oleracea* L. *var. Botrytis*) and rocket salad (*Eruca sativa* L.) (*Brassicaceae*)	Lyophilized, finely ground. Extraction: MeOH/water 70:30 × 2, 70–80 °C + ultrasound. No clean-up	HPLC: Discovery C18 4.6 × 250 mm, 5 μm 0.1% formic acid/ACN. 20 min.	ESI-LTQ-FTICR MS/MS	Untargeted. Identification. Validation: linearity and LOD for 5 standards.	24 glucosinolates	(Lelario et al., 2012)
Broccoli leaf samples (*Parthenon*, *Nubia*, and *Naxos*)	Lyophilized. Extraction: ASE EtOH/water. No clean-up	UHPLC: BEH Shield C18; 2.1 × 1000 mm, 1.7 μm; water/ACN + 05% formic acid. 15 min.	ESI-Q-TOF MS/MS	Targeted; validation: linearity, recovery, MDL, MQL recovery, ME for 7 standards.	15 glucosinolates	(Ares et al., 2015)
Cranberries (*Vaccinium macrocarpon*), blueberries (*Vaccinium myrtillus*), grapes and raisins (*Vitis vinifera*)	Lyophilized. Method 1: MeOH:H_2O (10:90 v/v) 0.2% HCl, + ultrasound, centrifugation (C18 SPE for extraction method 1.2). Method 2: MeOH/H_2O (85:15 v/v), + ultrasound, centrifugation Method 3: acetone/H_2O/HCl 70:29.9:0.1 v/v/v, + ultrasound, centrifugation. Clean-up by C18 SPE	UHPLC: Kinetex C18 4.6 × 100 mm, 2.6 μm; water + 0.1% formic acid/MeOH; 36 min. CZE was also employed	ESI-LTQ-Orbitrap Velos MS/MS	Untargeted. Identification and confirmation.	Phenolics including A, B, C-type procyanidin and proanthocyanidin trimers and tetramers	(Navarro et al., 2014)

(continued on next page)

Table 2 *(continued)*

Matrix	Extraction	Chromatography	MS	Identification/validation/quantitation	Compounds	Ref.
Three artichoke varieties (*Cynara scolymus*)	Lyophilized, ground Extraction: MeOH/water 60:40 + ultrasound. No clean-up	UPLC: HSS T3 1.0 × 100 mm, 1.8 μm; Water/ACN + 0.1% formic acid. ~20 min	ESI- Q-TOF MS/MS ION TRAP DAD	Untargeted metabolomic approach. Identification.	40 compounds	(Farag et al., 2013)
Six Chilean berries (*L. apiculate, L. chequén, U. molinae, A. meli, V. corymbosum, B. microphylla*)	Lyophilized, pulverized. Defatting. Extraction: acidified MeOH, ultrasound. Clean-up: XAD7	HPLC: Luna C18 4.6 × 250 mm, 5 μm Water/ACN + 0.1% formic acid. 60 min.	ESI +/− TOF MS DAD	Untargeted. Comparison between 6 berries species. Identification	31 anthocyanins + some minor phenolics	(Ramirez et al., 2015)
Sim fruit (*Rhodomyrtus tomentosa*)	Lyophilized. Extraction: acetone: water: acetic acid 50:49:1, × 3 under shaking 1 h, 37 °C. Clean-up by C18 SPE, fractions	HPLC: Gemini C18, 4.6 × 150 mm, 3 μm; water 5% formic acid/ACN. 60 min	ESI- LTQ-Orbitrap-XL MS/MS DAD	Untargeted. Identification and confirmation. Validation.	19 phenolic compounds including stilbenes and ellagitannins	(Lai et al., 2013)
Potato (*Solanum tuberosum* L.)	Lyophilized. Extraction: MeOH/water 80:20, 1% acetic acid + ultrasound (× 4). No clean-up	UPLC: Poroshell 120 EC-C18 4.6 × 100 mm, 2.7 μm. Water + 0.5% acetic acid/ACN; 26 min.	ESI +/− Q-TOF MS/MS DAD	Untargeted. Comparison of flesh, peel, and whole tuber of two potato cultivars. Tentative identification. Quantitation by UV–vis data	24 phenolic and other polar compounds	(López-Cobo et al., 2014)
Peppers (*Capsicum annuum* L.)	Lyophilized. Extraction: MeOH/water (80:20, v/v) + ultrasounds. No clean-up	UPLC: Zorbax Eclipse plus C18 4.6 × 150 mm, 1.8 μm. Water + 0.5% acetic acid/ACN; 70 min.	ESI- TOF MS DAD	Untargeted. Three different pepper varieties. Tentative identification.	45 phenolic and other polar compounds	(Morales-Soto et al., 2013)
Blueberry (*Vaccinium cyanococcus*), and red radish (*Raphanus sativus* var. *longipinnatus* '*Shinrimei*')	Lyophilized, ground Extraction: MeOH/water/formic acid (60:40:1, v/v/v) + ultrasound, 60 min at RT. The extract was dried and alkaline hydrolyzed. Clean-up by OASIS HLB	UHPLC: Hypersil Gold C18 2.1 × 200 mm, 1.9 μm; Water + 0.1% formic acid/ACN + 0.1% formic acid; 65 min.	ESI- LTQ Orbitrap XL MS/MS	Untargeted. Identification. No quantification.	Anthocyanins and non-anthocyanin phenolic compounds	(Sun et al., 2012)
Brassica microgreens	Lyophilized, ground Extraction: MeOH/water (60:40, v/v) + ultrasound, 60 min at RT. No clean-up	UHPLC: Hypersil Gold AQ RP-C18 2.1 × 200 mm, 1.9 μm; Water + 0.1% formic acid/ACN + 0.1% formic acid; 65 min.	ESI +/− LTQ Orbitrap XL MS/MS DAD	Untargeted. Tentative identification. No quantification.	64 polyphenols (30 anthocyanins, 105 flavonol glycosides, and 29 hydroxycinnamic acid and hydroxybenzoic acid derivatives)	(Sun et al., 2013)
Red mustard greens (*Brassica juncea Coss* variety)	Lyophilized, ground Extraction: MeOH/water/(60:40, v/v) + ultrasound, 60 min at RT. The extract was dried and alkaline hydrolyzed. Clean-up by OASIS HLB	UHPLC: Hypersil Gold AQ RP-C18 2.1 × 200 mm, 1.9 μm; Water + 0.1% formic acid/ACN + 0.1% formic acid; 65 min.	ESI +/− LTQ Orbitrap XL MS/MS DAD	Untargeted. Identification. No quantification.	67 anthocyanins, 102 flavonol glycosides, and 40 hydroxycinnamic acid derivatives	(Lin et al., 2011)
Two diploid strawberries (*Fragaria vesca*)	Lyophilized, ground. Extraction: MeOH/water/formic acid (60:40:1, v/v/v) Homogenization + ultrasound. No clean-up	UHPLC: Hypersil Gold AQ RP-C18 2.1 × 200 mm, 1.9 μm; Water + 0.1% formic acid/ACN + 0.1% formic acid; 161 min.	ESI +/− LTQ Orbitrap XL MS/MS DAD	Untargeted. Identification. No quantification.	About 58 phenolic compounds	(Sun et al., 2014)
12 purple-fleshed sweet potato (Jihei 1, Xuzi 3, Xuzi 6 and Zhezi 4)	Lyophilized Extraction: 1% formic acid in ethanol + incubation for 2 h × 3 times. No clean-up	UPLC: BEH C18 2.1 × 100 mm, 1.7 μm. Water 2% formic acid/ACN. 40 min.	ESI + Q-TOF MS/MS DAD	Quantification.	Thirteen acylated anthocyanins	(He et al., 2016)
Onion	Lyophilized Extraction: MeOH 80% mixer mill 10 min 4 °C. No clean-up	HPLC: Xselect CSH C18, 2.1 × 100 mm, 3.5 μm water/ACN + 0.1% formic acid 60 min	ESI + FTICR MS and MS/MS	Untargeted metabolomics. Identification.	22 (identified) 45 (suspected) sulfur containing compounds	(Nakabayashi et al., 2013)

(continued on next page)

Table 2 *(continued)*

Matrix	Extraction	Chromatography	MS	Identification/validation/quantitation	Compounds	Ref.
Queule fruits	Lyophilized Extraction: MeOH (× 3) in the dark 1 h 25 °C. Clean-up by XAD7	HPLC: Luna C18 4.6 × 250 mm, 5 μm; water/ACN + 0.1% formic acid; 80 min	ESI +/− TOF MS DAD	Untargeted metabolomics Identification. Quantitation.	32 (detected) 26 (identified) 8 (isolated and quantified) compounds.	(Simirgiotis et al., 2013)
Two amaranth *cv* (*Amaraanthus hybridus* and *Amaranthus hypochondriacus*)	Lyophilized and ground Extraction: MeOH/water (60:40, v/v) under shaking 30 min. 20 °C. No clean-up	HPLC: Ascentis express F5 4.6 × 150 mm, 5 μm; water 0.05% formic acid/ACN; 33 min	ESI +/− TOF MS Ion trap MS/MS	Untargeted. Identification. Confirmed by MS/MS after preparative LC	Saponins	(Zehring et al., 2015)
Roots of red beets (*Beta vulgaris* L.)	Lyophilized and milled Extraction: MeOH 80% × 3 ultrasound. After defatting. Double clean-up by C18 SPE	UPLC: BEH C18, 2.1 × 100 mm, 1.7 μm; water 0.1% formic acid/ACN 13 min.	ESI- Q-Exactive Orbitrap Ion trap MS/MS	Untargeted. Tentative identification.	44 (detected) 27 (tentatively identified) saponins	(Mikołajczyk-Bator et al., 2016)
Carrot steam peels	Lyophilized and added with 1 g of $NaHCO_3$ and water. £0 min rest under N2. Extraction: ASE with 70:30 acetone:MeOH + 0.1% butylated hydroxytoluene at 40 °C (3 ×); LLE with hexane	Six different UHPLC stationary phases (ACQUITY UPLC BEH C18, ACQUITY UPLC HSS T3, ACQUITY UPLC HSS C18 SB, ACQUITY UPLC CSH C18, ACQUITY UPLC BEH Shield RP C18, ACQUITY UPLC BEH Phenyl); one HPLC YMC C30	APCI-Q-Exactive Orbitrap	Targeted (for column comparison)	15 carotenoids (5 carotenes, 10 xanthophylls)	(Bijttebier, D'Hondt et al., 2014)
Cucumber (*Cucumis sativus* L.)	Lyophilized and ground Extraction: MeOH/water (80:20, v/v) ultrasound 30 min, 22 °C. No clean-up	UHPLC: Zorbax C18 4.6 × 150 mm, 1.8 μm; water 0.5% acetic acid/ACN 30 min	ESI- Q-TOF MS/MS	Untargeted metabolomics Tentative identification.	73 phenolic compounds	(Abu-Reidah et al., 2012)
Lettuce (*Lactuca sativa*)	Lyophilized and ground Extraction: MeOH/water (80:20, v/v) ultrasound 30 min, 22 °C. No clean-up	UHPLC: Zorbax C18 4.6 × 150 mm, 1.8 μm; water 0.5% acetic acid/ACN 30 min	ESI- Q-TOF MS/MS	Untargeted metabolomics. Tentative identification.	159 compounds belonging to different classes	(Abu-Reidah et al., 2013)
Green asparagus (*Asparagus officinalis*)	Lyophilized and ground Extraction: MeOH/water (80:20, v/v) ultrasound 20 min, 37 °C. No clean-up	UHPLC: Zorbax SB C18, 2.1 × 100 mm, 1.8 μm; water 0.5% acetic acid/ACN 60 min	ESI- Q-TOF MS/MS	Untargeted metabolomics Tentative identification.	94 polar and semi-polar compounds belonging to different classes	(Jiménez-Sánchez et al., 2016)
Potato (*Solanum tuberosum cv* 'Urenika')	Lyophilized and ground Extraction: 80% ethanol under stirring 2 h, RT. No clean-up	UHPLC: Zorbax C18 4.6 × 150 mm, 1.8 μm; water 0.5% formic acid/MeOH 14 min	ESI +/− Q-TOF MS/MS	Targeted/untargeted Identification. Quantitation with 25 standard	31 compounds belonging to different classes	(Chong et al., 2013)
Faba bean (*Vicia faba* L.)	Lyophilized and ground Extraction: MeOH/water (80:20, v/v) ultrasound 30 min, 22 °C. No clean-up	UHPLC: Zorbax C18 4.6 × 150 mm, 1.8 μm; water 0.5% acetic acid/ACN 30 min	ESI- Q-TOF MS/MS	Untargeted metabolomics. Tentative identification.	179 compounds; 104 phenolics	(Abu-Reidah et al., 2014)
Skin of 34 hybrid grape varieties (*V. vinifera, Vitis riparia, Vitis labrusca, Vitis lincecumii* and *Vitis rupestris*)	Homogenized. Extraction: MeOH/water/formic acid (50/48.5/1.5, v/v/v) anthocyanins removed by Oasis MCX SPE	UPLC: RRHD SB-C18 3.0 × 150 mm, 1.8 μm water/ACN + 0.1% formic acid; ~15 min	ESI- Q-TOF MS/MS QqQ	Untargeted, identification and confirmation. No validation	24 flavonol glycosides	(De Rosso et al., 2014)
Skin, pulp and seeds of Albariño grape (*Vitis vinifera* L.)	Frozen, separation of pulp, skin, and seed. Homogenized. Extraction: EtOH/water (80/20, v/v) pH 3.5 × 2. No clean-up	HPLC: Nucleosil 120 C18 4.0 × 250 mm, 5 μm water with 0.8% acetic acid/20:80 aqueous:ACN 45 min.	ESI- Q-TOF MS/MS QqQ	Untargeted. Identification, confirmation. Quantification DAD No validation.	43 phenolics	(Di Lecce et al., 2014)
Tomato (*Lycopersicon esculentum Mill.*)	Homogenized. Extraction: EtOH/water + 0.1% formic acid (80:20, v/v) + ultrasound × 2 times. Clean-up by Oasis MAX SPE	HPLC: Luna C18 2.0 × 50 mm, 5 μm. Water/ACN + 0.1% formic acid. 20 min.	ESI- LTQ-Orbitrap Velos MSMS QqQ	Untargeted. Identification and confirmation. No validation	38 phenolics	(Vallverdú-Queralt et al., 2010)

(continued on next page)

Table 2 *(continued)*

Matrix	Extraction	Chromatography	MS	Identification/validation/quantitation	Compounds	Ref.
Tomato (*Lycopersicon esculentum Mill.*)	Frozen. Homogenized. Extraction: EtOH/water (80/20, v/v) × 2 times. Clean-up by Oasis MAX SPE	HPLC: Luna C18 2.0 × 50 mm, 5 μm. Water/ACN + 0.1% formic acid. 20 min.	ESI QTOF MS/MS QqQ	Untargeted, identification. Targeted, 10 comp. quantified. Validation: LOQs 0.4–2 μg g^{-1}; recoveries 78–98%	34 phenolics	(Vallverdú-Queralt et al., 2012)
Tangerine tomato (*Solanum lycopersicum* L.), and red tomato (*Solanum lycopersicum* L.) juice	Polar fraction: juice added with MeOH (sonication); lipophilic fraction: LLE of pellet from previous extraction with hexane:acetone (1/1, v/v, 3 ×); LLE on pooled supernatants with water.	HPLC: C30 250 × 4.6 mm, 3 5 μm. MeOH/2-methyl-2-methoxypropane + 2% ammonium acetate 21 min.	APCI (lipophilic fraction) or ESI (polar fraction) Q-TOF MS/MS	Untargeted phytochemical and metabolite analysis	423 compounds detected in the lipophilic fraction 474 compounds in the polar fraction	(Cichon et al., 2017)
Two red grape varieties (*V. vinifera*)	Frozen and homogenized. Extracted immediately, MeOH under stirring. No clean-up	UHPLC: Zorbax RRHD SB-C18 3.0 × 150 mm, 1.8 μm. Water/ACN + 0.1% formic acid. ~40 min	ESI +- Q-TOF MS/MS	Targeted	18 stilbene derivatives	(Flamini et al., 2013)
Blueberry (*Vaccinium angustifolium* Aiton), cranberry (*Vaccinium macrocarpon* Aiton), chestnut (*Aesculus turbinata* Blume)	(Frozen)-homogenized. Extraction: acetone/water (70:30, v/v) Clean-up: SEC + RP chromatography. Thiolytic cleavage of polymers	UPLC: Cosmosil 2.5C18-MS-II 2.0 × 75 mm, 2.5 μm. Water/ACN + 0.1% formic acid. 10 min.	ESI- Q-TOF MS/MS DAD	Untargeted. Structural information.	Highly polymeric proanthocyanidins.	(Kimura et al., 2011)
Olive (*Olea europaea* L., Carolea, Coratina, Nocellara del Belice and Leccino *cv*)	Sample crushed + ultrasound with MeOH. Defatting with hexane.	Semipreparative fractionation: Luna C18 4.6 × 250 mm; 5 mM NH4 CH3COO/MEOH	ESI + Q-TOF Infusion of fractions	Untargeted.	Glycosylated polyphenolics and secoiridoids	(Di Donna et al., 2007)
Strawberry (*Fragaria* × *ananassa*)	Homogenized. Extraction: MeOH, ultrasound for 30 min. Ascorbic acid 1% × 2 Fractions mixed and dried. No clean-up	HPLC: Luna C18, 2.0 × 150 mm, 3 μm; water/MeOH, 0.1% formic acid, ~90 min.	Q Exact. Orbitrap MS/MS QqQ DAD	Untargeted. Identification.	66 nonanthocyanin phenolic compounds and citric acid	(Álvarez-Fernández et al., 2015)
Strawberry (*Fragaria* × *ananassa*)	Frozen and pulverized Extraction: ultrapure water + ultrasound. No clean-up	HPLC: Kromasil C18 column, 4.6 × 150 mm, 5 μm; water/ACN 0.1% formic acid, 45 min.	Q-TOF-MS/MS QqQ	Targeted. Identification. Quantitation with QqQ	Procyanidins, cyanidins and flavonols	(Blanch et al., 2012)
Leek (*Allium porrum* L.)	Sample crushed +15 min stirring, +15 min ultrasound with MeOH. Defatting	Luna C18 10 × 250 mm, 5 μm semipreparative water 0.1% formic acid/MeOH	ESI + Q-TOF MS/MS Infusion of fractions	Targeted. Identification.	Malonyl flavonol	(Di Donna et al., 2014)
Various samples: blueberries (*Vaccinium*), red cabbage (*Brassica oleracea* L.), red radish (*Raphanus raphanistrum* subsp. *sativus*), grape skins (*Vitis vinifera*) and black beans (*Phaseolus vulgaris*)	Sample frozen and ground Extraction: MeOH/formic acid/water (50:5:45, v/v/v) Clean-up OASIS HLB	UPLC: HILIC, Acquity BEH amide, 1.0 × 150 mm, 1.7 μm; water/ACN 0.4% TFA ~40 min	ESI + Q-TOF MS/MS UV	Untargeted. Identification.	71 anthocyanins	(Willemse et al., 2013)
Various samples: blueberries (*Vaccinium*), red cabbage (*Brassica oleracea* L.), red radish (*Raphanus raphanistrum* subsp. *sativus*), grape skins (*Vitis vinifera*) and black beans (*Phaseolus vulgaris*)	Sample frozen and ground Extraction: MeOH/formic acid/water (50:5:45, v/v/v) Clean-up OASIS HLB	Bidimensional off line: HILIC XBridge BEH amide, 4.6 × 150 mm, 2.5 μm; RPLC Kinetex C18 4.6 × 50 mm, 2.6 μm	ESI + Q-TOF MS/MS UV	Untargeted. Identification.	Many anthocyanins	(Willemse et al., 2014)
Grape (*Vitis vinifera* L. *cv.* Muscat of Alexandria)	Frozen and homogenized. Extraction: citrate buffer pH 5 Vortexed ~10 min Clean-up C18	UHPLC Poroshell phen-Hex 2.1 × 150 mm, 2.7 μm; water/ACN 0.1% acetic acid; 22 min	ESI- Q-TOF MS/MS	Untargeted. Identification	13 monoterpenol glyco-conjugates	(Hjelmeland et al., 2015)
Potato (*Solanum tuberosum*)	Homogenized; extraction: MeOH (× 3). Clean-up: SPE Strata C18	HPLC: Luna PFP, 2.0 × 150 mm, 3 μm; water/MeOH, 0.01% formic acid, 13 min.	ESI + LTQ-Orbitrap MS/MS	Targeted, quantitation. Validation.	α-Chaconine, α-solanine, demissidine, solasodine	(Caprioli et al., 2014)

(continued on next page)

Table 2 *(continued)*

Matrix	Extraction	Chromatography	MS	Identification/validation/quantitation	Compounds	Ref.
Strawberry (*Fragaria* × *ananassa*)	Frozen, homogenized; extraction: MeOH 80% under shaking 15 min, RT No clean-up	UHPLC: Zorbax Eclipse XDB-C18 2.1 × 100 mm, 1.8 μm; water/MeOH, 0.1% formic acid 14.5 min	ESI- Q-TOF MS/MS	Untargeted metabolite profiling.	60 sensory active metabolites	(Kårlund et al., 2015)
Grape (*Vitis vitifera* L.); cocoa beans (*Cacao Theobroma*) and apples (*Malus domestica*, red Starking variety)	Flash-frozen, ground, extraction: acetone70% in water (v/v) × 3 and of 70% MeOH in water (v/v) × 3. The extracts were combined. Cocoa sample defacted. No clean-up	HILIC: Develosil Diol-100, 1.0 × 250 mm, 5 μm. ACN/acetic acid 99:1; MeOH/water/acetic acid 94.05:4.95:1, 100 min. RP-LC: Zorbax SB-C18, 4.6 × 50 mm, 1.8 μm. Water 0.1% formic acid/ACN 2 min.	ESI- Q-TOF MS/MS UV	Untargeted LC × LC Online (Kalili et al., 2013). Off line (Kalili & de Villiers, 2009). Identification.	78 proanthocyanidins (Kalili et al., 2013), many compounds (Kalili & de Villiers, 2009)	(Kalili & de Villiers, 2009; Kalili et al., 2013)
Sorghum (*Sorghum bicolor* L. Moench)	Frozen. Ground; extraction: MeOH 70%, under sonication 1 h × 2 Clean-up C18	HPLC: Luna C18 21 × 250 mm, 4 μm, semipreparative	ESI + LTQ Orbitrap Infusion Fraction QqQ	Targeted.	Substituted deoxyanthocyanidins	(Bai et al., 2014)
Tomato (*Solanum lycopersicum* L.)	LLE with MeOH/tert-butyl methyl ether (1:1, v/v) No clean-up	HPLC: Acclaim C30 column 4.6 × 150 mm, 3 μm, quaternary solvent system, 10 min	APCI + Orbitrap Exactive MS QqQ DAD	Targeted. Identification. Quantitation	Lutein, zeaxanthin, β-carotene, α-carotene, and lycopene	(Van Meulebroek et al., 2014)
Soybeans (*Glycine max*)	Ground and sieved Extraction: MeOH/water 75:25 under sonication. Clean-up: OASIS HLB	HPLC: Zorbax SB-C18 chip, 75 μm × 150 mm, 5 μm. Water/ACN + 1% formic acid. 30 min	ESI + Q-TOF MS/MS LTQ-Orbitrap	Untargeted. Identification. Validation: surrogate standards	73 flavonoids, 16 with discriminating ability.	(Chang et al., 2012)
Beans *cv* Sarconi (*Phaseolus vulgaris* L.)	Finely ground Extraction: EtOH 2 h sonication RT No clean-up	UHPLC: Accucore 150 C18, 4.6 × 150 mm, 2.6 μm. Water/ACN + 0.1% formic acid. 30 min	ESI + FTICR MS/MS	Untargeted. Identification. Laser induced dissociation	Many soya saponins.	(Bianco et al., 2015)
Wheat (*Triticum aestivum* L.)	Milled. Free fraction extraction: chilled ethanol (× 2). Bonded fraction: alkaline + acid hydrolysis, acidified and extracted with ethyl acetate.	UHPLC: Zorbax Eclipse plus 4.6 × 150 mm, 1.8 μm. Water 0.5% acetic acid/ACN. 72 min	ESI- TOF MS	Untargeted. Tentative identification.	104 phenolics	(Dinelli et al., 2009, 2011)
Cocoa beans (*Cacao Theobroma*)	Milled to a powder; defacting Extraction: MeOH 70% stirring overnight	No separation	ESI- FTICR MS/MS	Untargeted metabolomics. Tentative identification.	About 3000 formula assigned	(Milev et al., 2014)
Cocoa beans (*Cacao Theobroma*)	Milled to a powder; defacting Extraction: MeOH 70% stirring overnight	HPLC: C18 Water 0.05% acetic acid/ACN. 65 min	ESI- TOF MS Ion trap	Untargeted. Identification.	32 phenolics	(Patras et al., 2014)
Walnut (*Juglans regia* L.)	Sample homogenized in an ice bath. Extraction: MeOH/water (60:40, v/v) × 3, ultrasound Defatted (× 2)	HPLC Atlantis T3 C18 2.1 × 100 mm, 3 μm. Water/ACN + 0.1% formic acid. 34 min	ESI- LTQ-Orbitrap Velos MS/MS	Untargeted. Identification	120 compounds	(Regueiro et al., 2014)
Four beer samples (Lager, Pilsen, Märzenbier and non-alcoholic beer)	Ultrasonication + stirring Clean-up by Oasis MAX SPE	HPLC: Luna C18 2.0 × 50 mm; 3 μm. Water/ACN + 0.1% formic acid. 25 min	ESI- LTQ Orbitrap Velos MS/MS DAD	Untargeted. Identification	47 phenolics	(Quifer-Rada et al., 2015)
Beer	Degassed under sonication	UHPLC: Zorbax Eclipse plus C18 2.1 × 150 mm, 1.8 μm. Water/ACN + 0.1% formic acid. 15 min	ESI + / − Q-TOF MS/MS	Untargeted metabolomics. Identification Differential analysis	150 compounds, some identification at level 1 and 2	(Hughey et al., 2016)

(continued on next page)

Table 2 *(continued)*

Matrix	Extraction	Chromatography	MS	Identification/validation/quantitation	Compounds	Ref.
Wine	Direct injection, only filtration	UPLC: Acquity BEH C18 2.1 × 200 mm, 1.7 μm; water/ACN + 7.5% formic acid.	ESI + Q-TOF MS/MS QqQ DAD	Untargeted. Identification.	Anthocyanidin-glucosides Anthocyanidin-di-glucosides, oligomeric anthocyanins	(Alberts et al., 2012; De Villiers et al., 2011)
Wine	Filtration	LC × LC. Poroshell 120 EC-C18 3 × 50 mm, 2.7 μm. Water/ACN 10 mMACNH4; ZIC HILIC 2.1 × 150 mm, 5 μm. ACN/water	ESI- TOF MS	Targeted	Some phenolics	(Greco et al., 2013)
Wine	Samples clean-up C18	UPLC Aquity C18 Water + 0.1% formic acid/MeOH. 30 min	ESI- FTICR MS DAD	Targeted/untargeted Identification	Many phenolics	(Roullier-Gall et al., 2014)
Orange (*Citrus sinensis*) juice	LLE with MeOH/ethyl acetate/petroleum ether (1:1:1, v/v/v) No clean-up	LC × LC, DP-RP: cyano 1.0 × 250 mm, 5 μm + RP18 4.6 × 100 mm monolithic	APCI + IT-TOF MS/MS DAD	Untargeted. Identification.	30 carotenoids identified	(Dugo et al., 2009)
Citrus bergamia juice	Lyophilization. Extraction: MeOH + ultrasound No clean-up	UHPLC: Kinetex C18 2.1 × 100 mm, 1.7 μm Water/ACN + 0.1% formic acid. ~7 min	ESI- IT-TOF MS/MS DAD	Untargeted. Identification. Quantification DAD No validation	17 flavonoids	(Sommella et al., 2013)
Cranberry syrup (*Vaccinium macrocarpon*)	Mixed with MeOH No clean-up	UHPLC: Zorbax Eclipse plus C18 4.6 × 150, 1.8 μm Water + 0.1% formic acid/MeOH. 30 min.	ESI +/− Q-TOF MS/MS DAD	Targeted/untargeted Validated for 4 class-representative compounds	34 phenolics	(Iswaldi et al., 2012)
Persimmon (*Diospyros kaki Thunb.*) juices	Mixed with MeOH 80%, stirring 1 h RT No clean-up	UHPLC: Zorbax Eclipse plus C18 4.6 × 150 mm, 1.8 μm. Water + 0.25% acetic acid/MeOH; 27 min.	ESI- TOF MS DAD	Untargeted. Tentative identification	74 (detected) 45 (tentatively identified) compounds	(Jiménez-Sánchez et al., 2015)
58 apple juice samples	Extraction: ethylacetate pH 7 + ethylacetate pH 1.5 No clean-up	HPLC: C18 4.6 × 250 mm, 5 μm; Water 2% acetic acid/ACN + MeOH 60 min	ESI- Q-Exactive Orbitrap MS/MS UV	Untargeted. Identification. Differential. Validation 6 standard	Phenolics	(Guo et al., 2013)
Cranberry syrups (*Vaccinium macrocarpon*)	Samples were diluted No clean-up	nanoHPLC: C18 75 μm × 100 mm, 3 μm Water + 0.1% formic acid/ACN 26 min	Nano ESI +/− TOF MS	Targeted/untargeted Identification	9 phenolics (identified) and 38 (tentatively identified)	(Contreras et al., 2015)
Tomato (*Solanum lycopersicum*)	Samples extracted with EtOH, 5 min sonication Clean-up by OASIS MAX	HPLC: Luna C18, 2.0 × 50 mm, 5 μm. Water/ACN + 0.1% formic acid.19 min.	ESI- Q-TOF MS/MS	Untargeted. Tentative identification	47 phenolics	(Vallverdú-Queralt et al., 2011)
Honey different botanical origin acacia (*Robinia pseudoacacia*), sunflower (*Helianthus annuus*), linden (*Tilia cordata*), basil (*Ocimum basilicum*), buckwheat (*Fagopyrum esculentum*), oilseed rape (*Brassica napus*), and goldenrod (*Solidago virgaurea*).	Sample diluted with NaCl solution Extraction: ethyl acetate (× 5) No clean-up	UHPLC: Hypersil Gold C18; 2.1 × 50 mm, 1.9 μm; water/ACN + 0.1% formic acid. 10 min.	ESI- LTQ Orbitrap MS/MS	Untargeted. Identification. Differential analysis	43 phenolics	(Kečkeš et al., 2013)
Virgin olive oil (VOO): 5 varieties of olive fruit called Arbequina, Picual, Hojiblanca, Lechín de Sevilla, and Cornicabra	SPE by diol phase	HPLC: XTerra MS 2.1 × 100 mm, 3.5 μm; 0.1% acetic ac./ACN, 45 min. CE: 85 cm × 50 μm; 25 mmol L^{-1} NH_4HCO_3, pH 9.0, 10 min	ESI- TOF MS	Untargeted. Identified. Confirmed. Not quantified CZE-TOF MS	HPLC: 19 known + 26 unknown phenolics, CE: 18 known + 28 unknown phenolics.	(Carrasco-Pancorbo et al., 2007)
VOO (5 varieties Cornezuelo, Manzanilla, Hojiblanca, Picual and Arbequisur)	SPE by diol phase	UHPLC: Zorbax Eclipse plus 4.6 × 150 mm, 1.8 μm. 0.5% acetic acid/ACN. 33 min	ESI- TOF MS IT-MS[n]	Untargeted. Identification.	19 phenolics	(Fu et al., 2009)

(continued on next page)

Table 2 (continued)

Matrix	Extraction	Chromatography	MS	Identification/validation/quantitation	Compounds	Ref.
VOO (3 varieties: Picual, Hojiblanca and Arbequina)	SPE by diol phase	Nano-LC: C18 BioSphere, 75 μm × 10 cm, 3 μm. 0.5% acetic acid/ACN. 45 min. UHPLC: Zorbax Eclipse plus 4.6 × 150 mm, 1.8 μm. 0.5% acetic acid/ACN. 33 min	ESI-TOF MS	Targeted/untargeted Identification Quantification and validation for 7 phenolics LOQs 2–4800 ng L^{-1}	18 (identified) (quantified) phenolics	(García-Villalba et al., 2010)
VOO (6 varieties: Picual, Lechín, Cornicabra, Arbequina, Hojiblanca, Picudo)	SPE by diol phase. Semi-preparative HPLC column, 10 × 250 mm, 5 fractions	Semipreparative HPLC: Gemini C18 10 × 250 mm, 5 μm. CE: 85 cm × 50 μm; 40 mmol L^{-1} NH_4HCO_3, pH 9.5, 10 min/fraction	ESI-TOF MS CZE Q-TOF MS/MS	Untargeted. Tentative identification. Quantification.	75 compounds, most phenolics	(García-Villalba et al., 2009)
VOO (14 varieties: 2 Hojiblanca, 7 Picual, 1 Corneznelo, 1 Manzanilla, 3 Arbequina)	SPE by diol phase	UHPLC: Zorbax Eclipse plus 4.6 × 150 mm, 1.8 μm. Water 0.25% acet. ac./MeOH ~ 35 min	ESI-TOF MS DAD	Untargeted. Identification. Quantification.	66 compounds, 32 phenolics	(Lozano-Sánchez et al., 2010)
25 VOO samples from Arbequina, Koroneiki, Arbosana, Grappolo, Manzanilla, Coratina, Frantoio and MGS Mariense varieties	LLE, partition between hexane and MeOH.	UHPLC: Zorbax Eclipse plus 4.6 × 150 mm, 1.8 μm. Water 0.25% acet. ac./MeOH ~ 35 min	ESI-TOF MS DAD	Targeted/untargeted Identification. Validation for 6 comp.	18 compounds	(Balus et al., 2015)
Extra VOO (cv Leccino, Frantoio, Carboncella)	LLE, optimized by experimental design. Comparison with SPE	UHPLC: Zorbax Eclipse plus 2.1 × 150 mm, 1.8 μm. Water/ACN + 0.1% formic acid. 30.1 min	ESI +/− Q-TOF MS/MS	Untargeted. Identification. No quantification	ESI −: 44 identified phenolics, 28 unidentified ESI +: many unidentif. comp. with a nitrogen atom.	(Capriotti et al., 2014)

metabolite polarity. Moreover, very different chromatographic and MS conditions are necessary for analysis; therefore, the "polar" and "non-polar" compounds are usually extracted and analyzed separately.

2.1.1. Extraction of semi-solid and solid samples

As the first step, the vegetal sample is lyophilized and ground when necessary, to ensure a homogeneous sampling from processing of small powdered aliquots (Cavaliere, Foglia, Pastorini, Samperi, & Laganà, 2005). At this point, compounds should be extracted from the matrix. The extraction is carried out under mixing and/or ultrasonication (see column "Extraction" of Table 2). In some cases, to increase the rate of dissolution, more drastic conditions can be used, such as homogenization combined with accelerated solvent extraction (ASE) (Ares, Bernal, Nozal, Turner, & Plaza, 2015; Bijttebier, Zhani et al., 2014), or Soxhlet extraction (Pierson et al., 2014). In ASE, analytes are extracted at high temperature (50–200 °C) and pressure (500–3000 psi) for a short time (5–10 min). This technique complies with the principles of green chemistry, as it combines the benefits of high-throughput, automation and low solvent consumption in closed systems, thus reducing the risks for both operators and environment, but expensive equipment is required (Armenta & de la Guardia, 2015). Such drastic conditions need to be evaluated carefully, since both time and energy transfer increase the risk of artifacts. In this respect, ultrasonication represents the mildest solution.

Solid samples with low water content, such as soybeans (Chang et al., 2012), common beans (Bianco, Buchicchio, & Cataldi, 2015) and wheat grains (Dinelli et al., 2009, 2011) are usually ground and extracted with a solvent by sonication (Bianco et al., 2015; Chang et al., 2012) or by blending (Dinelli et al., 2009, 2011). Solid fatty samples, such as cocoa beans, are usually ground and defatted by means of an organic solvent, such as hexane under stirring, prior to analyte extraction (Kalili & de Villiers, 2009; Milev, Patras, Dittmar, Vrancken, & Kuhnert, 2014; Patras, Milev, Vrancken, & Kuhnert, 2014).

Fruit samples, such as grape, tomato and berries, are frozen and homogenized with the extraction solvent, keeping samples in the dark and/or at low temperature when degradation may occur; in some cases the extraction is assisted by sonication (see column "Extraction" of Table 2).

Given the diversity in analyte features, solvent characteristics, solvent-to-sample ratio, time of extraction and temperature are important variables to be considered during extraction. A compromise should be found but, when dealing with HRMS analysis, in only one work the extraction variables were actually optimized by an experimental design (Van Meulebroek, Vanden Bussche, Steppe, & Vanhaecke, 2014). This would enable a strong reduction of the number of experiments needed for optimization of the most relevant variables, taking into account the selectivity of the extraction procedure in terms of interfering compounds and matrix effect (ME). However, it should be pointed out that, despite the advantage just described, an experimental design could be useful only for relatively homogeneous matrices and for a class of compounds with similar physico-chemical characteristics. In fact, the solvent to be used for extraction strongly depends on the polarity of the compound class(es) to be extracted. As an example, for extraction of phenolics, mixtures of methanol (MeOH), ethanol or acetone and water, either acidified or not, are commonly employed, whereas less polar solvents, such as MeOH or MeOH/acetone are preferred for the extraction of a larger set of phytochemicals. By using solvents with different polarities, sometimes it is possible to obtain extracts with different composition and complementary information (Navarro, Núñez, Saurina, Hernández-Cassou, & Puignou, 2014). For instance, for metabolomic analysis in tomato juices, a series of liquid-liquid extractions (LLEs), first with polar (MeOH) and then apolar (hexane/acetone mixture) solvents were performed and fractions individually analyzed (Cichon, Riedl, & Schwartz, 2017).

A final remark is necessary for bound compounds, which need a chemical treatment to release them from the matrix. As an example, in

whole grain bound phenolics could be determined only after basic and acid hydrolysis (Dinelli et al., 2009, 2011). After each hydrolytic step, free phenolics were extracted with ethyl acetate and were analyzed as the free fraction.

2.1.2. Clean-up step

When utilizing targeted analysis or a predefined class profiling, sample preparation can comprise a clean-up step, usually consisting in a LLE or solid phase extraction (SPE). However, to avoid possible degradation and to save time, it is important to maintain the procedure as simple and rapid as possible. In this respect, the greater sensitivity in full scan mode typical of the HRMS analyzers allows working with less concentrated crude extracts with reduced ME, thus making a clean-up step unnecessary in many cases.

This choice is obvious for metabolomic approaches and, in general, when quantification is outside the scope of the work and only identification and confirmation of compounds are pursued; indeed, for a quantitative determination, the ME could cause inaccuracy for many matrices. However, a clean-up step is sometimes added to the procedure also for non-quantitative investigations. LLE has been used as pretreatment either discarding the apolar phase for defatting purpose (Di Donna, Mazzotti, Taverna, Napoli, & Sindona, 2014; Di Donna et al., 2007; Mikołajczyk-Bator, Błaszczyk, Czyżniejewski, & Kachlicki, 2016; Pierson et al., 2014; Ramirez, Zambrano, Sepúlveda, Kennelly, & Simirgiotis, 2015), or discarding the polar phase to eliminate polar compounds (Bijttebier, Zhani et al., 2014). Defatting can be performed on extracts as well, as the case of walnut kernels; they were first homogenized, then extracted with a polar solvent and the extract defatted by LLE with hexane (Regueiro et al., 2014). As a whole, LLE is not a very selective technique, thus a more refined matrix component removal may be needed. In this regard, SPE is the technique of choice. Besides, SPE provides an additional advantage over LLE, since the investigated compounds can be more selectively retained by the stationary phase and concentrated in the eluate. The classical C18 adsorbent has been employed to isolate single classes of compounds (Caprioli, Cahill, Vittori, & James, 2014; Hjelmeland, Zweigenbaum, & Ebeler, 2015; Mikołajczyk-Bator et al., 2016) but also to pre-fractionate complex extracts, by means of both disposable cartridges (Bai et al., 2014; Lai et al., 2013; Navarro et al., 2014; Passo Tsamo et al., 2015) or semi-preparative columns (Di Donna et al., 2014; Pierson et al., 2014). However, for SPE several phases are commercially available and the choice among them depends on the physico-chemical nature of the investigated compounds and on the matrix components which need to be removed. For instance, mixed polar-non polar resins, such as OASIS HLB, were employed to isolate anthocyanins (Lin, Sun, Chen, & Harnly, 2011; Sun, Lin, & Chen, 2012; Willemse, Stander, & de Villiers, 2013; Willemse, Stander, Tredoux, & de Villiers, 2014), while the Amberlite XAD7 resin, a nonionic macroreticular resin that adsorbs and releases ionic species through hydrophobic and polar interactions, was employed for isolating both anthocyanins (Ramirez et al., 2015) and less polar phenolics (Simirgiotis, Ramirez, Schmeda Hirschmann, & Kennelly, 2013). The mixed mode cation exchange OASIS MCX was used to remove anthocyanins, while other phenolic compounds were not retained (De Rosso et al., 2014). On the contrary, OASIS MAX was used to retains phenolics, which then were selectively recovered by an acidified eluent (Vallverdú-Queralt, Jáuregui, Medina-Remón, Andrés-Lacueva, & Lamuela-Raventós, 2010; Vallverdú-Queralt, Jáuregui, Medina-Remón, & Lamuela-Raventós, 2012). Size exclusion chromatography was employed in proanthocyanidins analysis, which are present in plants and fruits as polymers (Kimura, Ogawa, Akihiro, & Yokota, 2011).

2.2. Extraction of liquid samples

When dealing with liquid samples, extraction is usually avoided and substituted by a simple filtration and/or degassing, depending on the sample characteristics.

Alcoholic beverages, including beer (Hughey, McMinn, & Phung, 2016; Quifer-Rada et al., 2015) and wine (Alberts, Stander, & De Villiers, 2012; De Villiers, Cabooter, Lynen, Desmet, & Sandra, 2011; Greco, Grosse, & Letzel, 2013; Roullier-Gall, Lucio, Noret, Schmitt-Kopplin, & Gougeon, 2014), and non-alcoholic beverages, such as fruit juice and syrup (Contreras, Arráez-Román, Fernández-Gutiérrez, & Segura-Carretero, 2015; Dugo, Giuffrida, Herrero, Donato, & Mondello, 2009; Guo, Yue, Yuan, & Wang, 2013; Iswaldi et al., 2012; Jiménez-Sánchez et al., 2015; Sommella et al., 2013), tomato sauce (Vallverdú-Queralt, Jáuregui, Di Lecce, Andrés-Lacueva, & Lamuela-Raventós, 2011) and honey (Kečkeš et al., 2013) have been treated in different ways before instrumental analysis. Beer was first degassed by ultrasonication, and then filtered for metabolomics study or an aliquot was cleaned-up by an OASIS MAX cartridge for the selective extraction of phenolics. Wine was directly analyzed without any pretreatment with good qualitative results, but quantitation was outside the purpose of the work, therefore the ME was not evaluated. Non-polar compounds, such as carotenoids, were extracted from orange juice with a non-polar solvent mixture (Dugo et al., 2009). For a broad polarity compound extraction, apple juice was repeatedly extracted with ethyl acetate at two pH values (Guo et al., 2013), while for the extraction of polar flavonoids from citrus juice, the sample was first lyophilized, then extracted with MeOH under sonication (Sommella et al., 2013). Cranberry syrup concentrate, persimmon juice and honey were simply diluted and filtered (Contreras et al., 2015; Iswaldi et al., 2012; Jiménez-Sánchez et al., 2015; Kečkeš et al., 2013). For tomato sauce, an additional clean-up step by OASIS MAX was added (Vallverdú-Queralt et al., 2011).

Among the different liquid samples, a special attention should to be paid to oils. Oil is a liquid fatty matrix with peculiar characteristics. Virgin olive oil (VOO) is unique among the vegetable oils because it is obtained from the olive drupes by solely mechanical means with no other treatment, and thus it contains phytochemicals that are usually eliminated from other vegetable oils during refining. Polar phytochemicals can be extracted from the oily matrix by normal phase SPE on diol modified silica or by LLE with a polar solvent, such as MeOH or MeOH/water mixture. The SPE extraction is a well-established procedure and uses a relatively large sample amount (60 mL) mixed with an equal volume of hexane. During SPE fats are eluted with sample loading and washing with hexane, whereas the compounds of interest are finally recovered with MeOH. The LLE works in a similar mode, but in this case only 1 g of oil, previously diluted with 1 mL of hexane, is extracted with 2 mL of MeOH. Despite the disadvantage provided by large volumes, SPE is the most widely used technique (Carrasco-Pancorbo, Neusüß, Pelsing, Segura-Carretero, & Fernandez Gutiérrez, 2007; Fu et al., 2009; García-Villalba et al., 2009, 2010; Lozano-Sánchez et al., 2010), whereas only one study uses the classical LLE by hexane/MeOH partitioning (Ballus et al., 2015). However, a recent paper demonstrated some limitations of the diol SPE extraction method, namely recoveries were lower than those obtained with MeOH/hexane LLE for both the most and the least polar phenolic compounds (Capriotti et al., 2014). Moreover, the silica-based diol phase irreversibly adsorbs some presumably basic compounds, the nature of which should be still elucidated and might be of importance in determining the VOO characteristics.

2.3. Pitfalls in sample preparation of natural compounds

As a general rule, samples should be immediately frozen to avoid enzymatic reactions. The most rapid and simple method is to place the samples in liquid nitrogen to freeze them shortly after harvesting. However, during all the analytical procedure steps, in particular during sample handling and extraction, degradation or transformation of analytes of interest could occur. For example, Feuereisen, Gamero Barraza, Zimmermann, Schieber, and Schulze-Kaysers (2017)

demonstrated that using pressurized liquid extraction at high temperatures (> 75 °C), caused phenolic artifact formation. Another example of natural compounds prone to degradation are the glucosinolates that are present in *Brassicaceae* family plants. Indeed, myrosinase is the only known enzyme that cleaves a thio-linked glucose found in nature; although myrosinase and glucosinolates are stored in separate and different cell types, however they come into contact when plant tissues are disrupted by mechanical treatments (chewing, chopping, or cutting). Thus, in the presence of water, myrosinase hydrolyzes glucosinolates to yield a variety of degradation products, including isothiocyanates (Bell & Wagstaff, 2014; Lee, Lim, Kim, & Lee, 2017). For this reason, glucosinolates extraction is generally performed with MeOH, water or their mixtures at temperatures > 70 °C to inhibit myrosinase activity (Ares, Nozal, Bernal, & Bernal, 2014).

In the analysis of olive oil, Karkoula, Skantzari, Melliou, and Magiatis (2012) described the reactivity of oleocanthal and oleacein with MeOH and water, which are commonly used both for the extraction of polyphenols and as components of mobile phase during their liquid chromatography analysis, leading to the formation of several artifacts. Therefore, even in the instrumental determination step, target analytes can be transformed. This is also the case when large conjugated compounds are analyzed by electrospray ionization (ESI)-MS. In fact, dealing with O-glycosylated flavonoids containing more sugar moieties, ESI in-source fragmentation is easily observed (Cavaliere et al., 2005), unless very mild conditions are chosen (e.g., low declustering potential values, gas pressures etc.). In some cases, in-source fragmentation can be favored to improve the MS/MS information on the aglycone moiety (La Barbera et al., 2017).

2.4. Final remarks

It is noteworthy that in the majority of the reviewed works, especially in untargeted approaches, sample preparation was performed by protocols based on consolidated experience, without any form of performance evaluation. Obviously, it is objectively difficult to apply the methodologies well established for the targeted analysis to samples containing unknown compounds and determine recovery, ME and intermediate precision evaluation. Still, changing the matrix may greatly affect the process yield, in targeted analysis as well. In this context, recently a clear analysis of the effect of sample preparation was performed in a very systematic manner (Bijttebier et al., 2016), taking as starting point the observation that, despite the importance of the extraction protocol design in obtaining accurate and reproducible data, "Current extraction methods developed for plant metabolomics do not allow collection of the entire plant metabolome, introducing inevitable compromises". The authors' purpose was to evaluate the currently available methods for comprehensive extraction of plant metabolites and, more importantly, to establish a validation procedure. For this reason, they considered seven extraction procedures working on two model samples. Then, the samples were analyzed under their best condition and the relative process efficiency, repeatability and intermediate precision were calculated. The authors concluded that none of the selected procedures was suitable for the whole range of analytes. Moreover, some discrepancies in process efficiency also indicated a dependence on the original matrix. The possible formation of artifacts was also considered and attributed to residual enzymatic activity. Most importantly, it was established that, in the field of untargeted metabolite analysis, high quality results strongly depend on validated sample preparation methods. Thus, to conclude, the impression is that researchers often underestimate the importance of the pre-instrumental steps.

3. Liquid chromatography

The analysis of individual phytochemicals in complex extracts requires efficient separation methods prior to their detection. In this regard, since the middle '80s HPLC was recognized as the most versatile technique. The latest developments of HPLC, including very pH-stable stationary phases, sub-2 μm particles and fused-core particles, have considerably improved the performance of HPLC systems in terms of resolution, peak capacity, reproducibility and ruggedness. Efficiencies exceeding 80,000 plates and peak capacities over 600 can actually be routinely attained; therefore, the gap with GC performances has been considerably reduced, with the advantage of simpler sample preparation requirements.

Fully porous, high-quality packing material with diameter sub-2 μm was introduced in the early 2000s and initially used for packing long capillary columns that were used in a homemade, very high pressure system (Mellors & Jorgenson, 2004). However, very soon both columns and suitable liquid chromatography apparatuses become commercially available. Two years later, UPLC coupled with both diode-array detection (DAD) and MS/MS was applied to phytochemical analysis for the first time (Zhou, Xu, Xue, Zhang, & Liang, 2006).

3.1. UHPLC

UHPLC, both with fully porous or fused-core particles, may represent the ideal complement to HRMS to achieve a broader coverage of the compounds present in complex samples (Kaufmann, 2014). However, some constraints should be considered for successfully coupling UHPLC with HRMS (Rodriguez-Aller, Gurny, Veuthey, & Guillarme, 2012):

- *Mobile phase flow rate compatibility with ESI source*. Without the technical solutions (modification of the source design, increase of pneumatic assistance) available in the latest generation instrumentation, the optimum mobile phase flow rate of ESI sources is generally between 50 and 300 μL min^{-1}. The optimal flow rate for 2.1 mm i.d. columns packed with 5 μm particles is about 200 μL min^{-1}, whereas for columns packed with sub-2 μm particles in UHPLC-MS systems it is approximately 3 times larger.
- *Extra-column band broadening*. Again, this problem is more severe in older generation instruments, due to both instrument dead volumes and MS detector dispersion (Grata et al., 2009). The new generation instruments have very low dead volumes as well as improved source design. However, the loss of peak capacity could be still > 30% for short columns and fast gradients.
- *Frictional heating effect*. This effect can be observed at elevated mobile phase linear velocity; it causes additional pressure drop and band dispersion due to a temperature gradient inside the column.
- *Acquisition rate*. Peaks generated using UHPLC columns are narrower than those generated in conventional HPLC in the same conditions. As at least eight acquisition points are needed to define a chromatographic peak with acceptable accuracy, faster acquisition speeds are required.

Latest generation instruments have been designed to overcome these constraints; liquid chromatographs are able to bear very high maximum pressure limits (up to 1500 bar), and are equipped with a thermostatted column compartment with mobile phase pre-heating. These technical solutions (as well as the new generation ESI sources) are compatible with 2.1 mm i.d. UHPLC columns. With older instruments, a flow splitter between the column and the source could be inserted, provided that it is first ascertained this device does not introduce additional band broadening. Alternatively, 1 mm i.d. columns can be employed; however, this column size is not available by all suppliers and/or for all the stationary phases. The last issue will be discussed in a next paragraph.

Nevertheless, in the field of phytochemistry research several applications exploit UHPLC both with fully porous (Abu-Reidah, Contreras, Arráez-Román, Fernández-Gutiérrez, & Segura-Carretero, 2014; Abu-Reidah, Contreras, Arráez-Román, Segura-Carretero, & Fernández-

Gutiérrez, 2013; Abu-Reidah et al., 2012; Alberts et al., 2012; Ares et al., 2015; Ballus et al., 2015; Bijttebier, Zhani et al., 2014; Capriotti et al., 2014; Chong, Mcghie, Heyes, & Stowell, 2013; De Paepe et al., 2013; De Rosso et al., 2014; De Villiers et al., 2011; Dinelli et al., 2009, 2011; Farag, El-Ahmady, Elian, & Wessjohann, 2013; Flamini et al., 2013; Fu et al., 2009; He et al., 2016; Hughey et al., 2016; Hurtado-Fernández et al., 2011; Iswaldi et al., 2012; Jiménez-Sánchez, Lozano-Sánchez, Rodríguez-Pérez, Segura-Carretero, & Fernández-Gutiérrez, 2016; Jiménez-Sánchez et al., 2015; Kårlund, Hanhineva, Lehtonen, Karjalainen, & Sandell, 2015; Kečkeš et al., 2013; Kimura et al., 2011; Lin et al., 2011; Lozano-Sánchez et al., 2010; Mikołajczyk-Bator et al., 2016; Morales-Soto, Gómez-Caravaca, García-Salas, Segura-Carretero, & Fernández-Gutiérrez, 2013; Passo Tsamo et al., 2015; Roullier-Gall et al., 2014; Sun, Liu, Yang, Slovin, & Chen, 2014; Sun et al., 2013, 2012; Willemse et al., 2013) or fused-core particles (Abrankó, Nagy, Szilvássy, Stefanovits-Bányai, & Hegedus, 2015; Bianco et al., 2015; Hjelmeland et al., 2015; López-Cobo, Gómez-Caravaca, Cerretani, Segura-Carretero, & Fernández-Gutiérrez, 2014; Navarro et al., 2014; Sommella et al., 2013). In most of the cases, the flow rates are lower than the optimal one for the column used, or higher than the optimal one for the ESI source (only in few cases the flow was split). When quantification is performed by UV and identification by MS, the effluent from the column can be split in order to obtain a suitable flow rate for both detectors. Mobile phases are usually mixtures of water/acetonitrile (ACN), acidified with small amounts of formic or acetic acid, or formic acid and ammonium formate. To obtain better peak shapes for anthocyanins and anthocyanidins, the formic acid percentage should be increased up to 7.5% (Alberts et al., 2012; De Villiers et al., 2011; Willemse et al., 2014). The slow interconversion between carbinol-pseudobase and flavylium cationic species in solution causes remarkable band broadening under conventional reversed phase (RP)-LC conditions. An increase in acidity and temperature causes an increase in transformation rate. Moreover, a very slow linear velocity should be used to further improve resolution. When using hydrophilic interaction chromatography (HILIC), a smaller percentage of a stronger acid at lower concentration (0.4% trifluoroacetic acid) should be used to avoid the drastic reduction of t_R due to the excessive polarity of the HILIC phase (Willemse et al., 2013). MeOH is rarely used as organic modifier (Ballus et al., 2015; Jaiswal & Kuhnert, 2014).

An unusually low polarity mobile phase system, consisting of 5 mmol L^{-1} ammonium acetate/MeOH/ACN/ethyl acetate as weak phase and ACN/ethyl acetate as strong phase, was used to elute very hydrophobic compounds, such as xanthin diesters (Bijttebier, Zhani et al., 2014) or carotenoids (Bijttebier, D'Hondt et al., 2014). For separation of carotenoids, a C30 column was used, with 2-methoxy-2-methylpropane as the less polar modifier (Cichon et al., 2017) due to its ability in separating cis/trans isomers, and it was compared to six UHPLC C18 columns (particle size 1.7 or 1.8 μm) (Bijttebier, D'Hondt et al., 2014); the authors concluded that while UHPLC analysis is more suited for rapid screening of carotenoids, for analysis of complex mixtures, the HPLC C30 column gives better resolution. The advantage of C30 column was offset by the fact that analysis on this column took about four times longer than UHPLC (100 min vs. 23 min, respectively). However, one must now consider that at the time when this comparison was performed, UHPLC C30 columns were not available, while now core-shell C30 columns with particle size 2.7 μm are commercially available and might become a first choice for xanthin and carotenoid separation.

3.2. HPLC

Although the advantages of UHPLC are fully recognized and this technique has become popular since ca. 2010, traditional HPLC (3–5 μm particle size) is still widely employed (see column "Chromatography" in Table 2). In this case, gradient times (including re-equilibration) are generally longer than for UHPLC systems, usually between 20 and 100 min for HPLC and 7–45 min for UHPLC, with significant exceptions. In few cases, conventional HPLC can be used to fractionate complex mixtures, as the case of fractionation of crude extracts by semipreparative HPLC. This process increases the final number of samples but decreases their complexity, thus fractions can be directly infused into the HRMS system (Di Donna et al., 2007, 2014).

Nano-HPLC with 75 μm i.d. column, in the form of 100 mm long column (Contreras et al., 2015; García-Villalba et al., 2010) or 150 mm chip (Chang et al., 2012), has been rarely employed. The 100 mm column performances were compared to those of an UPLC column (4.6 × 150 mm, 1.8 μm particle size), highlighting its better limits of detection but also pointing out the worse reproducibility.

3.3. Two-dimension chromatography

Despite the very high efficiency, sometimes one dimension LC does not provide sufficient resolution for complex samples separation. Two-dimension comprehensive chromatography (LC × LC) is a technique involving two different separation processes exploited in sequence with high level of orthogonality.

An on-line LC × LC system has been used for a better separation of the complex pattern of epoxycarotenoids esters in orange juice (Dugo et al., 2009). In this work, the first dimension fractionation was performed by normal phase LC with a relatively long and small i.d. column, working at a very low flow rate and a fast mobile phase program. Two-minute fractions (20 μL) were automatically recovered and re-injected by a switching valve into a larger i.d. RP monolithic column, working with an elevate flow rate and very fast solvent program. In this way, the analysis in the second dimension is completed (including re-equilibration time) before the injection of a new fraction coming from the first dimension. The very high flow rate of the second dimension ensures compatibility between the mobile phases. The orthogonality of the two techniques ensures better resolution than with an optimized one-dimension method, although both columns do not work under their best conditions.

Another on-line LC × LC orthogonal system was created by coupling HILIC with RPLC. The rationale of column coupling was similar to that reported above: proanthocyanidins were separated in the first dimension by a relatively long, small i.d. column working at very low flow rate and fast mobile phase program, while the second dimension was performed with a larger i.d. column and fast gradient (Kalili, Vestner, Stander, & De Villiers, 2013). To ensure solvent compatibility, only a fraction of the effluent collected from the first column was injected into the second column by means of a splitter.

A serial coupling of RP (first dimension) with HILIC, which used a simple T mixer, was exploited for better separation of polar and non-polar phenols in wine. The flow rate of the RP column was ten times lower than that of the HILIC column. In this way, the polar compounds, eluted unretained from the first column with the water-rich phase, were separated by HILIC with a mobile phase ten time richer in ACN, while the less polar compounds, separated in the first dimension, travelled unretained through the second dimension. Both solvent programs must be accurately designed to avoid reverse gradients, with a consequent peak broadening (Greco et al., 2013).

A different configuration (parallel instead of serial) was proposed by Pyke et al. (2015) for a generic profiling of small molecules. In addition, in this case the sample was injected into the RP precolumn at very low % of ACN and flow rate. The very polar compounds were not retained and were diluted to > 90% ACN before entering the HILIC column. Then, the elution program started, while the RP column was still under low flow rate at the initial conditions. The flow rate was increased and the gradient started only after the elution from HILIC column was completed. This configuration is very interesting, but it has never been used for phytochemical analysis, yet.

Off-line LC × LC was also performed for separation of proanthocyanidins (Kalili & de Villiers, 2009) and anthocyanins (Willemse et al.,

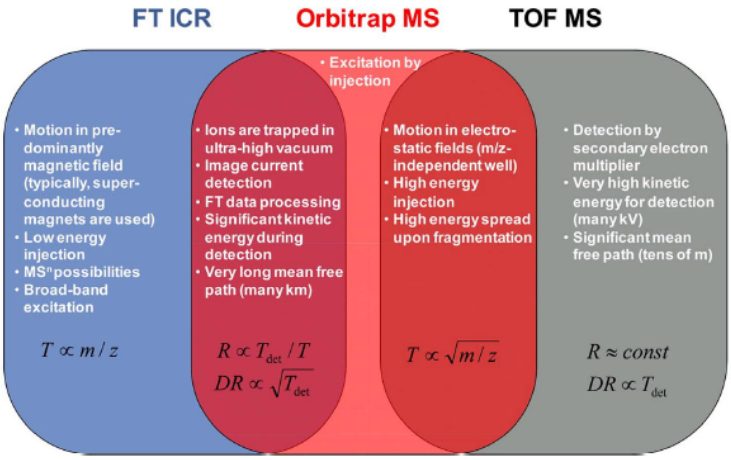


Fig. 1. Scheme comparing the main features of HRMS techniques (T is period of oscillation, R is resolving power, DR is in-spectrum dynamic range, T_{det} is duration of detection per one spectrum). Reprinted with permission from (Zubarev & Makarov, 2013). Copyright (2013) American Chemical Society.

2014). In both cases, HILIC was used for the first dimension and RP for the second dimension. The off-line setting certainly facilitates the coupling, overcoming a degree of rigidity of on-line systems. However, the flow rate of the first column still needs to be maintained low and only a small portion of the injected sample can be transferred to the second separation, decreasing sensitivity.

Off-line HPLC × capillary zone electrophoresis (CZE) was also achieved and applied to extra VOO extract (García-Villalba et al., 2009). Seventeen LC fractions were collected and then analyzed by CZE-HRMS/MS.

4. High resolution mass spectrometry and tandem mass spectrometry

A detailed discussion of this issue is outside the purpose of this review. For a simple, but exhaustive overview, the reader may refer to the feature article by Zubarev and Makarov (2013), the review by Forcisi et al. (2013) and the tutorial by Scigelova, Hornshaw, Giannakopulos, and Makarov (2011).

4.1. General remarks

Mass spectrometry has been used for a long time in a variety of food applications. Recent years, however, have witnessed a remarkable increase in the employment of high resolution mass spectrometers. Depending on the physical principle of operation, high resolution analyzers are classified into time of flight (TOF), Fourier transform ion cyclotron resonance (FTICR) and Orbitrap (Zubarev & Makarov, 2013). They could be further combined with quadrupole and quadrupole traps to allow analysis of both intact ions coming from analytes and their fragmentation pattern. Alternatively, the same analyzer can perform MS and MS^2, $MS^{n>2}$, as for radiofrequency or static electromagnetic FT ion traps.

The mass resolution is the measure of the ability to distinguish two peaks of different *m/z*, and it is measured as full width at half maximum (FWHM). The narrower the peak width, the higher the resolution. The mass accuracy is the ratio of the *m/z* measurement error to the true *m/z* value, and it is usually expressed in parts per million (ppm). The advantage of measuring a compound mass with a suitable high accuracy arises from the fact that, in this way, it is possible to directly determine its elemental composition or, at least, largely reduce the possible combinations. Therefore, accurate mass acts as a powerful filter for confirming the identity of a compound or for identifying an unknown one, when such information is combined with that from different structural investigation techniques. Other aspects of the usefulness of mass accuracy for small molecule analysis include isotopic pattern. An accurate isotopic pattern determination can further reduce the possible accurate mass matching for different elemental compositions up to 95%.

Resolution and accurate mass are strictly related to each other. High resolution is needed because one cannot accurately perform mass measurement if some compounds remain insufficiently resolved. Moreover, there is an intrinsic relation between mass resolution and the precision of mass measurement. In this context, the resolution settings define the confidence interval within which the mass measurement matches the exact mass of the analyte.

However, what do we intend for high resolution and mass accuracy, today? There is no rule in this sense; moreover, also considering that only few laboratories have the state-of-the-art instrumentation, we can work out a reasonable lower limit, namely 15,000 FWHM and ± 5 ppm, respectively. Obviously, better performances lead to a better certainty of identification. Ultimately, the mass analyzers suitable for high mass accuracy determination are the ion beam based TOF and the image current based FTICR and Orbitrap instruments.

4.2. Fourier transform based instruments

The FTICR and Orbitrap analyzers outweigh any other mass analyzer with respect to the achievable maximum mass resolution and accuracy. In both FT analyzers, resolution (R) depends on the ratio of the detection time (T_d) to the period of main oscillations (T_ω); however, in FTICR ions orbit in the magnetic field of superconducting magnets, therefore T_ω is directly proportional to *m/z* and R inversely proportional to *m/z*. In the Orbitrap analyzer, the ion motion is determined by the electrostatic field, which leads to T_ω being proportional (and R inversely proportional) to the square root of *m/z* (Fig. 1). The consequence of this difference is that, for any FTICR and any Orbitrap device, there is a critical *m/z* below which the resolving power achieved for the same T_ω is higher for FTICR but above which the Orbitrap analyzer starts to show higher R. This value was about 4000 for the 15T FTICR and the standard Orbitrap, but became about 300 for the high field compact Orbitrap.

In FTICR instruments, field uniformity of superconducting magnets and very high accuracy of frequency measurements lead to unmatched mass resolution and mass accuracy (Fig. 1). Similarly to FTICR, in the Orbitrap mass analyzer the image current is also transformed into the frequency domain and then converted into a mass spectrum. However, as only electrical fields are employed to trap ions, Orbitrap is a small

size analyzer and makes it possible to use it in benchtop instruments, justifying the wider spread of Orbitrap instruments over FTICR ones. Besides, enhanced FT improves the performance and, for instance, offers the choice of twice the speed at the same resolving power or twice the resolving power at the same speed.

The combination of FTMS with ion selection and fragmentation devices represents a very suitable approach for small molecules analyses. Even if in theory, the Orbitrap can be operated to perform MS/MS operation, in practice it requires cycles of pressure changes inside the trap, thus the need of a fragmentation device and the possibility of working in data dependent mode. Initially, a linear trap quadrupole (LTQ) was introduced, then it was substituted by a collision cell, which enabled higher-energy collisional dissociation (HCD).

Improvement in Orbitrap technology comprise a new concept Orbitrap (Exactive, 2008), which, in addition to high mass accuracy, fast scanning, wide dynamic range as well as high sensitivity, is capable of generating fragmentation information in a non-selective manner (Bateman, Kellmann, Muenster, Papp, & Taylor, 2009). The last evolution of Orbitrap technology produced the Q-Exactive, which is equipped with a quadrupole analyzer, and Orbitrap Fusion, a three hybrid instrument in which a dual stage LTQ has been added after the collision cell and offers unmatched flexibility (Fig. 2).

4.3. Time of flight based instruments

In TOF analyzers the time of flight of the ion depends on the square root of its m/z, and resolution depends on the capacity of the instrument of eliminating the initial spread of kinetic energy of the injected ions. Since in TOF instruments detection is accomplished by secondary electron multipliers, the maximum detection efficiency is reached at a very high kinetic energy of ions; this makes R independent from detection time (T_d) or m/z but makes dynamic range directly dependent on T_d, therefore the shorter the acquisition time, the smaller the dynamic range (Fig. 1). Typically, the highest resolution available in commercial TOF devices is several times lower than the highest resolution in Orbitrap and FTICR instruments. At the same time, TOF resolution is largely independent from m/z also in the MS/MS mode, whereas in the Orbitrap analyzer the resolution in MS/MS mode is often sacrificed to increase scan speed. However, the higher resolution of TOF instruments in MS/MS mode does not always engender higher mass accuracy, due to the reduced transmission of TOF analyzers which employ orthogonal acceleration of continuously flowing ions.

4.4. Which is the best instrument?

A question of great interest is which of FT-based or TOF-based instruments is the most suited for plant metabolite studies, or whether both may be equally suitable for that purpose. A comparison between FT-based instruments (especially Orbitrap, which is supplanting the FTICR in the majority of the applications) and the TOF-based ones is neither simple nor straightforward. Seldom, the performances appear similar when comparing instrument of the similar technological level. Probably, for the instrument that can operate in MS/MS mode the true positive identifications by Orbitrap are higher than those obtainable by TOF, but chromatography can play a crucial role as well. However, at best of our knowledge, only one paper reports a comparison between a Q-TOF instrument and a single stage (Exactive plus) Orbitrap, and for a specific application in differential metabolomics (Glauser, Veyrat, Rochat, Wolfender, & Turlings, 2013) reaching the conclusion that "Surprisingly, both systems provided very similar outputs in terms of detected markers and sample classification".

Thus, given that the chromatographic system coupled to the MS instrumentation also influences the performance of this analytical step, then familiarity with the technology may play a significant role on the quality of the results. In this regard, modern hybrid high-resolution instruments have several data acquisition options, the most commonly employed being the so-called data dependent acquisition (DDA). In DDA, tandem mass spectra are acquired only when some preselected parameters, such as inclusion list, threshold intensity, charge state, etc. are satisfied. Other not DDA acquisition modes, which could produce more information, have been explored for proteomics (Chapman, Goodlett, & Masselon, 2014; Geiger, Cox, & Mann, 2010), more rarely in metabolomics (Rathahao-Paris, Alves, Junot, & Tabet, 2016). Given the above, most of the papers on the analysis of phytochemicals in food do not seem to fully test or exploit results provided by fields where HRMS has been extensively investigated, and the data acquisition methods can be even omitted at all. Besides, conventional MS is still the most employed approach and HRMS is seldom performed, mainly for confirmation of compound identity alone.

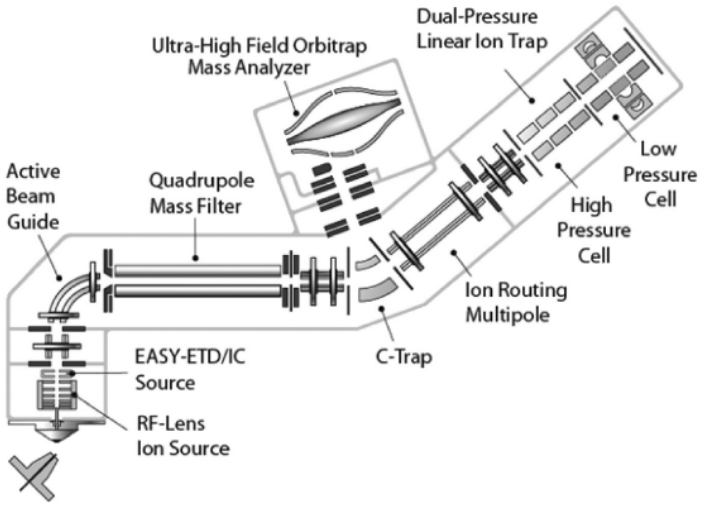


Fig. 2. Schematic representation of the structure of an Orbitrap Fusion mass spectrometer. Reprinted with permission from (Senko et al., 2013). Copyright (2013) American Chemical Society.

5. Application to real samples

5.1. Phenols, polyphenols and phenolic acids

Phenolics are by far the most investigated phytochemicals in food of vegetal origin. Sometimes the analysis is restricted to only a subclass, such as flavonols, isoflavones, stilbenes, anthocyanins, deoxyanthocyanidins, oligomeric proanthocyanidins, flavonoids. Usually, a comprehensive phenolics profiling is performed (see column "Compounds" in Table 2).

The quantitative analysis of phenolics content is really challenging; for instance, in grapes it changed significantly depending on the kind of water supply and supply rates, even for samples harvested in the same vineyard (Capriotti et al., 2012). Moreover, another study performed on grape berries evidenced that during preservation about half of the polyphenol content was lost (Cavaliere et al., 2008).

Both TOF- and, to a lesser extent, Orbitrap-based instruments have been used. As previously stated, a comparison of the usefulness of the two technologies is not possible, as instruments with very different performances have been employed and no study ever dealt with this issue.

5.1.1. Phenolics by single TOF analyzers

The simplest instrument used a single reflectron TOF analyzer, such as the micrOTOF (Bruker Daltonics) and the LCT premier XE TOF (Waters Micromass), which were benchtop robust instruments with a mass resolution exceeding 10,000 FWHM, particularly suitable for formula determination of small molecules by accurate mass (error < 5 ppm) and isotopic pattern measurements. Structural information is not provided by this kind of instrument. With only the accurate mass and isotopic abundance at hand, 34 phenolic compounds (104 taking into account isomeric compounds) were tentatively identified in different wheat and durum wheat varieties (Dinelli et al., 2009, 2011). The identification was supported by the data reported in the literature. This kind of instrumentation was also used for phenolics identification in VOO (Ballus et al., 2015; Carrasco-Pancorbo et al., 2007; Lozano-Sánchez et al., 2010). Only 18 compounds were considered as identified, based on authentic standards or consolidated literature. This type of approach, where the identification of compounds is based only on accurate mass measurement and comparison with the literature, is questionable and generally not accepted. In-source fragmentation (pseudo MS/MS spectra) and UV spectra can help bypass such limitation and acquire some useful structural information. Such an approach was employed for anthocyanins, which can lose the glycosidic moiety. In this way 31 anthocyanins were tentatively identified based on their accurate mass, the aglycone in-source fragment accurate mass and UV adsorption maxima; in addition, some compounds were confirmed by comparison with available standards (Ramirez et al., 2015). MS^n technique can solve the above issue, by repeating the analysis with an ion trap mass spectrometer and combining the formula obtained by the TOF analyzer with a careful interpretation of the collision induced dissociation MS^n spectra (Fu et al., 2009; Jaiswal & Kuhnert, 2014; Patras et al., 2014). This complementary information was used for the identification of known and new flavonoids in raw fermented cocoa beans (Patras et al., 2014), and phenolic acid glycosides in bottle gourd (Jaiswal & Kuhnert, 2014), providing 14 new flavonoids and 12 simple and complex phenolic acid glycosides.

5.1.2. Phenolics by hybrid TOF analyzers

In the recent years, the most widely used instrumentation comprised Q-TOF hybrid mass spectrometers, namely instruments that, with respect to single TOF analyzers, allow tandem mass spectra acquisition. However, as the sensitivity in the MS/MS mode is generally lower than that in simple MS, structural information for trace amount compounds could be lost. This problem is less restrictive with the last generation, more sensitive instruments. Based on their accurate mass, isotopic abundances and fragmentation pattern, 24 phenolics were identified in whole potato (López-Cobo et al., 2014), 73 in cucumber (Abu-Reidah et al., 2012), 27 in cranberry (Iswaldi et al., 2012), 47 in tomato-containing sauces (Vallverdú-Queralt et al., 2011). However, there is the need for a more rigorous identification standard procedure to confirm identifications.

As for methods based on Q-TOF technology, DAD is almost universally used because of the characteristic absorption wavelength of different classes of phenolics. Additional supplementary information was obtained by triple quadrupole (QqQ) mass spectrometer (Álvarez-Fernández, Cerezo, Cañete-Rodríguez, Troncoso, & García-Parrilla, 2015; Vallverdú-Queralt et al., 2010) or CZE instead of LC (Navarro et al., 2014). In some cases, HRMS was also used in combination with chemiometry to assess sample authenticity or sample origin (Ballus et al., 2015; Guo et al., 2013; Kečkeš et al., 2013; Lozano-Sánchez et al., 2010; Navarro et al., 2014; Roullier-Gall et al., 2014).

Although less common, the IT-TOF hybrid mass spectrometer could also be used for phenolics analysis. This instrumental configuration was successfully used for the tentative identification of flavonoids in *Citrus bergamia* juice by a 5 min-long chromatographic run (Sommella et al., 2013). A total of 17 flavonoids were identified, with only two of them based on comparison to authentic standards. Although some are very expensive, a considerable number of phenolics standard are commercially available and seldom 15–20 of them are used to confirm identifications.

In some instances, samples were analyzed by both a Q-TOF and a QqQ instrument, using product ion and neutral loss scans for confirmation, and multiple reaction monitoring (MRM) scan for quantitation purpose when pure standard of some identified compounds are available (De Rosso et al., 2014; Di Lecce et al., 2014; Vallverdú-Queralt et al., 2012). However, the neutral loss scan mode suffers from a reduced sensitivity and relatively long scan times.

5.1.3. Phenolics by Orbitrap analyzers

Compared to TOF-based technology, Orbitrap-based one has been much less employed for the identification of phenolics in food. The capability for quantitative determination of phenolics in apple samples by the Orbitrap Exactive was assessed by analyzing 39 phenolics, for which the standards were available (De Paepe et al., 2013). The Exactive model can reach a resolution up to 240,000 FWHM at m/z 200 and was used in this work at scan speed of 2 Hz, which enables a still good resolution (50,000) and the necessary scan time for peak definition. The instrumental method was fully validated in terms of trueness and precision, as well as mass accuracy (< 2 ppm); however, structural information was not available and isobaric compounds could not be differentiated.

The LTQ-Orbitrap XL was used for the identification of 15 phenolics in plantain (Passo Tsamo et al., 2015), 19 in sim fruit (*Rhodomyrtus tomentosa*) (Lai et al., 2013), 207 in red mustard greens (Sun et al., 2013), 167 in 5 *Brassica* species (Lin et al., 2011), 52 in diploid strawberries (Sun et al., 2014), 56 in fermented strawberries (Álvarez-Fernández et al., 2015), and 43 in honey (Kečkeš et al., 2013). With this Orbitrap model it is possible to determine the accurate mass of precursor ions by Orbitrap and fragment nominal masses by LTQ. As for all Orbitrap models, with reduced scan speed it is possible to have the accurate mass of both precursors and products. Eighteen compounds were identified, with an error in accurate mass measurement < 3 ppm. Compared to other papers, Sun et al. (2013) and Lin et al. (2011) reported more identifications. However, it should be taken into account that: a) both positive and negative ESI were employed; b) the identifications are putative; c) a long time gradient was used; d) *Brassicaceae* are phenolics-rich vegetables.

The LTQ Orbitrap Velos was used to characterize polyphenols in tomato (Vallverdú-Queralt et al., 2010) and in beer (Quifer-Rada et al., 2015) by the same research group. Thirty-eight phenolics were identified in tomato samples (Vallverdú-Queralt et al., 2010), only 10 of

which by comparison to standards; this is about the same number (but with only partial overlap) as in the work reported above (Vallverdú-Queralt et al., 2012), in which a Q-TOF mass spectrometer was employed. For the beer sample analysis, the LTQ Orbitrap Velos was used in both DDA mode, which resulted in the MS/MS acquisition of sufficiently intense ions, and MS^n mode for some less intense signals that were not selected in the DDA. Forty-seven phenolics were identified by comparison to standards or to the literature, and spectra interpretation. A considerably larger number of phenolics (120) was identified in walnuts (Regueiro et al., 2014). This difference could be related to the higher complexity of the sample, especially for tannins related compounds and, possibly, to less restrictive identification criteria.

Orbitrap Q-Exactive has been rarely employed (Guo et al., 2013; La Barbera et al., 2017; Zenezini Chiozzi et al., 2017). La Barbera et al. (2017) identified 113 compounds in strawberries at various confidence level, some of them never identified in strawberries before or never identified at all. This result was obtained by optimizing both the chromatographic and MS conditions.

5.1.4. Proanthocyanidins and anthocyanins

Some classes of polyphenols are usually analyzed separately. Proanthocyanidins and anthocyanins are two examples. Proanthocyanidins, also known as condensed tannins, are polymeric phenolic compounds formed through condensation of flavan-3-ol units (Kalili & de Villiers, 2009; Kalili et al., 2013; Kimura et al., 2011). Proanthocyanidins are a very complex mixture of oligomeric/isomeric compounds and fragmentation patterns are very complex and difficult to be interpreted. Anthocyanins are analyzed separately due to their structural peculiarity (see above) and better ionization in positive ESI, other than other phenolics for which negative ESI gives a better sensitivity. Anthocyanins are glycosylated anthocyanidins and exist as many isomeric compounds. The glycosidic moiety can be esterified in various positions by acids, such as acetic, hydroxybenzoic, pyruvic, and phenylpropanoics. The complementary information available from UV detection is useful, because of the uncommon absorption wavelength at about 500 nm, characteristic of these compounds. Examples of this type of studies include 13 acylated anthocyanins determined in purple red potatoes variety cultivated in China (He et al., 2016). Willemse et al. detected up to 71 anthocyanins in various vegetal matrices, with red grape the sample containing the largest variety of these compounds (Willemse et al., 2013, 2014). A strong matrix dependence is also evident. Similarly to red grape, red wine is also very rich in anthocyanins together with a variety of related compounds, such as diglucoside oligomers, flavan-3-ols adducts and reaction products with unsaturated compounds to form pyranoanthocyanins, which could be generated during wine aging. Therefore, the anthocyanins determination in wine is more complex than in grape and optimized chromatographic conditions are mandatory, also considering the presence of isomers. De Villiers and coworkers, using optimized UHPLC conditions, tentatively identified 101 anthocyanins-related compounds in red wine on the basis of accurate mass, fragmentation pattern and relative retention time (De Villiers et al., 2011). In addition, 36 proanthocyanidins were also tentatively identified in the same run. The identification process was improved due to the use of two chromatographic runs, one for the QqQ analyzer in the neutral loss mode and one in product ion scan; finally, a Q-TOF was employed for accurate mass confirmation (Alberts et al., 2012). In this way, 123 compounds were tentatively identified.

5.2. Other phytochemicals

5.2.1. Glucosinolates

Compounds different from phenolics have been investigated in edible fruits and vegetables to a much lesser extent. Glucosinolates profiling in *Brassicaceae* was obtained by ESI-LTQ-FTICR (Lelario, Bianco, Bufo, & Cataldi, 2012). The 7T FTICR was operated in MS mode at resolution of 100,000 at m/z 400 and allowed a mass error < 2 ppm. The identification of chromatographic peaks was based either on accurate mass measurements and spectra comparison with those obtained from pure standards, when commercially available, or on accurate mass measurement and interpretation of MS/MS data, resulting in 24 identifications. In a later work, a Q-TOF mass spectrometer was used, but it provided fewer identifications (Ares et al., 2015). This difference may be due to: a) analyte content dependence on sample origin; 2) different extraction methods; 3) the higher resolution and accuracy of the FTICR instrument which permitted separation of signals that otherwise merged as a fused signal, and the small m/z windows that, in turn, decreased S/N. Recently, a simple informatics tool named GLS-Finder was introduced for rapid profiling and/or identification of glucosinolates in *Brassicaceae* (Sun, Zhang, & Chen, 2016). The strategy is based on in-database searching of molecular formula, mass fragmentation behavior, and isotopic distribution using data generated in lab by UHPLC–$HRMS^n$. This program was applied to identify glucosinolates in 49 *Brassica* vegetable samples from data generated by a UHPLC-LTQ Orbitrap and significantly improved the identification rate than in previous works (reaching up to 146 identified compounds). Results were also manually verified. Glucosinolates, because of their structure and fragmentation behavior, are particularly suited for a relatively simple data handling.

5.2.2. Saponins

Saponins are a large group of complex compounds the structure of which consists of a triterpenoid moiety bonded to a complex variety of sugar units. They can be found in a variety of vegetables, as well as in some marine organisms belonging to the phylum *echinodermata*, and exert several biological activities. Twenty-five saponins were identified in amaranth species leaves by a combination of HPLC-TOF MS and direct infusion ion trap MS^n (Zehring et al., 2015). Putative identifications were based on MS^n spectra examination after selection of the candidate compounds by HRMS in both positive and negative ESI, considering that saponins can produce $[M + NH_4]^+$ and $[M\text{-}H]^-$ ions in positive and negative ionization, respectively. In another study, 44 compounds were putatively identified in *Beta vulgaris* roots by UHPLC-Q-Exactive Orbitrap MS/MS and supplementary information was obtained by UHPLC ion trap MS^n (Mikołajczyk-Bator et al., 2016). In this case, only negative ESI was performed. Such a large number of compound identifications, 27 of which identified for the first time, was reached by analyzing an extract very rich in the target compounds, obtained after a long pretreatment process. In addition, the identification of saponins based solely on MS is not unequivocal, because many isomers show similar fragmentation patterns.

Alkaloids and glycoalkaloids were analyzed in potato samples by positive ESI-LTQ Orbitrap. The approach was targeted and two species-characteristic glycosides (larger amount) and the corresponding aglycones (small amount) were identified and quantified by MRM in the LTQ (Caprioli et al., 2014). The developed method was fully validated.

5.2.3. Carotenoids and other terpenoid compounds

Carotenoids are a large class of secondary metabolites that represent the most diffuse plant pigments in nature and exert numerous biological activities. Due to their non-polar nature, this class is usually analyzed separately because it requires very different conditions for extraction, separation and ionization before MS. Carotenoids were analyzed in red-ripe and green tomato (Van Meulebroek et al., 2014). Five compounds were determined quantitatively by QqQ in MRM mode and the method was validated, using authentic standards, in positive atmospheric pressure chemical ionization (APCI). The samples were also analyzed by Orbitrap Exactive and, by comparison with a database containing 137 different chemical formulas, up to 97 positive matching hits were found. Identification was solely based on the presence of two diagnostic ions: the $[M + H]^+$ ion and the corresponding ^{13}C isotopic ion. Although unequivocal identification cannot be guaranteed due to the presence of isomers, this result demonstrates the feasibility of the

approach. Epoxy compounds, which can be considered markers of product aging, were analyzed in orange juice by positive APCI-IT-TOF MS (Dugo et al., 2009). Carotenoids were analyzed in their intact, fatty-acid esterified forms; the identification was based on the accurate mass measurement (error < 11 ppm) and fragmentation pattern in the IT, resulting in 30 putatively identified compounds including oxidized and non-oxidized ones.

Monoterpenol glycosides in grape were identified and roughly quantified by Q-TOF MS/MS (Hjelmeland et al., 2015). In this approach, an in-house database was first produced by summing up the available MS information on this class of compounds. Then, compounds present in the specific sample were identified by comparison with the database, to which t_{Rs} were also added. In this way, peak alignment could be performed and the composition of different samples could be compared.

5.2.4. Metabolomics of non-polar and polar fractions

Other published methods for the identification of phytochemicals in edible plants are claimed to be or can be considered untargeted metabolomics-based approaches (Abu-Reidah et al., 2013, 2014; Bijttebier, Zhani et al., 2014; Chong et al., 2013; Farag et al., 2013; Hughey et al., 2016; Hurtado-Fernández et al., 2011; Jiménez-Sánchez et al., 2016; Kårlund et al., 2015; Milev et al., 2014; Nakabayashi et al., 2013; Pierson et al., 2014; Simirgiotis et al., 2013). In metabolomics there is no distinction between primary and secondary metabolism, but a classification can be based on the polarity of the analyzed compounds.

From a practical point of view, only two papers deal with a generic approach for the identification of non-polar compounds. The first method was developed and tested with chili peppers (Bijttebier, Zhani et al., 2014). The sample was extracted by ASE; then, non-polar compounds were isolated LLE. Orbitrap Exactive was used for MS and only the in-source fragmentation was used for structural information. Other information was obtained from the UV spectra. By this method, about 50 carotenoids were identified also combining the spectral and MS information with comparison with standard, presence reported for the phytogenetically affine species, and previous identification from the literature. Moreover, other 70 later eluting compounds, mainly non-polar phospholipids and acylglycerols, were also identified. In the second method (Cichon et al., 2017), non-polar compounds in two types of tomato juice were analyzed by Q-TOF. Of the 423 compounds detected in the lipophilic fraction, 352 were found to be significantly different. For the polar fraction 474 compounds were detected, and 346 were found to be significantly different. Among such large number of detected phytochemicals, only few compounds were identified based on authentic standards, accurate mass, and characteristic UV–visible spectra.

Other procedures were used to analyze polar/semi-polar metabolites extracted by MeOH or MeOH (EtOH)/water mixture. In all these works, TOF-based instruments were used, and in many cases, the UV spectra supported identifications, which were in large part tentative. ESI ionization was performed in negative mode or in both positive and negative mode (see Table 2); however, the number of compounds identified did not appear related to this choice. Many compounds can be identified in both modes. However, some compounds which can be ionized only in positive mode were not detected at all if only negative ionization was used (Fu et al., 2009). Surprisingly, unlike medicinal plants, only few plants and fruits used for food have been analyzed with a non-targeted metabolomic approach (see Table 2) and only one work deals with the metabolomic characterization of a drink as popular as beer (Hughey et al., 2016).

The isotopic ratio is one of the parameters usually considered for compound identification. The reliability of this parameter could increase as the number of C atoms increases. However, the number of possible molecular formulae increases as well and some ambiguities cannot be solved, even by the very high resolution and mass accuracy achievable by FTICR instruments. The problem of determining unambiguously the number of C atoms of an ion was resolved by growing the same plant in presence of natural CO_2 and $^{13}CO_2$, thereby, the corresponding ions (if monocharged) will appear in the two mass chromatograms at $(m/z)_2 - (m/z)_1 = n$ ($m^{13}C - m^{12}C$). This expedient was used to determine the formula of sulfur containing compounds in onion (Nakabayashi et al., 2013). The isotopic ratio and relative abundance are very informative and, combined with the number of C atoms in the formula, MS/MS, and the very high resolution (260,000) and mass accuracy (< 1 ppm) of the FTICR instrument, enabled the unambiguous formula extraction of 22 sulfur containing compounds out of a dataset with 4693 features.

5.3. The quantification issue

A classical quantification method for phenolics is based on DAD, which can be used on-line or off-line with the mass spectrometer (De Paepe et al., 2013; Di Lecce et al., 2014; He et al., 2016; Lai et al., 2013; López-Cobo et al., 2014; Passo Tsamo et al., 2015; Simirgiotis et al., 2013; Sommella et al., 2013) at different wavelengths, such as 260 nm for isoflavones, 280 nm for benzoic acid derivatives, flavan-3-ols and procyanidins, 320 nm for hydroxycinnamic derivatives, 350–365 nm for flavonols, 500–520 nm for anthocyanins-anthocyanidins. The advantages of the absorbance-based approach are: the large linear range, the negligible ME and the molar extinction coefficient, substantially independent from the glycosidic moiety for many glycosylated compounds. The disadvantage is that, in presence of possible coelution, quantification is no longer possible.

Quantitative analysis is one of the most important and universal applications of MS. However, only compounds for which the standards are available may be quantified by MS. In addition, due to unpredictable ME and unavailability of blanks, the standard addition method is the most suitable one, as it circumvents the problem. MRM acquisition with QqQ instruments is a well-established procedure which has been also used for quantification when accurate mass was measured by HRMS for compound identification (Abrankó et al., 2015; Blanch, Alvarez, Sanchez-Ballesta, Escribano, & Merodio, 2012; Vallverdú-Queralt et al., 2012).

TOF instruments suffered from poor reproducibility and reduced dynamic ranges, requiring further mathematical algorithms, such as time-to-digital converter, in order to extend the linear range. This seriously limited the usability of TOF instruments for quantitative analysis. However, the problem of poor reproducibility and small linear range that affected the first high resolution TOF mass spectrometers has largely been overcome and good calibrations can be obtained with these instruments. Although this capability has been exploited for the determination of contaminants, such as pesticides in fruit and vegetables (Gómez-Ramos, Ferrer, Malato, Agüera, & Fernández-Alba, 2013), the same level of implementation was not reflected for endogenous phytochemical analysis. Only in few studies TOF MS was used for quantification (Ares et al., 2015; Ballus et al., 2015; Chong et al., 2013; Farag et al., 2013; García-Villalba et al., 2009, 2010; Hurtado-Fernández et al., 2011; Iswaldi et al., 2012; Lozano-Sánchez et al., 2010). Usually, only those compounds for which standards are available were quantified; in some cases, when standards were not available (or are not stable), such compounds were extracted and purified from a sample and later used as standards for quantitation (García-Villalba et al., 2009; Lozano-Sánchez et al., 2010). The Orbitrap-based instruments can provide reproducible quantification results (at R ≥ 50,000) and large linear range calibration (Gómez-Ramos et al., 2013); nevertheless, they are rarely exploited for quantification purpose (Bijttebier, Zhani et al., 2014; Caprioli et al., 2014; Guo et al., 2013; Van Meulebroek et al., 2014). One research reports a quantitative determination by both Orbitrap Exactive and UV (De Paepe et al., 2013). The best fit was obtained with non-linear ($1/x$) curve, but this fact cannot be considered a general rule. For complex samples, usually containing several similar compounds (oligomers, homologues, isomers) for which standards are

often not available, a complex quantification strategy has to be established. This was devised for the quantification of carotenoids in red chili peppers (Bijttebier, Zhani et al., 2014), where both UV and accurate mass data were exploited. Compounds having the same polyene backbone have also the same molar extinction coefficient, therefore, in order to determine the concentration in weight unit, the molecular mass was determined by MS and the concentration by calibration performed with only one standard belonging to one class. Compounds, for which standards were available, were quantified by calibration with the accurate mass provided by Orbitrap Q-Exactive.

When standards are not available, quantification by a standard with a similar structure has been performed (Chong et al., 2013; Farag et al., 2013). However, this operation should be handled with care as in many cases the ion desorption in ESI may be significantly influenced by small differences in structure. In comparison to this approach, more reliable approach might be the quantification of compounds based on MS when the standard is available and by UV for compounds with similar structure (Hurtado-Fernández et al., 2011).

5.4. Waste products valorization

Vegetables, fruits, and oilseeds processing steps lead to high amounts of waste materials, for instance peels, seeds, and oilseed meals. Nowadays the disposal of these materials represents a huge problem aggravated by legal restrictions. In fact, plant waste usually is subjected to microbial spoilage; thus, a drying step is necessary before further exploitation. The cost of drying, storage, and transport implies additional economical limitations to waste utilization. Therefore, agro-industrial waste is often utilized as feed or fertilizer. However, it is well known that by-products represent an important source of sugars, minerals, organic acids, dietary fiber, as well as bioactive compounds, such as the phenolic compounds. Thus, new aspects concerning the use of these wastes as by-products for further exploitation on the production of high nutritional value food additives or supplements have gained increasing interest because their recovery may be economically attractive (Ravindran & Jaiswal, 2016). As an example, Fig. 3 portraits an overview of the recovery of various dietary components from fruit pomace (Djilas, Čanadanović-Brunet, & Ćetković, 2009).

Grape is the largest fruit crop in the word; almost 80% of it is used in wine making and produces a huge amount of waste. A small amount is used for animal feed, or as fertilizers but, due to problems associated with such use (Fontana, Antoniolli, & Bottini, 2013), most of this waste still represents a disposal problem.

Apple production in the world is about forty millions of metric tons, a remarkable portion of which is intended for the production of both alcoholic and nonalcoholic beverages and generates a left-over biomass in huge quantities. Production of olive oil is one of the fastest growing agro-food sectors in the world. Only about 15–20% of the harvested olives can be transformed into olive oil, while the remnant constitutes the production waste.

The products reported above were the first to be addressed for the reuse of agro-industries manufacturing wastes, but more recently many other production residues were interested by the concept of green industry. The phytochemical characterization of byproducts is an important starting point to design waste reuse procedures. Nevertheless, HRMS has experienced little attention in this field and only in the most recent years, a relatively small number of investigations have been carried out by this technique.

Method of extraction and determination of phenolics in grape pomace were recently reviewed (Fontana et al., 2013). Extraction methods were very similar to that used for grape, starting preferably from a lyophilized sample. Then, separation was performed by HPLC and identification/quantitation were usually done by DAD or low resolution MS and MS/MS. HPLC-FTICR-MS was used to increase information about flavan-3-ols compositions (Rockenbach et al., 2012). Flavan-3-ols are a very complex mixture of procyanidins and galloyl substituted procyanidins, and 251 compounds were tentatively identified by integrating the information from the three different approaches. The analysis time was > 130 min. This is not surprising, considering the long scan time needed by FTICR to maintain high resolution.

Anthocyanins composition was determined in wine grape pomace by HPLC DAD/positive ESI Orbitrap XL MS^3 and fractionation followed by MALDI-TOF (Oliveira, Alhinho Da Silva, Teixeira, De Freitas, & Salas, 2015). MALDI TOF was used to confirm the presence of anthocyanins dimers, as they can be generated in ESI but not in MALDI ionization. The number of compounds identified was much lower than that in wine (Alberts et al., 2012; De Villiers et al., 2011). This difference might be due to the low sensitivity of positive ESI and/or to the increase of complexity during wine aging.

Olive pomace was qualitatively analyzed by HPLC-DAD-ESI^- TOF MS/MS after ASE, giving rise to the putative identification of about 50 compounds, including those that are commonly found in olive oil extract and some of their glycosides (Peralbo-Molina, Priego-Capote, & Luque de Castro, 2012). A very complete and complex investigation of all the matrices involved in olive oil processing, including pomace and waste water, was performed by UHPLC-DAD-positive/negative ESI-TOF MS/MS. More than 80 phenolics were tentatively identified, being some glycosylated compounds detected in both pomace and waste water (Jerman Klen, Golc Wondra, Vrhovšek, & Mozetič Vodopivec, 2015). However, structures should be confirmed, especially for newly identified compounds.

A method developed for the identification of phenolics in apple and cider (Ramirez-Ambrosi et al., 2013) based on UHPLC-DAD-positive/negative-ESI-TOF pseudo-MS^3 (in source fragmentation is not equivalent to MS/MS) was also applied to pomace of the same apples, producing the same identification with exception of two compounds: 51 compounds were tentatively identified, with the limitation that the compounds for which standards were unavailable should be confirmed.

Phenolic compounds are highly present in byproducts from many other vegetable and fruits and are thus a valuable source for valorization toward phenolic-rich extracts; among them, *Brassicaceae* species are the most studied. The cauliflower (*Brassica oleracea* L. *var. botrytis*) leaves are a rich source of phenolics, however a considerable amount is present in a bound, non-extractable form. Huynh, Smagghe, Gonzales, Van Camp, and Raes (2014) studied how to release phenolic compounds from outer leaves of cauliflower using two commercially available polysaccharide-degrading enzymes, Viscozyme L cellulolytic enzyme mixture and Rapidase vegetable juice, while in another study non-extractable phenolics (NEP) were released from the matrix by alkaline hydrolysis followed by sonication (Gonzales, Smagghe, Raes, & Van Camp, 2014). Both studies used the same method, which consisted in sample clean-up by C18 SPE followed by UHPLC-DAD-positive/negative-ESI-TOF MS/MS. In the same conditions 13 compounds, mainly kaempferol conjugates, were identified and quantitatively estimated. Hydrolysis systematically gave larger amounts of phenolics than the non-hydrolyzed samples. In another study by the same group (Gonzales, Raes et al., 2014), after methanolic extraction under homogenization, the extract was subdivided in three aliquots: one unchanged, whereas the other two were subjected to acid and alkaline hydrolysis, respectively. Then, the aliquots were cleaned up by C18 SPE. Samples were analyzed by UHPLC-DAD-positive/negative-ESI-TOF-ion mobility (IM)-MS/MS. IM was added to separate positional isomers, while other conditions were the same. Nine standards, 5 of which were glycosylated, were used to study fragmentation. Hydrolysis was used for structure confirmation, because acidic hydrolysis cleaves the glycosidic bond, while alkaline hydrolysis cleaves the ester bond. A total of 24 compounds was identified, including 5 suspected kaempferol-O-C glycosides. The same approach of Gonzales, Smagghe et al. (2014) was applied to red cabbage and Brussels sprout wastes (Gonzales et al., 2015). The conditions of alkaline hydrolysis were further improved, while the instrumental setup was maintained. A total of 51 chromatographic peaks were tentatively identified, most of them being

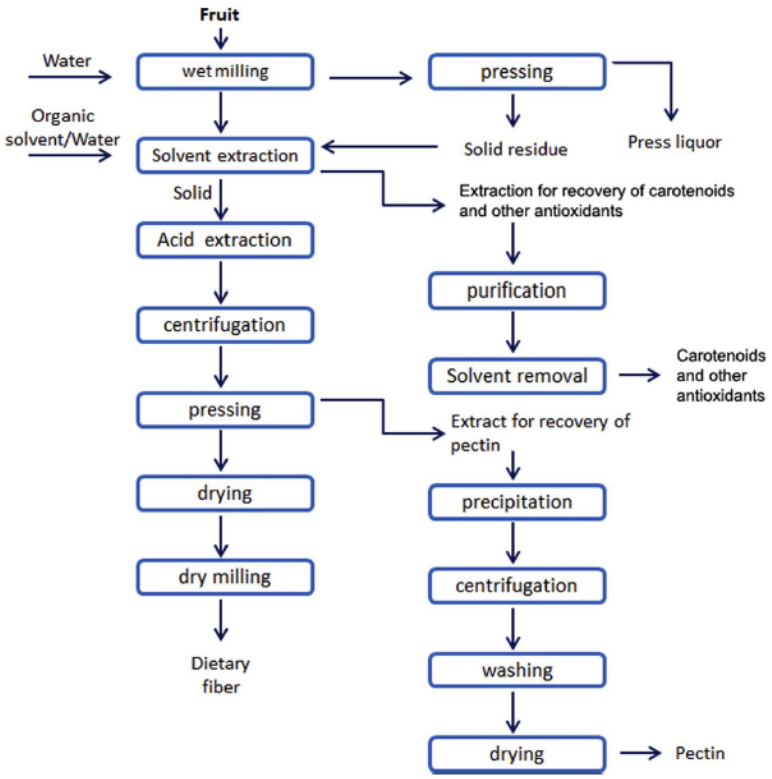


Fig. 3. Schematic diagram for fruit pomace handling procedure. The diagram illustrates some procedures for recovery of valuable compounds form fruit wastes. (Adapted from reference (Djilas et al., 2009).)

flavonoid glycosides and acylated flavonoid glycosides. However, some phenolic acid derivatives, glucosinolates and anthocyanins were also identified.

Finally, three more byproduct samples, such as mango, broad bean and black carrot waste, were characterized for bioactive phytochemical composition. Microwave assisted extraction was performed on mango byproducts (peel and seed) after lyophilization. The extracted phenolic compounds were analyzed by HPLC-ESI-Q-TOF MS and 30 compounds were separated and tentatively identified based on their accurate mass, isotopic abundances and fragmentation pattern; these compounds were differently distributed among the three different varieties of mango (Dorta, González, Lobo, Sánchez-Moreno, & de Ancos, 2014). Fresh whole pods of large bean (*Vicia faba*) were lyophilized, then extracted under sonication with a MeOH/water mixture; the analysis was performed by UHPLC-negative ESI-Q-TOF MS/MS (Abu-Reidah et al., 2014). An untargeted metabolomic approach for polar and semi-polar compounds was used and 142 compounds were tentatively identified or at least characterized (only 9 standards were available), including phenolics, organic acids, amino acid and peptide derivatives, and terpenoids. The black carrot waste characterization problem was tackled from two points of view: polyphenolic content and bioaccessibility (Kamiloglu et al., 2016). Extracts were analyzed by the same instrumental setup as previously described, in addition an in vitro gastrointestinal digestion was simulated to determine the fraction available for intestinal absorption. The aim was not the identification of a large number of compounds, as the quantitative determination of the most relevant compounds was much more important. It was found that, after simulated digestion, the amount of bioavailable phenolics was higher for pomace and peel than for commercial black carrots derived products, such as juice and fermented beverages. This suggests that unprocessed raw black carrot waste, such as pomace and peel, should be used to enrich food products. However, it is difficult to imagine its addition to any kind of food products without any further processing.

In conclusion, HRMS has been scarcely exploited in this field. Probably, a detailed investigation might be considered not necessary, but this is a questionable opinion. In fact, the authenticity of functionalized food might at least be proved. In addition, a metabolomic approach can surely extensively characterize any agro-industrial waste, suggesting the more appropriate processes to reutilize such waste.

5.5. *Metabolomics: present and future*

Targeted metabolomics is a well-established methodology focused on the analysis of known compounds related to one or more metabolic pathways. The purpose of this approach is to discover specific modifications, both qualitative and quantitative, in metabolites that may be related to a particular condition, such as different genotypes, different origin, cultural practices etc. By analyzing a sufficient number of samples with the aid of simple chemometric tools, such as the principal component analysis, targeted metabolomics enables to discriminate and classify samples into groups. For this approach, HRMS is not strictly necessary, but the connection with UHPLC enables high-throughput analysis avoiding false positives.

Untargeted metabolomics produces datasets that should cover, theoretically, the whole metabolome of a biological sample. In practice, however, the huge amount of metabolites in the plant kingdom and the great diversity of chemical and physical characteristics of such compounds make, at present, impossible to acquire the entire metabolome in a single experiment, despite the progresses in both chromatography and MS. Metabolic profiles are restricted to a single pathway or class of

compounds. Metabolic profiling of phytochemicals in food of vegetal origin is not as much developed as in other fields, including medicinal plants. This gap can be expected to be bridged within the next few years. Suspect screening analysis is a mid-way approach, in which the identification of compounds relies on specific available information, such as their molecular formula and isotopic pattern. Thus, specific databases should be available and providing confident identification can be upgraded over time.

What is the state-of-the-art in HRMS-based plant metabolomics? HRMS alone, without any fractionation step, has been employed sporadically using Direct Analysis in Real Time (DART), TOF-MS (Kim, Park, & Jang, 2014; Prchalová, Kovařík, Ševčík, Čížková, & Rajchl, 2014) and in vivo nanoESI/Q-Orbitrap MS (Chang, Peng, Dan, Shuai, & Hu, 2015). Direct HRMS is a very fast, high throughput technique; however, ME severely affects the ionization efficiency. Therefore, only few most abundant compounds can be detected and quantitation is almost impossible; only fingerprints of a sample for fast classification can be possible. In this context, it is important to devise new, more efficient ion sources to enable developments in this field.

As the objective of untargeted metabolomics is to obtain accurate, detailed metabolic fingerprints, the role of the selectivity of the combined LC-MS system is determinant in order to maximize the number of identified metabolites (Ortmayr, Causon, Hann, & Koellensperger, 2016). Currently, most of MS setups employed in metabolomics are based on DDA. However, this option is limited by the duty cycle necessary to acquire MS/MS spectra (Wolfender et al., 2015). In this contest, the two steps cannot be optimized separately, and a compromise is necessary to obtain the maximum possible number of MS/MS features. By the way, it is possible to substantially increase the number of putatively identified compounds still maintaining accurate and reproducible peak areas (La Barbera et al., 2017). This approach is relatively simple and it could be used in the future.

Data independent acquisition (DIA) mode is not limited to the detection of a restricted number of MS/MS features of selected (for example most abundant) precursor ions in the spectrum. Without a previous selection, all the precursors are fragmented at any time of the chromatographic separation.

This approach is much more complex but it has the advantage of allowing a more comprehensive dataset and retrospective mining. Although some different strategies have been developed (Wolfender et al., 2015), it has never been applied to food phytochemicals. However, considering the inherent advantages, future developments could be expected.

Comprehensive two-dimension LC combining orthogonal separation techniques appears very attractive when both polar and non-polar compounds have to be analyzed contemporaneously. However, it was rarely used (see Section 3.3), mainly due to the fact that the two most suitable techniques for coupling, namely RPLC and HILIC, have mobile phase compatibility issues, in addition to longer equilibration times for HILIC than for RPLC (Ortmayr et al., 2016). Some combinations RP × RP have been sporadically studied, however the extent of orthogonality is poor. For the future, owing the inherent resolving power, improvements in both column and instrumentation technologies could facilitate the diffusion of this technique.

IM-MS is not a new MS technique (Mason & Schamp, 1958); however, only recently it was implemented into a very performing HRMS instrument. IM can be regarded as an orthogonal dimension to both LC and MS, since it separates ions based on the difference in shape; in this way, it is possible to separate isomeric compounds with the same m/z, which otherwise cannot be separated by chromatography alone. For basic principles of time dispersive IM-MS, state-of-the-art instrumentation and some applications, the reader may refer to other reviews (May & McLean, 2015; Ortmayr et al., 2016). One obstacle to a larger diffusion of IM-HRMS could be the instrumentation cost. A second obstacle is provided by very complex datasets which, though very informative, are complicated as IM does separate also different adducts of the same compound; this may create difficulties in identification which arise when adducts, in which the charge can be localized in more than one position, form and when their interconversion is slow compared to the drift time. When these two shortcomings will be overcome, the inclusion of IM-MS in the workflow will be helpful for elucidation of isomeric compounds, such as isomeric glycosides.

5.6. Compound identification challenges

Compound identification is considered the bottleneck of untargeted metabolite profiling. The process of identification is complex and dependent on the analytical platform, as well as the robustness of the method and databases. Therefore, confidence of identification can vary widely. The first recommended minimum reporting standards for chemical analysis in metabolomics were established by the Chemical Analysis Working Group (CAWG) in 2007 (Sumner et al., 2007). Four levels of identification confidence were established:

level 1: identified compounds (using authentic standards);
level 2: putatively annotated compounds (verifying MS and MS/MS spectra);
level 3: putatively annotated compound class (deciphering MS and MS/MS spectra);
level 4: unidentified or unclassified metabolites (which can be differentiated based on spectral data).

Moreover, authors should clearly differentiate and state the level of identification rigor for all reported metabolites. This classification can be considered still valid although, recently, a group of specialists invited the metabolomics community to debate the assumption of more articulated standards (Creek et al., 2014). In practice, this recommendation is often disregarded in the field that is the subject of this review. Only mass accuracy and MS/MS fragments are reported, or reference to other previous works (the confidence of the original identification remaining unknown) or to the database containing the similar MS/MS spectrum, without score. It is desirable that in the near future all researchers using a metabolomic approach will use the same criteria.

The first step of HRMS data analysis is the selection of ions of interest from the many signals in the spectrum, a process which involves feature detection (defined as pairs consisting of the retention time and the m/z value), alignment and normalization. The result is a dataset (i.e., a list of ions with their m/z values, abundances and retention times). Several computational tools can be used for this purpose (Misra & van der Hooft, 2016; Rathahao-Paris et al., 2016).

To reduce the number of possible molecular formulas, different heuristic filters can be applied including restrictions on the number of elements and the elemental ratios of nitrogen, oxygen, phosphorus and sulfur vs carbon. However, this strategy typically leads to multiple putative identities for each detected compound. Fragmentation patterns are highly informative and can be readily compared to the literature data or dedicated databases, typically containing few hundreds of MS/MS spectra. Proper usage and development of MS-based spectral databases is essential for metabolomics to reach the status of other omic sciences. In this sense, there were many challenges for software developers, but one of the most fundamental is the growing accessibility of plant metabolite databases, since current databases are still not comprehensive of all known metabolites, due to the relatively small number of metabolites commercially available as pure standards, not to mention the large number of metabolites with unknown chemical structures that remain to be identified and characterized. As far as phytochemicals is concerned, a suitable database mainly devoted to plant metabolomics is ReSpect (Sawada et al., 2012), which is a curated downloadable public database, which, however, contains at least 50% data from QqQ instrumentation. For an updated discussion on databases in metabolomics, please refer to (Gil de la Fuente, Grace

Armitage, Otero, Barbas, & Godzien, 2017). However, this approach might be invalidated because MS/MS spectra are instrument-dependent. Other than in GC–MS techniques, the use of databases in LC-MS is hindered mainly due to the lack of standardization of the processes of ionization (ESI, APCI) compared to the EI (electron impact ionization) analysis, resulting in significant differences in the spectra of the same compounds acquired using different ionization conditions (Kachlicki, Piasecka, Stobiecki, & Marczak, 2016). Despite these limitations, the use of reference spectral databases is still one of the best approaches to annotate the structure of known metabolites when full isolation and structure determination by NMR or X-ray crystallography is not possible. In other cases, accurate databases are built by the users based on their own set of available standards and instruments. Alternatively, novel computational tools that heuristically predict MS fragmentation patterns in silico have been developed to assist with identification of metabolites for which tandem MS data are not available, yet.

Continuous improvements in data (pre)processing software (Rathahao-Paris et al., 2016) and MS/MS database (Nakabayashi & Saito, 2013) make the untargeted metabolomics approach more and more affordable, therefore, an increasing application of such an approach in food phytochemistry is also expected.

6. Conclusions

For a long time, the chemical analysis employing HPLC coupled to low resolution MS^n has been widely used for identification of phytochemicals. However, the uncertainty of the technique required stringent confirmation with pure standard or high informative complementary techniques, such as NMR.

More recently, the tremendous technological progress in both HPLC and MS greatly improved the probability of correct identification of a compound, based on MS data alone. However, there are still difficulties for mass analyzers to fully support the separation capability of UHPLC, as the required chromatographic and mass spectrometric resolutions essentially move in opposite directions. Although providing high resolution, TOF and FT instrumentations have very different designs. Typically, the highest resolution available in TOF instruments is several times lower than the ultimate resolution in the high field Orbitrap. However, in practice, TOF is used at the maximum resolution of 40,000 and Orbitrap is used at the same resolution to reach the same scan speed. Mass accuracy in MS mode is comparable for the two instruments, although Orbitrap is more stable and with a wider dynamic range. In MS/MS mode Q-TOF offers better accuracy but lower sensitivity, thus the actual reported true positive identification rate achieved with Orbitrap mass analyzers is probably the same as TOF. In the end, both latest generation TOF and Orbitrap systems basically provide a similar performance and both are suitable for plant phytochemical analysis. However, both technologies are still evolving: as an example, the latest commercialized Orbitrap Fusion, to our knowledge, has never been used in this field.

Two different approaches with different levels of accomplishable information could be carried out: target analysis and untargeted analysis. In the first approach, a very limited number of target compounds is identified and/or quantified, and the comparison with standards is a straightforward operation. When a large number of unknown compounds should be identified, the main strategy for automatic identification or peak annotation could be based on database search. In these searches, MS, HRMS, MS/MS spectra are compared to a large number of existing web databases. Chemical formula assignments can be obtained by matching the experimental accurate mass with the theoretical one and comparing the observed isotopic distributions. Access to ultra-HRMS is sometimes necessary to avoid false positive interpretations. Even when mass accuracies < 1 ppm are obtained, a definitive molecular formula determination remains a difficult task, particularly for compounds with MW > 500 Da. Fragmentation patterns are highly informative and can be readily compared to the literature data or dedicated databases, but this is currently hindered by the dependence of MS/MS spectra on the employed instrument and lack of real bioinformatics tools, as developed for other fields, which allow automatic compound identification without extensive manual check; thus, practically databases need to be built by the users based on their own set of available standards.

Finally, from the examination of the literature, it can be concluded that, with respect to medicinal plants and diet integrators, phytochemical analysis in food of plant origin has been much lesser exploited. Only phenolics in some fruits and plant-derived beverages were quite extensively studied.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Abrankó, L., Nagy, Á., Szilvássy, B., Stefanovits-Bányai, É., & Hegedus, A. (2015). Genistein isoflavone glycoconjugates in sour cherry (*Prunus cerasus* L.) cultivars. *Food Chemistry, 166*, 215–222. http://dx.doi.org/10.1016/j.foodchem.2014.06.007.

Abu-Reidah, I. M., Arráez-Román, D., Quirantes-Piné, R., Fernández-Arroyo, S., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2012). HPLC-ESI-Q-TOF-MS for a comprehensive characterization of bioactive phenolic compounds in cucumber whole fruit extract. *Food Research International, 46*(1), 108–117. http://dx.doi.org/10.1016/j.foodres.2011.11.026.

Abu-Reidah, I. M., Contreras, M.d. M., Arráez-Román, D., Fernández-Gutiérrez, A., & Segura-Carretero, A. (2014). UHPLC-ESI-QTOF-MS-based metabolic profiling of *Vicia faba* L. (*Fabaceae*) seeds as a key strategy for characterization in foodomics. *Electrophoresis, 35*(11), 1571–1581. http://dx.doi.org/10.1002/elps.201300646.

Abu-Reidah, I. M., Contreras, M. D. M., Arráez-Román, D., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2013). Reversed-phase ultra-high-performance liquid chromatography coupled to electrospray ionization-quadrupole-time-of-flight mass spectrometry as a powerful tool for metabolic profiling of vegetables: *Lactuca sativa* as an example of its application. *Journal of Chromatography A, 1313*, 212–227. http://dx.doi.org/10.1016/j.chroma.2013.07.020.

Alberts, P., Stander, M. A., & De Villiers, A. (2012). Advanced ultra high pressure liquid chromatography-tandem mass spectrometric methods for the screening of red wine anthocyanins and derived pigments. *Journal of Chromatography A, 1235*, 92–102. http://dx.doi.org/10.1016/j.chroma.2012.02.058.

Allouche, Y., Warleta, F., Campos, M., Sánchez-Quesada, C., Uceda, M., Beltrán, G., & Gaforio, J. J. (2011). Antioxidant, antiproliferative, and pro-apoptotic capacities of pentacyclic triterpenes found in the skin of olives on MCF-7 human breast cancer cells and their effects on DNA damage. *Journal of Agricultural and Food Chemistry, 59*(1), 121–130. http://dx.doi.org/10.1021/jf102319y.

Álvarez-Fernández, M. A., Cerezo, A. B., Cañete-Rodríguez, A. M., Troncoso, A. M., & García-Parrilla, M. C. (2015). Composition of nonanthocyanin polyphenols in alcoholic-fermented strawberry products using LC-MS (QTRAP), high-resolution MS (UHPLC-Orbitrap-MS), LC-DAD, and antioxidant activity. *Journal of Agricultural and Food Chemistry, 63*(7), 2041–2051. http://dx.doi.org/10.1021/jf506076n.

Antunes-Ricardo, M., Moreno-García, B. E., Gutiérrez-Uribe, J. A., Aráiz-Hernández, D., Alvarez, M. M., & Serna-Saldivar, S. O. (2014). Induction of apoptosis in colon cancer cells treated with isorhamnetin glycosides from *Opuntia ficus-indica* pads. *Plant Foods for Human Nutrition, 69*(4), 331–336. http://dx.doi.org/10.1007/s11130-014-0438-5.

Ares, A. M., Bernal, J., Nozal, M. J., Turner, C., & Plaza, M. (2015). Fast determination of intact glucosinolates in broccoli leaf by pressurized liquid extraction and ultra high performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry. *Food Research International, 76*, 498–505. http://dx.doi.org/10.1016/j.foodres.2015.06.037.

Ares, A. M., Nozal, M. J., & Bernal, J. (2013). Extraction, chemical characterization and biological activity determination of broccoli health promoting compounds. *Journal of Chromatography A, 1313*, 78–95. http://dx.doi.org/10.1016/j.chroma.2013.07.051.

Ares, A. M., Nozal, M. J., Bernal, J. L., & Bernal, J. (2014). Optimized extraction, separation and quantification of twelve intact glucosinolates in broccoli leaves. *Food Chemistry, 152*, 66–74. http://dx.doi.org/10.1016/j.foodchem.2013.11.125.

Armenta, S., & de la Guardia, M. (2015). Green chromatography for the analysis of animal origin foods. *TrAC Trends in Analytical Chemistry, 80*, 517–530. http://dx.doi.org/10.1016/j.trac.2015.06.012.

Bai, Y., Findlay, B., Sanchez Maldonado, A. F., Schieber, A., Vederas, J. C., & Gänzle, M. G. (2014). Novel pyrano and vinylphenol adducts of deoxyanthocyanidins in sorghum sourdough. *Journal of Agricultural and Food Chemistry, 62*(47), 11536–11546. http://dx.doi.org/10.1021/jf503330b.

Ballus, C. A., Quirantes-Piné, R., Bakhouche, A., Da Silva, L. F. D. O., De Oliveira, A. F., Coutinho, E. F., ... Godoy, H. T. (2015). Profile of phenolic compounds of Brazilian virgin olive oils by rapid resolution liquid chromatography coupled to electrospray ionisation time-of-flight mass spectrometry (RRLC-ESI-TOF-MS). *Food Chemistry, 170*, 366–377. http://dx.doi.org/10.1016/j.foodchem.2014.08.054.

Bateman, K. P., Kellmann, M., Muenster, H., Papp, R., & Taylor, L. (2009). Quantitative-qualitative data acquisition using a benchtop orbitrap mass spectrometer. *Journal of the American Society for Mass Spectrometry, 20*(8), 1441–1450. http://dx.doi.org/10.1016/j.jasms.2009.03.002.

Bell, L., & Wagstaff, C. (2014). Glucosinolates, myrosinase hydrolysis products, and

flavonols found in rocket (*Eruca sativa* and *Diplotaxis tenuifolia*). *Journal of Agricultural and Food Chemistry*. http://dx.doi.org/10.1021/jf501096x.
Bhattacharya, U., Halder, B., Mukhopadhyay, S., & Giri, A. K. (2009). Role of oxidation-triggered activation of JNK and p38 MAPK in black tea polyphenols induced apoptotic death of A375 cells. *Cancer Science*, *100*(10), 1971–1978. http://dx.doi.org/10.1111/j.1349-7006.2009.01251.x.
Bianco, G., Buchicchio, A., & Cataldi, T. R. I. (2015). Structural characterization of major soyasaponins in traditional cultivars of Fagioli di Sarconi beans investigated by high-resolution tandem mass spectrometry. *Analytical and Bioanalytical Chemistry*, *407*(21), 6381–6389. http://dx.doi.org/10.1007/s00216-015-8810-3.
Bijttebier, S., D'Hondt, E., Noten, B., Hermans, N., Apers, S., & Voorspoels, S. (2014). Ultra high performance liquid chromatography versus high performance liquid chromatography: Stationary phase selectivity for generic carotenoid screening. *Journal of Chromatography A*, *1332*, 46–56. http://dx.doi.org/10.1016/j.chroma.2014.01.042.
Bijttebier, S., Van der Auwera, A., Foubert, K., Voorspoels, S., Pieters, L., & Apers, S. (2016). Bridging the gap between comprehensive extraction protocols in plant metabolomics studies and method validation. *Analytica Chimica Acta*, *935*, 136–150. http://dx.doi.org/10.1016/j.aca.2016.06.047.
Bijttebier, S., Zhani, K., D'Hondt, E., Noten, B., Hermans, N., Apers, S., & Voorspoels, S. (2014). Generic characterization of apolar metabolites in red chili peppers (*Capsicum frutescens* L.) by orbitrap mass spectrometry. *Journal of Agricultural and Food Chemistry*, *62*(20), 4812–4831. http://dx.doi.org/10.1021/jf500285g.
Blanch, M., Alvarez, I., Sanchez-Ballesta, M. T., Escribano, M. I., & Merodio, C. (2012). Increasing catechin and procyanindin accumulation in high-CO_2-treated *Fragaria vesca* strawberries. *Journal of Agricultural and Food Chemistry*, *60*(30), 7489–7496. http://dx.doi.org/10.1021/jf301547t.
Caprioli, G., Cahill, M. G., Vittori, S., & James, K. J. (2014). Liquid chromatography-hybrid linear ion trap-high-resolution mass spectrometry (LTQ-Orbitrap) method for the determination of glycoalkaloids and their aglycons in potato samples. *Food Analytical Methods*, *7*(7), 1367–1372. http://dx.doi.org/10.1007/s12161-013-9758-6.
Capriotti, A. L., Caruso, G., Cavaliere, C., Foglia, P., Laganà, A., & Samperi, R. (2012). Polyphenol content in white table grape (*Vitis vinifera*) berries of cultivar Italia: Interactive effect of irrigation, delayed harvest and storage. *Natural Product Research*, *26*(19), 1787–1795. http://dx.doi.org/10.1080/14786419.2011.606623.
Capriotti, A. L., Cavaliere, C., Crescenzi, C., Foglia, P., Nescatelli, R., Samperi, R., & Laganà, A. (2014). Comparison of extraction methods for the identification and quantification of polyphenols in virgin olive oil by ultra-HPLC-QToF mass spectrometry. *Food Chemistry*, *158*, 392–400. http://dx.doi.org/10.1016/j.foodchem.2014.02.130.
Carrasco-Pancorbo, A., Neusüß, C., Pelsing, M., Segura-Carretero, A., & Fernandez Gutiérrez, A. (2007). CE- and HPLC-TOF-MS for the characterization of phenolic compounds in olive oil. *Electrophoresis*, *28*(5), 806–821. http://dx.doi.org/10.1002/elps.200600382.
Cavaliere, C., Foglia, P., Gubbiotti, R., Sacchetti, P., Samperi, R., & Laganà, A. (2008). Rapid-resolution liquid chromatography/mass spectrometry for determination and quantitation of polyphenols in grape berries. *Rapid Communications in Mass Spectrometry*, *22*(20), 3089–3099. http://dx.doi.org/10.1002/rcm.3705.
Cavaliere, C., Foglia, P., Pastorini, E., Samperi, R., & Laganà, A. (2005). Identification and mass spectrometric characterization of glycosylated flavonoids in *Triticum durum* plants by high-performance liquid chromatography with tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *19*(21), 3143–3158. http://dx.doi.org/10.1002/rcm.2185.
Chang, Q., Peng, Y., Dan, C., Shuai, Q., & Hu, S. (2015). Rapid in situ identification of bioactive compounds in plants by in vivo nanospray high-resolution mass spectrometry. *Journal of Agricultural and Food Chemistry*, *63*(11), 2911–2918. http://dx.doi.org/10.1021/jf505749n.
Chang, Y., Zhao, C., Wu, Z., Zhou, J., Zhao, S., Lu, X., & Xu, G. (2012). Chip-based nanoflow high performance liquid chromatography coupled to mass spectrometry for profiling of soybean flavonoids. *Electrophoresis*, *33*(15), 2399–2406. http://dx.doi.org/10.1002/elps.201100581.
Chapman, J. D., Goodlett, D. R., & Masselon, C. D. (2014). Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrometry Reviews*, *33*(6), 452–470. http://dx.doi.org/10.1002/mas.21400.
Cherng, J.-M., Shieh, D.-E., Chiang, W., Chang, M.-Y., & Chiang, L.-C. (2007). Chemopreventive effects of minor dietary constituents in common foods on human cancer cells. *Bioscience, Biotechnology, and Biochemistry*, *71*(6), 1500–1504. http://dx.doi.org/10.1271/bbb.70008.
Chong, E. S. L., Mcghie, T. K., Heyes, J. A., & Stowell, K. M. (2013). Metabolite profiling and quantification of phytochemicals in potato extracts using ultra-high-performance liquid chromatography-mass spectrometry. *Journal of the Science of Food and Agriculture*, *93*(15), 3801–3808. http://dx.doi.org/10.1002/jsfa.6285.
Cichon, M. J., Riedl, K. M., & Schwartz, S. J. (2017). A metabolomic evaluation of the phytochemical composition of tomato juices being used in human clinical trials. *Food Chemistry*, *228*, 270–278. http://dx.doi.org/10.1016/j.foodchem.2017.01.118.
Contreras, M. D. M., Arráez-Román, D., Fernández-Gutiérrez, A., & Segura-Carretero, A. (2015). Nano-liquid chromatography coupled to time-of-flight mass spectrometry for phenolic profiling: A case study in cranberry syrups. *Talanta*, *132*, 929–938. http://dx.doi.org/10.1016/j.talanta.2014.10.049.
Creek, D. J., Dunn, W. B., Fiehn, O., Griffin, J. L., Hall, R. D., Lei, Z., ... Wolfender, J.-L. (2014). Metabolite identification: Are you sure? And how do your peers gauge your confidence? *Metabolomics*, *10*(3), 350–353. http://dx.doi.org/10.1007/s11306-014-0656-8.
De Paepe, D., Servaes, K., Noten, B., Diels, L., De Loose, M., Van Droogenbroeck, B., & Voorspoels, S. (2013). An improved mass spectrometric method for identification and quantification of phenolic compounds in apple fruits. *Food Chemistry*, *136*(2), 368–375. http://dx.doi.org/10.1016/j.foodchem.2012.08.062.
De Rosso, M., Tonidandel, L., Larcher, R., Nicolini, G., Dalla Vedova, A., De Marchi, F., ... Flamini, R. (2014). Identification of new flavonols in hybrid grapes by combined liquid chromatography-mass spectrometry approaches. *Food Chemistry*, *163*, 244–251. http://dx.doi.org/10.1016/j.foodchem.2014.04.110.
De Villiers, A., Cabooter, D., Lynen, F., Desmet, G., & Sandra, P. (2011). High-efficiency high performance liquid chromatographic analysis of red wine anthocyanins. *Journal of Chromatography A*, *1218*(29), 4660–4670. http://dx.doi.org/10.1016/j.chroma.2011.05.042.
Di Donna, L., Mazzotti, F., Napoli, A., Salerno, R., Sajjad, A., & Sindona, G. (2007). Secondary metabolism of olive secoiridoids. New microcomponents detected in drupes by electrospray ionization and high-resolution tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *21*(3), 273–278. http://dx.doi.org/10.1002/rcm.2830.
Di Donna, L., Mazzotti, F., Taverna, D., Napoli, A., & Sindona, G. (2014). Structural characterisation of malonyl flavonols in leek (*Allium porrum* L.) using high-performance liquid chromatography and mass spectrometry. *Phytochemical Analysis*, *25*(3), 207–212. http://dx.doi.org/10.1002/pca.2493.
Di Lecce, G., Arranz, S., Jáuregui, O., Tresserra-Rimbau, A., Quifer-Rada, P., & Lamuela-Raventós, R. M. (2014). Phenolic profiling of the skin, pulp and seeds of Albariño grapes using hybrid quadrupole time-of-flight and triple-quadrupole mass spectrometry. *Food Chemistry*, *145*, 874–882. http://dx.doi.org/10.1016/j.foodchem.2013.08.115.
Dillard, C. J., & Bruce German, J. (2000). Phytochemicals: Nutraceuticals and human health. *Journal of the Science of Food and Agriculture*, *80*, 1744–1756. http://dx.doi.org/10.1002/1097-0010(20000915)80:12<1744::AID-JSFA725>3.0.CO;2-W.
Dinelli, G., Carretero, A. S., Di Silvestro, R., Marotti, I., Fu, S., Benedettelli, S., ... Gutiérrez, A. F. (2009). Determination of phenolic compounds in modern and old varieties of durum wheat using liquid chromatography coupled with time-of-flight mass spectrometry. *Journal of Chromatography A*, *1216*(43), 7229–7240. http://dx.doi.org/10.1016/j.chroma.2009.08.041.
Dinelli, G., Segura-Carretero, A., Di Silvestro, R., Marotti, I., Arráez-Román, D., Benedettelli, S., ... Fernadez-Gutierrez, A. (2011). Profiles of phenolic compounds in modern and old common wheat varieties determined by liquid chromatography coupled with time-of-flight mass spectrometry. *Journal of Chromatography A*, *1218*(42), 7670–7681. http://dx.doi.org/10.1016/j.chroma.2011.05.065.
Djilas, S., Čanadanović-Brunet, J., & Ćetković, G. (2009). By-products of fruit processing as a source of phytochemicals. *Chemical Industry and Chemical Engineering Quarterly*, *15*(4), 191–202. http://dx.doi.org/10.2298/CICEQ0904191D.
Dorta, E., González, M., Lobo, M. G., Sánchez-Moreno, C., & de Ancos, B. (2014). Screening of phenolic compounds in by-product extracts from mangoes (*Mangifera indica* L.) by HPLC-ESI-QTOF-MS and multivariate analysis for use as a food ingredient. *Food Research International*, *57*, 51–60. http://dx.doi.org/10.1016/j.foodres.2014.01.012.
Dugo, P., Giuffrida, D., Herrero, M., Donato, P., & Mondello, L. (2009). Epoxycarotenoids esters analysis in intact orange juices using two-dimensional comprehensive liquid chromatography. *Journal of Separation Science*, *32*(7), 973–980. http://dx.doi.org/10.1002/jssc.200800696.
Farag, M. A., El-Ahmady, S. H., Elian, F. S., & Wessjohann, L. A. (2013). Metabolomics driven analysis of artichoke leaf and its commercial products via UHPLC-q-TOF-MS and chemometrics. *Phytochemistry*, *95*, 177–187. http://dx.doi.org/10.1016/j.phytochem.2013.07.003.
Feuereisen, M. M., Gamero Barraza, M., Zimmermann, B. F., Schieber, A., & Schulze-Kaysers, N. (2017). Pressurized liquid extraction of anthocyanins and biflavonoids from *Schinus terebinthifolius Raddi*: A multivariate optimization. *Food Chemistry*, *214*, 564–571. http://dx.doi.org/10.1016/j.foodchem.2016.07.002.
Fiehn, O. (2002). Metabolomics - The link between genotypes and phenotypes. *Plant Molecular Biology*, *48*(1–2), 155–171. http://dx.doi.org/10.1023/A:1013713905833.
Flamini, R., De Rosso, M., De Marchi, F., Dalla Vedova, A., Panighel, A., Gardiman, M., ... Bavaresco, L. (2013). An innovative approach to grape metabolomics: Stilbene profiling by suspect screening analysis. *Metabolomics*, *9*(6), 1243–1253. http://dx.doi.org/10.1007/s11306-013-0530-0.
Fontana, A. R., Antoniolli, A., & Bottini, R. (2013). Grape pomace as a sustainable source of bioactive compounds: Extraction, characterization, and biotechnological applications of phenolics. *Journal of Agricultural and Food Chemistry*, *61*(38), 8987–9003. http://dx.doi.org/10.1021/jf402586f.
Forcisi, S., Moritz, F., Kanawati, B., Tziotis, D., Lehmann, R., & Schmitt-Kopplin, P. (2013). Liquid chromatography-mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling. *Journal of Chromatography A*, *1292*, 51–65. http://dx.doi.org/10.1016/j.chroma.2013.04.017.
Fu, S., Segura-Carretero, A., Arráez-Román, D., Menéndez, J. A., De La Torre, A., & Fernández-Gutiérrez, A. (2009). Tentative characterization of novel phenolic compounds in extra virgin olive oils by rapid-resolution liquid chromatography coupled with mass spectrometry. *Journal of Agricultural and Food Chemistry*, *57*(23), 11140–11147. http://dx.doi.org/10.1021/jf901590n.
García-Villalba, R., Carrasco-Pancorbo, A., Vázquez-Martín, A., Oliveras-Ferraros, C., Menéndez, J. A., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2009). A 2-D-HPLC-CE platform coupled to ESI-TOF-MS to characterize the phenolic fraction in olive oil. *Electrophoresis*, *30*(15), 2688–2701. http://dx.doi.org/10.1002/elps.200800807.
García-Villalba, R., Carrasco-Pancorbo, A., Zurek, G., Behrens, M., Bäßmann, C., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2010). Nano and rapid resolution liquid chromatography-electrospray ionization-time of flight mass spectrometry to identify and quantify phenolic compounds in olive oil. *Journal of Separation Science*, *33*(14), 2069–2078. http://dx.doi.org/10.1002/jssc.201000184.

Gebhardt, R. (1997). Antioxidative and protective properties of extracts from leaves of the artichoke (*Cynara scolymus* L.) against hydroperoxide-induced oxidative stress in cultured rat hepatocytes. *Toxicology and Applied Pharmacology*, *144*(2), 279–286. http://dx.doi.org/10.1006/taap.1997.8130.

Geiger, T., Cox, J., & Mann, M. (2010). Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Molecular & Cellular Proteomics*, *9*(10), 2252–2261. http://dx.doi.org/10.1074/mcp.M110.001537.

Gil de la Fuente, A., Grace Armitage, E., Otero, A., Barbas, C., & Godzien, J. (2017). Differentiating signals to make biological sense - A guide through databases for MS-based non-targeted metabolomics. *Electrophoresis*, *0*, 1–16. http://dx.doi.org/10.1002/elps.201700070.

Glauser, G., Veyrat, N., Rochat, B., Wolfender, J. L., & Turlings, T. C. J. (2013). Ultra-high pressure liquid chromatography-mass spectrometry for plant metabolomics: A systematic comparison of high-resolution quadrupole-time-of-flight and single stage Orbitrap mass spectrometers. *Journal of Chromatography A*, *1292*, 151–159. http://dx.doi.org/10.1016/j.chroma.2012.12.009.

Gómez-Ramos, M. M., Ferrer, C., Malato, O., Agüera, A., & Fernández-Alba, A. R. (2013). Liquid chromatography-high-resolution mass spectrometry for pesticide residue analysis in fruit and vegetables: Screening and quantitative studies. *Journal of Chromatography A*. http://dx.doi.org/10.1016/j.chroma.2013.02.065.

Gonzales, G. B., Raes, K., Coelus, S., Struijs, K., Smagghe, G., & Van Camp, J. (2014). Ultra (high)-pressure liquid chromatography-electrospray ionization-time-of-flight-ion mobility-high definition mass spectrometry for the rapid identification and structural characterization of flavonoid glycosides from cauliflower waste. *Journal of Chromatography A*, *1323*, 39–48. http://dx.doi.org/10.1016/j.chroma.2013.10.077.

Gonzales, G. B., Raes, K., Vanhoutte, H., Coelus, S., Smagghe, G., & Van Camp, J. (2015). Liquid chromatography-mass spectrometry coupled with multivariate analysis for the characterization and discrimination of extractable and nonextractable polyphenols and glucosinolates from red cabbage and Brussels sprout waste streams. *Journal of Chromatography A*, *1402*, 60–70. http://dx.doi.org/10.1016/j.chroma.2015.05.009.

Gonzales, G. B., Smagghe, G., Raes, K., & Van Camp, J. (2014). Combined alkaline hydrolysis and ultrasound-assisted extraction for the release of nonextractable phenolics from cauliflower (*Brassica oleracea var. botrytis*) waste. *Journal of Agricultural and Food Chemistry*, *62*(15), 3371–3376. http://dx.doi.org/10.1021/jf500835q.

Grata, E., Guillarme, D., Glauser, G., Boccard, J., Carrupt, P. A., Veuthey, J. L., ... Wolfender, J. L. (2009). Metabolite profiling of plant extracts by ultra-high-pressure liquid chromatography at elevated temperature coupled to time-of-flight mass spectrometry. *Journal of Chromatography A*, *1216*(30), 5660–5668. http://dx.doi.org/10.1016/j.chroma.2009.05.069.

Greco, G., Grosse, S., & Letzel, T. (2013). Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine. *Journal of Separation Science*, *36*(8), 1379–1388. http://dx.doi.org/10.1002/jssc.201200920.

Guo, J., Yue, T., Yuan, Y., & Wang, Y. (2013). Chemometric classification of apple juices according to variety and geographical origin based on polyphenolic profiles. *Journal of Agricultural and Food Chemistry*, *61*(28), 6949–6963. http://dx.doi.org/10.1021/jf4011774.

He, W., Zeng, M., Chen, J., Jiao, Y., Niu, F., Tao, G., ... He, Z. (2016). Identification and quantitation of anthocyanins in purple-fleshed sweet potatoes cultivated in China by UPLC-PDA and UPLC-QTOF-MS/MS. *Journal of Agricultural and Food Chemistry*, *64*(1), 171–177. http://dx.doi.org/10.1021/acs.jafc.5b04878.

Hjelmeland, A. K., Zweigenbaum, J., & Ebeler, S. E. (2015). Profiling monoterpenol glycoconjugation in *Vitis vinifera* L. cv. Muscat of Alexandria using a novel putative compound database approach, high resolution mass spectrometry and collision induced dissociation fragmentation analysis. *Analytica Chimica Acta*, *887*, 138–147. http://dx.doi.org/10.1016/j.aca.2015.06.026.

Hughey, C. A., McMinn, C. M., & Phung, J. (2016). Beeromics: From quality control to identification of differentially expressed compounds in beer. *Metabolomics*, *12*(1), 1–13. http://dx.doi.org/10.1007/s11306-015-0885-5.

Hurtado-Fernández, E., Pacchiarotta, T., Gómez-Romero, M., Schoenmaker, B., Derks, R., Deelder, A. M., ... Fernández-Gutiérrez, A. (2011). Ultra high performance liquid chromatography-time of flight mass spectrometry for analysis of avocado fruit metabolites: Method evaluation and applicability to the analysis of ripening degrees. *Journal of Chromatography A*, *1218*(42), 7723–7738. http://dx.doi.org/10.1016/j.chroma.2011.08.059.

Huynh, N. T., Smagghe, G., Gonzales, G. B., Van Camp, J., & Raes, K. (2014). Enzyme-assisted extraction enhancing the phenolic release from cauliflower (*Brassica oleracea* L. *var. botrytis*) outer leaves. *Journal of Agricultural and Food Chemistry*, *62*(30), 7468–7476. http://dx.doi.org/10.1021/jf502543c.

Ignat, I., Volf, I., & Popa, V. I. (2011). A critical review of methods for characterisation of polyphenolic compounds in fruits and vegetables. *Food Chemistry*. http://dx.doi.org/10.1016/j.foodchem.2010.12.026.

Iswaldi, I., Gómez-Caravaca, A. M., Arráez-Román, D., Uberos, J., Lardón, M., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2012). Characterization by high-performance liquid chromatography with diode-array detection coupled to time-of-flight mass spectrometry of the phenolic fraction in a cranberry syrup used to prevent urinary tract diseases, together with a study of its antibacterial activity. *Journal of Pharmaceutical and Biomedical Analysis*, *58*(1), 34–41. http://dx.doi.org/10.1016/j.jpba.2011.09.027.

Jaiswal, R., & Kuhnert, N. (2014). Identification and characterization of the phenolic glycosides of *Lagenaria siceraria* stand. (bottle gourd) fruit by liquid chromatography-tandem mass spectrometry. *Journal of Agricultural and Food Chemistry*, *62*(6), 1261–1271. http://dx.doi.org/10.1021/jf4053989.

Jerman Klen, T., Golc Wondra, A., Vrhovšek, U., & Mozetič Vodopivec, B. (2015). Phenolic profiling of olives and olive oil process-derived matrices using UPLC-DAD-ESI-QTOF-HRMS analysis. *Journal of Agricultural and Food Chemistry*, *63*(15), 3859–3872. http://dx.doi.org/10.1021/jf506345q.

Jian, L., Du, C. J., Lee, A. H., & Binns, C. W. (2005). Do dietary lycopene and other carotenoids protect against prostate cancer? *International Journal of Cancer*, *113*(6), 1010–1014. http://dx.doi.org/10.1002/ijc.20667.

Jiménez-Sánchez, C., Lozano-Sánchez, J., Marti, N., Saura, D., Valero, M., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2015). Characterization of polyphenols, sugars, and other polar compounds in persimmon juices produced under different technologies and their assessment in terms of compositional variations. *Food Chemistry*, *182*, 282–291. http://dx.doi.org/10.1016/j.foodchem.2015.03.008.

Jiménez-Sánchez, C., Lozano-Sánchez, J., Rodríguez-Pérez, C., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2016). Comprehensive, untargeted, and qualitative RP-HPLC-ESI-QTOF/MS^2 metabolite profiling of green asparagus (*Asparagus officinalis*). *Journal of Food Composition and Analysis*, *46*, 78–87. http://dx.doi.org/10.1016/j.jfca.2015.11.004.

Kachlicki, P., Piasecka, A., Stobiecki, M., & Marczak, Ł. (2016). Structural characterization of flavonoid glycoconjugates and their derivatives with mass spectrometric techniques. *Molecules*, *21*(11), 1494–1514. http://dx.doi.org/10.3390/molecules21111494.

Kalili, K. M., & de Villiers, A. (2009). Off-line comprehensive 2-dimensional hydrophilic interaction × reversed phase liquid chromatography analysis of procyanidins. *Journal of Chromatography A*, *1216*(35), 6274–6284. http://dx.doi.org/10.1016/j.chroma.2009.06.071.

Kalili, K. M., Vestner, J., Stander, M. A., & De Villiers, A. (2013). Toward unraveling grape tannin composition: Application of online hydrophilic interaction chromatography × reversed-phase liquid chromatography-time-of-flight mass spectrometry for grape seed analysis. *Analytical Chemistry*, *85*(19), 9107–9115. http://dx.doi.org/10.1021/ac401896r.

Kamiloglu, S., Capanoglu, E., Bilen, F. D., Gonzales, G. B., Grootaert, C., Van De Wiele, T., & Van Camp, J. (2016). Bioaccessibility of polyphenols from plant-processing by-products of black carrot (*Daucus carota* L.). *Journal of Agricultural and Food Chemistry*, *64*(12), 2450–2458. http://dx.doi.org/10.1021/acs.jafc.5b02640.

Karkoula, E., Skantzari, A., Melliou, E., & Magiatis, P. (2012). Direct measurement of oleocanthal and oleacein levels in olive oil by quantitative ^{1}H NMR. Establishment of a new index for the characterization of extra virgin olive oils. *Journal of Agricultural and Food Chemistry*, *60*(47), 11696–11703. http://dx.doi.org/10.1021/jf3032765.

Kårlund, A., Hanhineva, K., Lehtonen, M., Karjalainen, R. O., & Sandell, M. (2015). Nontargeted metabolite profiles and sensory properties of strawberry cultivars grown both organically and conventionally. *Journal of Agricultural and Food Chemistry*, *63*(3), 1010–1019. http://dx.doi.org/10.1021/jf505183j.

Kaufmann, A. (2014). Combining UHPLC and high-resolution MS: A viable approach for the analysis of complex samples? *TrAC Trends in Analytical Chemistry*. http://dx.doi.org/10.1016/j.trac.2014.06.025.

Kečkeš, S., Gašić, U., Veličković, T.Ć., Milojković-Opsenica, D., Natić, M., & Tešić, Ž. (2013). The determination of phenolic profiles of Serbian unifloral honeys using ultra-high-performance liquid chromatography/high resolution accurate mass spectrometry. *Food Chemistry*, *138*(1), 32–40. http://dx.doi.org/10.1016/j.foodchem.2012.10.025.

Khoddami, A., Wilkes, M. A., & Roberts, T. H. (2013). Techniques for analysis of plant phenolic compounds. *Molecules*, *18*(2), 2328–2375. http://dx.doi.org/10.3390/molecules18022328.

Kim, H. J., Park, S. R., & Jang, Y. P. (2014). Extraction-free in situ derivatisation of timosaponin AIII using direct analysis in real time TOF/MS. *Phytochemical Analysis*, *25*(4), 373–377. http://dx.doi.org/10.1002/pca.2488.

Kim, H. K., & Verpoorte, R. (2010). Sample preparation for plant metabolomics. *Phytochemical Analysis*. http://dx.doi.org/10.1002/pca.1188.

Kimura, H., Ogawa, S., Akihiro, T., & Yokota, K. (2011). Structural analysis of A-type or B-type highly polymeric proanthocyanidins by thiolytic degradation and the implication in their inhibitory effects on pancreatic lipase. *Journal of Chromatography A*, *1218*(42), 7704–7712. http://dx.doi.org/10.1016/j.chroma.2011.07.024.

Krauss, M., Singer, H., & Hollender, J. (2010). LC-high resolution MS in environmental analysis: From target screening to the identification of unknowns. *Analytical and Bioanalytical Chemistry*, *397*(3), 943–951. http://dx.doi.org/10.1007/s00216-010-3608-9.

La Barbera, G., Capriotti, A. L., Cavaliere, C., Piovesana, S., Samperi, R., Zenezini Chiozzi, R., & Laganà, A. (2017). Comprehensive polyphenol profiling of a strawberry extract (*Fragaria* × *ananassa*) by ultra-high-performance liquid chromatography coupled with high-resolution mass spectrometry. *Analytical and Bioanalytical Chemistry*, *409*(8), 2127–2142. http://dx.doi.org/10.1007/s00216-016-0159-8.

Lai, T. N. H., Herent, M. F., Quetin-Leclercq, J., Nguyen, T. B. T., Rogez, H., Larondelle, Y., & André, C. M. (2013). Piceatannol, a potent bioactive stilbene, as major phenolic component in *Rhodomyrtus tomentosa*. *Food Chemistry*, *138*(2–3), 1421–1430. http://dx.doi.org/10.1016/j.foodchem.2012.10.125.

Lambert, J. D., & Elias, R. J. (2010). The antioxidant and pro-oxidant activities of green tea polyphenols: A role in cancer prevention. *Archives of Biochemistry and Biophysics*, *501*(1), 65–72. http://dx.doi.org/10.1016/j.abb.2010.06.013.

Lee, J. G., Lim, S., Kim, J., & Lee, E. J. (2017). The mechanism of deterioration of the glucosinolate-myrosynase system in radish roots during cold storage after harvest. *Food Chemistry*, *233*, 60–68. http://dx.doi.org/10.1016/j.foodchem.2017.04.104.

Leiherer, A., Mündlein, A., & Drexel, H. (2013). Phytochemicals and their impact on adipose tissue inflammation and diabetes. *Vascular Pharmacology*, *58*(1–2), 3–20. http://dx.doi.org/10.1016/j.vph.2012.09.002.

Lelario, F., Bianco, G., Bufo, S. A., & Cataldi, T. R. I. (2012). Establishing the occurrence of major and minor glucosinolates in *Brassicaceae* by LC-ESI-hybrid linear ion-trap and Fourier-transform ion cyclotron resonance mass spectrometry. *Phytochemistry*, *73*, 74–83. http://dx.doi.org/10.1016/j.phytochem.2011.09.010.

Lin, L. Z., Sun, J., Chen, P., & Harnly, J. (2011). UHPLC-PDA-ESI/HRMS/MSn analysis of

anthocyanins, flavonol glycosides, and hydroxycinnamic acid derivatives in red mustard greens (*Brassica juncea* Coss variety). *Journal of Agricultural and Food Chemistry*, *59*(22), 12059–12072. http://dx.doi.org/10.1021/jf202556p.

Loizou, S., Lekakis, I., Chrousos, G. P., & Moutsatsou, P. (2010). Beta-sitosterol exhibits anti-inflammatory activity in human aortic endothelial cells. *Molecular Nutrition & Food Research*, *54*, 551–558. http://dx.doi.org/10.1002/mnfr.200900012.

López-Cobo, A., Gómez-Caravaca, A. M., Cerretani, L., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2014). Distribution of phenolic compounds and other polar compounds in the tuber of *Solanum tuberosum* L. by HPLC-DAD-q-TOF and study of their antioxidant activity. *Journal of Food Composition and Analysis*, *36*(1–2), 1–11. http://dx.doi.org/10.1016/j.jfca.2014.04.009.

Lourenço, C. F., Gago, B., Barbosa, R. M., De Freitas, V., & Laranjinha, J. (2008). LDL isolated from plasma-loaded red wine procyanidins resist lipid oxidation and tocopherol depletion. *Journal of Agricultural and Food Chemistry*, *56*(10), 3798–3804. http://dx.doi.org/10.1021/jf0733259.

Lozano-Sánchez, J., Segura-Carretero, A., Menendez, J. A., Oliveras-Ferraros, C., Cerretani, L., & Fernández-Gutiérrez, A. (2010). Prediction of extra virgin olive oil varieties through their phenolic profile. Potential cytotoxic activity against human breast cancer cells. *Journal of Agricultural and Food Chemistry*, *58*(18), 9942–9955. http://dx.doi.org/10.1021/jf101502q.

Mannarino, M. R., Ministrini, S., & Pirro, M. (2014). Nutraceuticals for the treatment of hypercholesterolemia. *European Journal of Internal Medicine*, *25*(7), 592–599. http://dx.doi.org/10.1016/j.ejim.2014.06.008.

Mason, E. A., & Schamp, H. W. (1958). Mobility of gaseous ions in weak electric fields. *Annals of Physics*, *4*(3), 233–270. http://dx.doi.org/10.1016/0003-4916(58)90049-6.

May, J. C., & McLean, J. A. (2015). Ion mobility-mass spectrometry: Time-dispersive instrumentation. *Analytical Chemistry*. http://dx.doi.org/10.1021/ac504720m.

McIntyre, B. S., Briski, K. P., Gapor, A., & Sylvester, P. W. (2000). Antiproliferative and apoptotic effects of tocopherols and tocotrienols on preneoplastic and neoplastic mouse mammary epithelial cells. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, *224*(4), 292–301. http://dx.doi.org/10.1046/j.1525-1373.2000.22434.x.

Mellors, J. S., & Jorgenson, J. W. (2004). Use of 1.5-μm porous ethyl-bridged hybrid particles as a stationary-phase support for reversed-phase ultrahigh-pressure liquid chromatography. *Analytical Chemistry*, *76*(18), 5441–5450. http://dx.doi.org/10.1021/ac049643d.

Mikołajczyk-Bator, K., Błaszczyk, A., Czyżniejewski, M., & Kachlicki, P. (2016). Characterisation and identification of triterpene saponins in the roots of red beets (*Beta vulgaris* L.) using two HPLC–MS systems. *Food Chemistry*, *192*, 979–990. http://dx.doi.org/10.1016/j.foodchem.2015.07.111.

Milev, B. P., Patras, M. A., Dittmar, T., Vrancken, G., & Kuhnert, N. (2014). Fourier transform ion cyclotron resonance mass spectrometrical analysis of raw fermented cocoa beans of cameroon and ivory coast origin. *Food Research International*, *64*, 958–961. http://dx.doi.org/10.1016/j.foodres.2014.07.012.

Misra, B. B., & van der Hooft, J. J. J. (2016). Updates in metabolomics tools and resources: 2014–2015. *Electrophoresis*, *37*(1), 86–110. http://dx.doi.org/10.1002/elps.201500417.

Morales-Soto, A., Gómez-Caravaca, A. M., García-Salas, P., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2013). High-performance liquid chromatography coupled to diode array and electrospray time-of-flight mass spectrometry detectors for a comprehensive characterization of phenolic and other polar compounds in three pepper (*Capsicum annuum* L.) samples. *Food Research International*, *51*(2), 977–984. http://dx.doi.org/10.1016/j.foodres.2013.02.022.

Nakabayashi, R., & Saito, K. (2013). Metabolomics for unknown plant metabolites. *Analytical and Bioanalytical Chemistry*, *405*(15), 5005–5011. http://dx.doi.org/10.1007/s00216-013-6869-2.

Nakabayashi, R., Sawada, Y., Yamada, Y., Suzuki, M., Hirai, M. Y., Sakurai, T., & Saito, K. (2013). Combination of liquid chromatography-fourier transform ion cyclotron resonance-mass spectrometry with 13C-labeling for chemical assignment of sulfur-containing metabolites in onion bulbs. *Analytical Chemistry*, *85*(3), 1310–1315. http://dx.doi.org/10.1021/ac302733c.

Navarro, M., Núñez, O., Saurina, J., Hernández-Cassou, S., & Puignou, L. (2014). Characterization of fruit products by capillary zone electrophoresis and liquid chromatography using the compositional profiles of polyphenols: Application to authentication of natural extracts. *Journal of Agricultural and Food Chemistry*, *62*(5), 1038–1046. http://dx.doi.org/10.1021/jf404776d.

Oliveira, J., Alhinho Da Silva, M., Teixeira, N., De Freitas, V., & Salas, E. (2015). Screening of anthocyanins and anthocyanin-derived pigments in red wine grape pomace using LC-DAD/MS and MALDI-TOF techniques. *Journal of Agricultural and Food Chemistry*, *63*(35), 7636–7644. http://dx.doi.org/10.1021/acs.jafc.5b00256.

Ortmayr, K., Causon, T. J., Hann, S., & Koellensperger, G. (2016). Increasing selectivity and coverage in LC-MS based metabolome analysis. *TrAC Trends in Analytical Chemistry*, *82*, 358–366. http://dx.doi.org/10.1016/j.trac.2016.06.011.

Oz, T. A., & Kafkas, E. (2017). Phytochemicals in fruits and vegetables. In V. Waisundara, & N. Shiomi (Eds.), *Superfood and functional food - An overview of their processing and utilization* (pp. 175–184). InTech. http://dx.doi.org/10.5772/63180.

Passo Tsamo, C. V., Herent, M. F., Tomekpe, K., Happi Emaga, T., Quetin-Leclercq, J., Rogez, H., ... Andre, C. (2015). Phenolic profiling in the pulp and peel of nine plantain cultivars (*Musa* sp.). *Food Chemistry*, *167*, 197–204. http://dx.doi.org/10.1016/j.foodchem.2014.06.095.

Patras, M. A., Milev, B. P., Vrancken, G., & Kuhnert, N. (2014). Identification of novel cocoa flavonoids from raw fermented cocoa beans by HPLC-MS^n. *Food Research International*, *63*, 353–359. http://dx.doi.org/10.1016/j.foodres.2014.05.031.

Peralbo-Molina, Á., Priego-Capote, F., & Luque de Castro, M. D. (2012). Tentative identification of phenolic compounds in olive pomace extracts using liquid chromatography–tandem mass spectrometry with a quadrupole–quadrupole-time-of-flight mass detector. *Journal of Agricultural and Food Chemistry*, *60*(46), 11542–11550. http://dx.doi.org/10.1021/jf302896m.

Pierson, J. T., Monteith, G. R., Roberts-Thomson, S. J., Dietzgen, R. G., Gidley, M. J., & Shaw, P. N. (2014). Phytochemical extraction, characterisation and comparative distribution across four mango (*Mangifera indica* L.) fruit varieties. *Food Chemistry*, *149*, 253–263. http://dx.doi.org/10.1016/j.foodchem.2013.10.108.

Prchalová, J., Kovařík, F., Ševčík, R., Čížková, H., & Rajchl, A. (2014). Characterization of mustard seeds and paste by DART ionization with time-of-flight mass spectrometry. *Journal of Mass Spectrometry*, *49*(9), 811–818. http://dx.doi.org/10.1002/jms.3419.

Pyke, J. S., Callahan, D. L., Kanojia, K., Bowne, J., Sahani, S., Tull, D., ... Roessner, U. (2015). A tandem liquid chromatography-mass spectrometry (LC-MS) method for profiling small molecules in complex samples. *Metabolomics*, *11*(6), 1552–1562. http://dx.doi.org/10.1007/s11306-015-0806-7.

Quifer-Rada, P., Vallverdú-Queralt, A., Martínez-Huélamo, M., Chiva-Blanch, G., Jáuregui, O., Estruch, R., & Lamuela-Raventós, R. (2015). A comprehensive characterisation of beer polyphenols by high resolution mass spectrometry (LC-ESI-LTQ-Orbitrap-MS). *Food Chemistry*, *169*, 336–343. http://dx.doi.org/10.1016/j.foodchem.2014.07.154.

Rai, B., Kaur, J., Jacobs, R., & Singh, J. (2010). Possible action mechanism for curcumin in pre-cancerous lesions based on serum and salivary markers of oxidative stress. *Journal of Oral Science*, *52*(2), 251–256. http://dx.doi.org/10.2334/josnusd.52.251.

Ramirez, J. E., Zambrano, R., Sepúlveda, B., Kennelly, E. J., & Simirgiotis, M. J. (2015). Anthocyanins and antioxidant capacities of six Chilean berries by HPLC-HR-ESI-ToF-MS. *Food Chemistry*, *176*, 106–114. http://dx.doi.org/10.1016/j.foodchem.2014.12.039.

Ramirez-Ambrosi, M., Abad-Garcia, B., Viloria-Bernal, M., Garmon-Lobato, S., Berrueta, L. A., & Gallo, B. (2013). A new ultrahigh performance liquid chromatography with diode array detection coupled to electrospray ionization and quadrupole time-of-flight mass spectrometry analytical strategy for fast analysis and improved characterization of phenolic compounds in apple products. *Journal of Chromatography A*, *1316*, 78–91. http://dx.doi.org/10.1016/j.chroma.2013.09.075.

Rathahao-Paris, E., Alves, S., Junot, C., & Tabet, J.-C. (2016). High resolution mass spectrometry for structural identification of metabolites in metabolomics. *Metabolomics*, *12*(1), 10. http://dx.doi.org/10.1007/s11306-015-0882-8.

Ravindran, R., & Jaiswal, A. K. (2016). Exploitation of food industry waste for high-value products. *Trends in Biotechnology*, *34*(1), 58–69. http://dx.doi.org/10.1016/j.tibtech.2015.10.008.

Regueiro, J., Sánchez-González, C., Vallverdú-Queralt, A., Simal-Gándara, J., Lamuela-Raventós, R., & Izquierdo-Pulido, M. (2014). Comprehensive identification of walnut polyphenols by liquid chromatography coupled to linear ion trap-Orbitrap mass spectrometry. *Food Chemistry*, *152*, 340–348. http://dx.doi.org/10.1016/j.foodchem.2013.11.158.

Roberts, L. D., Souza, A. L., Gerszten, R. E., & Clish, C. B. (2012). Targeted metabolomics. *Current Protocols in Molecular Biology*, *1*. http://dx.doi.org/10.1002/0471142727.mb3002s98.

Rockenbach, I. I., Jungfer, E., Ritter, C., Santiago-Schübel, B., Thiele, B., Fett, R., & Galensa, R. (2012). Characterization of flavan-3-ols in seeds of grape pomace by CE, HPLC-DAD-MSn and LC-ESI-FTICR-MS. *Food Research International*, *48*(2), 848–855. http://dx.doi.org/10.1016/j.foodres.2012.07.001.

Rodriguez-Aller, M., Gurny, R., Veuthey, J.-L., & Guillarme, D. (2012). Coupling ultra high-pressure liquid chromatography with mass spectrometry: Constraints and possible applications. *Journal of Chromatography A*, *1292*, 2–18. http://dx.doi.org/10.1016/j.chroma.2012.09.061.

Roullier-Gall, C., Lucio, M., Noret, L., Schmitt-Kopplin, P., & Gougeon, R. D. (2014). How subtle is the "terroir" effect? Chemistry-related signatures of two "climats de Bourgogne". *PLoS ONE*, *9*(5), http://dx.doi.org/10.1371/journal.pone.0097615.

Russo, M., Nigro, P., Rosiello, R., D'Arienzo, R., & Russo, G. L. (2007). Quercetin enhances CD95- and TRAIL-induced apoptosis in leukemia cell lines. *Leukemia*, *21*(5), 1130–1133. http://dx.doi.org/10.1038/sj.leu.2404610.

Sawada, Y., Nakabayashi, R., Yamada, Y., Suzuki, M., Sato, M., Sakata, A., ... Saito, K. (2012). RIKEN tandem mass spectral database (ReSpect) for phytochemicals: A plant-specific MS/MS-based data resource and database. *Phytochemistry*, *82*, 38–45. http://dx.doi.org/10.1016/j.phytochem.2012.07.007.

Sbrana, C., Avio, L., & Giovannetti, M. (2014). Beneficial mycorrhizal symbionts affecting the production of health-promoting phytochemicals. *Electrophoresis*, *35*(11), 1535–1546. http://dx.doi.org/10.1002/elps.201300568.

Scalbert, A., Andres-Lacueva, C., Arita, M., Kroon, P., Manach, C., Urpi-Sarda, M., & Wishart, D. (2011). Databases on food phytochemicals and their health-promoting effects. *Journal of Agricultural and Food Chemistry*, *59*(9), 4331–4348. http://dx.doi.org/10.1021/jf200591d.

Scheibner, O., Squibb, A., Greco, G., & Steiner, F. (2016). Target, suspected-target, and non-target LC-MS(/MS) screening: New (practical) strategies for CECs in water bodies. *Assessing transformation products of chemicals by non-target and suspect screening - Strategies and workflows volume 2. 1242. Assessing transformation products of chemicals by non-target and suspect screening - Strategies and workflows volume 2* (pp. 143–154). American Chemical Society. http://dx.doi.org/10.1021/bk-2016-1242.ch008.

Scigelova, M., Hornshaw, M., Giannakopulos, A., & Makarov, A. (2011). Fourier transform mass spectrometry. *Molecular & Cellular Proteomics: MCP*, *10*(7), M111.009431. http://dx.doi.org/10.1074/mcp.M111.009431.

Senko, M. W., Remes, P. M., Canterbury, J. D., Mathur, R., Song, Q., Eliuk, S. M., ... Zabrouskov, V. (2013). Novel parallelized quadrupole/linear ion trap/orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. *Analytical Chemistry*, *85*(24), 11710–11714. http://dx.doi.org/10.1021/ac403115c.

Shehzad, A., Lee, J., & Lee, Y. S. (2013). Curcumin in various cancers. *BioFactors*, *39*(1), 56–68. http://dx.doi.org/10.1002/biof.1068.

Shouk, R., Abdou, A., Shetty, K., Sarkar, D., & Eid, A. H. (2014). Mechanisms underlying the antihypertensive effects of garlic bioactives. *Nutrition Research*, *34*(2), 106–115. http://dx.doi.org/10.1016/j.nutres.2013.12.005.

Simirgiotis, M. J., Ramirez, J. E., Schmeda Hirschmann, G., & Kennelly, E. J. (2013). Bioactive coumarins and HPLC-PDA-ESI-ToF-MS metabolic profiling of edible queule fruits (*Gomortega keule*), an endangered endemic Chilean species. *Food Research International*, *54*(1), 532–543. http://dx.doi.org/10.1016/j.foodres.2013.07.022.

Sommella, E., Pepe, G., Pagano, F., Tenore, G. C., Dugo, P., Manfra, M., & Campiglia, P. (2013). Ultra high performance liquid chromatography with ion-trap TOF-MS for the fast characterization of flavonoids in *Citrus bergamia* juice. *Journal of Separation Science*, *36*(20), 3351–3355. http://dx.doi.org/10.1002/jssc.201300591.

Stefani, M., & Rigacci, S. (2014). Beneficial properties of natural phenols: Highlight on protection against pathological conditions associated with amyloid aggregation. *BioFactors*, *40*(5), 482–493. http://dx.doi.org/10.1002/biof.1171.

Sumner, L., Amberg, A., Barrett, D., Beale, M., Beger, R., Daykin, C., ... Viant, M. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics*, *3*(3), 211–221. http://dx.doi.org/10.1007/s11306-007-0082-2.

Sun, J., Lin, L. Z., & Chen, P. (2012). Study of the mass spectrometric behaviors of anthocyanins in negative ionization mode and its applications for characterization of anthocyanins and non-anthocyanin polyphenols. *Rapid Communications in Mass Spectrometry*, *26*(9), 1123–1133. http://dx.doi.org/10.1002/rcm.6209.

Sun, J., Liu, X., Yang, T., Slovin, J., & Chen, P. (2014). Profiling polyphenols of two diploid strawberry (*Fragaria vesca*) inbred lines using UHPLC-HRMSn. *Food Chemistry*, *146*, 289–298. http://dx.doi.org/10.1016/j.foodchem.2013.08.089.

Sun, J., Xiao, Z., Lin, L. Z., Lester, G. E., Wang, Q., Harnly, J. M., & Chen, P. (2013). Profiling polyphenols in five brassica species microgreens by UHPLC-PDA-ESI/HRMSn. *Journal of Agricultural and Food Chemistry*, *61*(46), 10960–10970. http://dx.doi.org/10.1021/jf401802n.

Sun, J., Zhang, M., & Chen, P. (2016). GLS-finder: A platform for fast profiling of glucosinolates in brassica vegetables. *Journal of Agricultural and Food Chemistry*, *64*(21), 4407–4415. http://dx.doi.org/10.1021/acs.jafc.6b01277.

Tang, L., Zirpoli, G., Jayaprakash, V., Reid, M., McCann, S., Nwogu, C., ... Moysich, K. (2010). Cruciferous vegetable intake is inversely associated with lung cancer risk among smokers: A case-control study. *BMC Cancer*, *10*(1), 162. http://dx.doi.org/10.1186/1471-2407-10-162.

Theodoridis, G. A., Gika, H. G., Want, E. J., & Wilson, I. D. (2012). Liquid chromatography-mass spectrometry based global metabolite profiling: A review. *Analytica Chimica Acta*, *711*, 7–16. http://dx.doi.org/10.1016/j.aca.2011.09.042.

Toniolo, P., Van Kappel, A. L., Akhmedkhanov, A., Ferrari, P., Kato, I., Shore, R. E., & Riboli, E. (2001). Serum carotenoids and breast cancer. *American Journal of Epidemiology*, *153*(12), 1142–1147. http://dx.doi.org/10.1093/aje/153.12.1142.

Valls, J., Millán, S., Martí, M. P., Borràs, E., & Arola, L. (2009). Advanced separation methods of food anthocyanins, isoflavones and flavanols. *Journal of Chromatography A*, *1216*(43), 7143–7172. http://dx.doi.org/10.1016/j.chroma.2009.07.030.

Vallverdú-Queralt, A., Jáuregui, O., Di Lecce, G., Andrés-Lacueva, C., & Lamuela-Raventós, R. M. (2011). Screening of the polyphenol content of tomato-based products through accurate-mass spectrometry (HPLC-ESI-QTOF). *Food Chemistry*, *129*(3), 877–883. http://dx.doi.org/10.1016/j.foodchem.2011.05.038.

Vallverdú-Queralt, A., Jáuregui, O., Medina-Remón, A., Andrés-Lacueva, C., & Lamuela-Raventós, R. M. (2010). Improved characterization of tomato polyphenols using liquid chromatography/electrospray ionization linear ion trap quadrupole Orbitrap mass spectrometry and liquid chromatography/electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *24*(20), 2986–2992. http://dx.doi.org/10.1002/rcm.4731.

Vallverdú-Queralt, A., Jáuregui, O., Medina-Remón, A., & Lamuela-Raventós, R. M. (2012). Evaluation of a method to characterize the phenolic profile of organic and conventional tomatoes. *Journal of Agricultural and Food Chemistry*, *60*(13), 3373–3380. http://dx.doi.org/10.1021/jf204702f.

Van Meulebroek, L., Vanden Bussche, J., Steppe, K., & Vanhaecke, L. (2014). High-resolution Orbitrap mass spectrometry for the analysis of carotenoids in tomato fruit: Validation and comparative evaluation towards UV–VIS and tandem mass spectrometry. *Analytical and Bioanalytical Chemistry*, *406*(11), 2613–2626. http://dx.doi.org/10.1007/s00216-014-7654-6.

Vinayavekhin, N., & Saghatelian, A. (2010). Untargeted metabolomics. *Current Protocols in Molecular Biology*. http://dx.doi.org/10.1002/0471142727.mb3001s90.

Wang, X. J., Zhang, A. H., Yan, G. L., Han, Y., & Sun, H. (2014). UHPLC-MS for the analytical characterization of traditional Chinese medicines. *TrAC Trends in Analytical Chemistry*, *63*, 180–187. http://dx.doi.org/10.1016/j.trac.2014.05.013.

Weidner, C., de Groot, J. C., Prasad, A., Freiwald, A., Quedenau, C., Kliem, M., ... Sauer, S. (2012). Amorfrutins are potent antidiabetic dietary natural products. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(19), 7257–7262. http://dx.doi.org/10.1073/pnas.1116971109.

Willemse, C. M., Stander, M. A., & de Villiers, A. (2013). Hydrophilic interaction chromatographic analysis of anthocyanins. *Journal of Chromatography A*, *1319*, 127–140. http://dx.doi.org/10.1016/j.chroma.2013.10.045.

Willemse, C. M., Stander, M. A., Tredoux, A. G. J., & de Villiers, A. (2014). Comprehensive two-dimensional liquid chromatographic analysis of anthocyanins. *Journal of Chromatography A*, *1359*, 189–201. http://dx.doi.org/10.1016/j.chroma.2014.07.044.

Wolfender, J.-L., Marti, G., Thomas, A., & Bertrand, S. (2015). Current approaches and challenges for the metabolite profiling of complex natural extracts. *Journal of Chromatography A*, *1382*, 136–164. http://dx.doi.org/10.1016/j.chroma.2014.10.091.

Woyengo, T. A., Ramprasath, V. R., & Jones, P. J. H. (2009). Anticancer effects of phytosterols. *European Journal of Clinical Nutrition*, *63*, 813–820. http://dx.doi.org/10.1038/ejcn.2009.29.

Wu, H., Guo, J., Chen, S., Liu, X., Zhou, Y., Zhang, X., & Xu, X. (2013). Recent developments in qualitative and quantitative analysis of phytochemical constituents and their metabolites using liquid chromatography-mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis*, *72*, 267–291. http://dx.doi.org/10.1016/j.jpba.2012.09.004.

Yang, C. S., Lambert, J. D., & Sang, S. (2009). Antioxidative and anti-carcinogenic activities of tea polyphenols. *Archives of Toxicology*, *83*(1), 11–21. http://dx.doi.org/10.1007/s00204-008-0372-0.

Yokota, K., Kimura, H., Ogawa, S., & Akihiro, T. (2013). Analysis of A-type and B-type highly polymeric proanthocyanidins and their biological activities as nutraceuticals. *Journal of Chemistry*352042. http://dx.doi.org/10.1155/2013/352042.

Zehring, J., Reim, V., Schröter, D., Neugart, S., Schreiner, M., Rohn, S., & Maul, R. (2015). Identification of novel saponins in vegetable amaranth and characterization of their hemolytic activity. *Food Research International*, *78*, 361–368. http://dx.doi.org/10.1016/j.foodres.2015.09.010.

Zenezini Chiozzi, R., Capriotti, A. L., Cavaliere, C., Ferraris, F., La Barbera, G., Piovesana, S., & Laganà, A. (2017). Evaluation of column length and particle size effect on the untargeted profiling of a phytochemical mixture by using UHPLC coupled to high-resolution mass spectrometry. *Journal of Separation Science*, *40*(12), 2541–2557. http://dx.doi.org/10.1002/jssc.201700135.

Zhou, D., Xu, Q., Xue, X., Zhang, F., & Liang, X. (2006). Identification of *O*-diglycosyl flavanones in *Fructus aurantii* by liquid chromatography with electrospray ionization and collision-induced dissociation mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis*, *42*(4), 441–448. http://dx.doi.org/10.1016/j.jpba.2006.05.015.

Zubarev, R. A., & Makarov, A. (2013). Orbitrap mass spectrometry. *Analytical Chemistry*, *85*(11), 5288–5296. http://dx.doi.org/10.1021/ac4001223.