CHROMATOGRAPHY HYPHENATED TO HIGH RESOLUTION MASS SPECTROMETRY IN UNTARGETED METABOLOMICS FOR INVESTIGATION OF FOOD (BIO)MARKERS

Leticia Lacalle-Bergeron[1], David Izquierdo-Sandoval[1], Juan V. Sancho[1], Francisco J. López[1], Félix Hernández[1], Tania Portolés[1*]

[1]Enviromental and Public Health Analytical Chemistry, Research institute for Pesticides and Water (IUPA), University Jaume I, Av. Sos Baynat S/N, 12071 Castellón de la Plana, Spain.

tportole@uji.es

ABSTRACT

Currently, there is a growing demand by our society, authorities and science to increase the knowledge about the quality of food and its relationship with health and disease. Untargeted metabolomics approaches are emerging as powerful tools for exploring metabolic changes in biological systems under different conditions with great potential in the food field. To this aim, it is necessary to apply advanced analytical techniques, such as chromatography hyphenated to high resolution mass spectrometry, which provides enough sensitivity and selectivity to cover a wide range of metabolites in complex samples, as food and biological samples. The objective of this work is to provide an overview of the most widely adopted strategies based on the use of high resolution mass spectrometry-based techniques for the identification of food (bio)markers through the untargeted metabolomics workflow. Detailed information is provided about the trends in each stage of the metabolomics process from updated literature with the objective to help researchers to select the most appropriate metabolic approaches.

Key words: untargeted metabolomics, mass spectrometry, LC-HRMS, GC-HRMS, food sciences, nutrition, biomarkers

Abbreviations

APCI	atmospheric pressure chemical ionization
BFI	biomarker of food intake
CI	chemical ionization
DDA	data dependent acquisition
DIA	data independent acquisition
EI	electron ionization
ESI	electrospray ionization
FS	full scan
GC	gas chromatography
HILIC	hydrophilic interaction chromatography

IMS	ion mobility spectrometry
IT	ion trap mass analyser
LC	liquid chromatography
LR	low resolution
MS	mass spectrometry
HRMS	high-resolution mass spectrometry
MS/MS	tandem mass spectrometry
MS^n	sequential mass spectrometry
NMR	nuclear magnetic resonance
OT	Orbitrap mass analyser
Q	quadrupole mass analyser
RPLC	reversed phase liquid chromatography
TOF	time-of-flight mass analyser
xC-HRMS	chromatographic techniques coupled to high resolution mass spectrometry

1 INTRODUCTION

Foodomics has been defined as a new discipline that studies food and nutrition domains combining the application of advanced analytical techniques (omics tools) and bioinformatics. The use of omics tools, such as genomics, transcriptomics, proteomics and/or metabolomics, is a requirement to address the challenges presented in emerging working areas included in foodomics studies [1]. Metabolomics can be defined as a non-selective, comprehensive analytical approach for the identification and quantification of metabolites in a biological system, typically those small molecules with a molecular weight below 1500 Da [2]. Metabolomics has become a powerful tool for the study of the complex interactions between diet and the human or animal organisms enabling to expand our knowledge of the subtle changes at metabolic level activated by foods, nutrients and disease. It has allowed significant improvements in the field of dietary assessment since it enables the identification of novel and robust biomarkers of food intake (BFIs) enhancing the accuracy and the objectivity in the measurement of dietary exposures and reducing the bias and errors associated with self-report methods [3]. On the other hand, the potential of metabolomics as a robust, efficient and sensitive analytical methodology in food safety, quality and traceability is widely recognized [4].

Metabolomic studies are challenging because of the aim to characterize complex and diverse biological matrices containing compounds with a wide range of polarities or volatilities. Carbohydrates, lipids, amino acids, amines, steroids, phenolic compounds, carotenoids, alkaloids or volatile compounds, are examples of compounds that constitute the metabolome [2]. This enormous diversity has led to the emergence of sub-areas within the metabolomics field to narrow down the search for compounds with similar physicochemical properties. As an example, lipidomics deals with the determination of lipid classes, subclasses and lipid signalling molecules, providing a tool for the assessment of changes in lipid metabolism [5]. On the other hand, volatolomics is the sub-unit of metabolomics responsible of the detection, characterization and quantification of volatile metabolites in a biological system [6].

In general, two complementary approaches are used in metabolic research: metabolic profiling (targeted metabolomics) and metabolic fingerprinting (untargeted metabolomics). Metabolite profiling focuses on the analysis of a group of metabolites such as those related to a specific metabolic pathway. In this approach, target metabolites are selected beforehand and they

are assessed using specific analytical methods. Technological advances have increased the number of metabolites that can be quantified simultaneously. Moreover, the results of metabolic profiling are independent of the technology used for data acquisition. Metabolic fingerprinting does not aim to identify the entire set of metabolites but rather to compare patterns or fingerprints of metabolites that change in response to an altered state promoted by endogenous (disease, genetics…) or exogenous (diet, environment…) conditions. It can be used as a tool to evaluate the state of a biological system by comparing, for example, control and disease subjects, or to assay the success of a particular treatment (prognosis/recovery). Once a differential pattern is discovered, further steps to identify the contributing compounds (qualitative) and to determine the absolute amounts of metabolites that participate in the processes studied (quantitative) must be followed. This issue is not trivial and prior to boarding on the task of discovering metabolic biomarkers, sufficiently sensitive and selective instruments and extensive compound libraries for metabolite identification must be available, while wide experience in data analysis and interpretation is also necessary [7]. Unlike the traditional analytical workflow, untargeted metabolomics is an hypothesis-driven methodology, that means that to address a biological question the experiment must be design with the broadest perspective as possible, and the hypothesis is generated from the result [5]. As large data sets are obtained from the results, potent statistical tools, as multivariate analysis, are necessary to reduce the data complexity and to reveal underlying trends from which it is hoped that hypothesis can be generated [8]. **Figure 1** shows a typical workflow followed in metabolomics fingerprinting. Concerning the detection and identification of metabolites, high-resolution mass spectrometry (HRMS)-based techniques are, undoubtedly, the most suitable option to deal with the vast diversity of small molecules with distinct physicochemical properties in complex biological matrixes that constitute the metabolome. The main advantages of HRMS-based metabolomics are the high sensitivity and selectivity as well as the accurate-mass full-spectrum acquired data, together with possibility to be coupled on-line to a separation technique. The hyphenation of separation techniques, mainly gas chromatography (GC) and liquid chromatography (LC), with HRMS reduces the complexity of the mass spectral data, enhancing the sensitivity of the detection and providing additional information about the physicochemical characteristics of the analysed molecules. Moreover, HRMS analysers can be used as a hybrid instrument allowing acquisitions in tandem mass spectrometry mode (MS/MS or MS^n)

incorporating fragmentation data of the metabolites and facilitating the confirmation of known, reported, compounds or assisting the elucidation of unknown metabolites.

The starting point of this work was a comprehensive search in Scopus database using the following keywords: “precision foods”, “ functional foods”, “precision nutrition”, “food intake”, “biomarkers of intake”, “nutritional assessment”, “dietary markers”, “nutrimetabolomics”, “food quality”, “food safety”, “food authenticity”, “food fraud” and “food traceability”; along with keywords related to the analytical technique and the methodology: “untargeted” (and synonyms “fingerprinting”, “untarget” and “non-target”), “metabolomics”, “mass spectrometry”, “HRMS” and “MS”; in papers published between 2017 and 2020. Reviews, trends, perspectives and book chapters were kept separately as a source of information. With the articles selected, a discussion of the trends in chromatography-HRMS-based metabolomics fingerprinting within in the context of foodomics is provided using as guideline the workflow shown in **Figure 1**.

2 STUDY DESIGN

Bearing in mind the objective of metabolomics fingerprinting, the experimental design requires careful consideration prior to laboratory work to ensure the quality and validity of the results. Within this approach, an appropriate experimental design must undertake the acquisition of data related to a specific biological question while ensuring that covariants or cofounders are not present or are well characterised [9].

A common feature in experimental design is that cohorts should be homogenous in those factors that are not included in the biological question, avoiding unnecessary errors, false leads and “statistical noise” [8]. For example, within the context of the discovery of fraudulent practices in the food industry, dead on arrival and regularly slaughtered chickens metabolite patterns were compared. All the chickens were grown on the same farm, were of the same age and were fed the same diet; moreover, the same tissues were analysed [10]. Sometimes, sample characterisation after sample collection is necessary to define in which cohort the sample belongs; for example, testing panels made up of professional tasters were used to assess quality in olive oil samples [11] and green tea samples [12]. In other study, the quality in berries of sea buckthorn was defined using a colorimeter [13]. In some occasions, when the objective is to look for the variability between geographic origin or variety, the characterisation is provided by the supplier [14].

Regardless the aim of the study, the collection of metadata during sample collection is crucial to avoid bias and the incorporation of data related to confounding factors into the statistical analysis [9]. In studies involving humans it is imperative to collect demographic/physiological/lifestyle metadata since such factors are difficult to control and inter-subject metabolic variation could be hiding the metabolic changes produced by the food or diet to be assessed. In animal or cellular assays, where there is commonly an extensive control over these factors, the inter-subject variation is usually negligible. Inclusion and exclusion criteria are commonly established in human studies to avoid incorporating subjects whose demographic/physiological/lifestyle characteristics can produce undesirable results in a specific study. For example, smokers used to be excluded due to the potential exposure to polycyclic aromatic hydrocarbons [15].

Two approaches can be considered when designing foodomics studies: intervention studies and observational studies. Dietary intervention study designs generally involve participants consuming a specific standardised diet or food product over a defined time. In this way, the variation introduced by food storage, preparation process, as well as the type of food and the nutritional value, is in usually controlled. Biofluids, urine or blood, are collected at specific time-points depending on the research interest. For example, blood samples were collected at baseline and after three and six weeks of treatment with the aim to compare the metabolite fingerprints at different levels of red-meat consumption [15]. On the other hand, twenty-four-hour urine samples were collected in a four-way cross-over intervention for the investigation of biomarkers in different kinds of meat consumed in a restricted diet during 48 h [16]. Within a cross-over design, such as the previous example, the participants receive all treatments reducing the inter-subject variation. When a succession of treatments is applied to the same participants is necessary to include washouts in the study design, which can consist in returning to the habitual diet or in excluding the food of interest for the study. The importance of washouts lies in returning to the basal metabolic levels avoiding carryover. Besides, in such studies that include blood collection, the washout duration must be longer to ensure the recovery of red blood cells and platelets [17]. Other elements to highlight are the randomisation of subjects and the nutritional and isocaloric equivalence between treatments. Generally, dietary intervention studies are expensive and laborious to conduct and some methodological compromises are required, such as limiting the sample size or reducing the time of study [3].

In observational studies, two groups, generally low and high consumers of the food(s) or diet of interest, are selected from food intake data collected by traditional dietary assessment methods such as food frequency questionnaires (FFQ), dietary diaries or other dietary assessment tools. Broadly, participants are selected from large cohorts to perform a cross-sectional study; in other words, groups of participants are compared at a single time point. For example, a cross-sectional design was applied to a subgroup of the SU.VI.MAX cohort, funded by the French National Cancer Institute. All participants were invited to complete a 24-h dietary record every two months up to a total of 10, covering all days of the week and all seasons of the year to assess their adherence to the French dietary recommendation [18]. A meticulous exclusion process was applied in the previous example, for selecting a limited number of participants from a large cohort. A stratified randomisation was performed to ensure that experimental groups are balanced concerning the confounding variables [19]. Observational work usually involves studies with a large number of samples and long study time; however, the limitations of traditional dietary assessment in providing reliable information could be a source of bias. Sometimes, cross-sectional studies in large cohorts are used to validate biomarkers identified by interventional designs [20].

3 SAMPLING AND SAMPLE PREPARATION

Once the experimental design is established, the next steps in the analytical process involve sample collection and sample preparation in the laboratory, including the shipping and storage of samples. It is essential to minimize sources of confounding factors, random or systematic errors during these stages to ensure the generation of robust and reproducible data, which only result in the variation between the different classes defined in the study design. Samples should be representative in terms of the biological question, defining factors such as the type and amount of sample, time of collection, and ensuring proper randomisation and group balancing within the sampling plan [21]. After collection of the sample, the metabolome may change because of many factors such as enzymatic activities, exposure to oxygen, UV light and temperature; so, optimum transport and storage conditions must be established to avoid sample losses, transformations, or contamination. Sample preparation in untargeted metabolomics aims transforming the physiochemical properties of the sample in a reproducible way to make it compatible with the analytical method. It should be as less selective as possible, maintaining the most the metabolomic composition of the sample, covering a wide range of compounds. Minimising the steps in sample

preparation avoids losses of metabolites and facilitates a high sample throughput [22]. In most cases, sample preparation is reduced to a straightforward solvent-extraction [4] or even a simple "dilute and shoot" in the case of less complex matrices when using LC separation [22].

Before the extraction, it is necessary to homogenise the sample and reduce its size, together with metabolism quenching. In solid samples (e.g. food, human and animal tissues), freeze-dried powder sample or frozen samples are commonly used; the sample can also be ground in a mortar with liquid nitrogen [23]. Vortex and ultrasounds sonication can be used to perform a more exhaustive extraction of metabolites during solvent-extraction in solid matrices [24–26], while in liquid samples (e.g. beverages and bio-fluids), stirring and aliquots are usually applied for homogenization and size reduction.

Hydro-organic mixtures containing water, methanol, acetonitrile and/or formic acid are a common choice for extraction since such versatile solvent systems provide enough solubility for covering polar and semipolar metabolites [21]. Hydrophobic extraction mixtures using organic solvents, such as chloroform or dichloromethane, are appropriate for the extraction of the non-polar fraction of the metabolome, as for example lipids or volatile compounds. Double extractions are sometimes applied to cover both polar and non-polar metabolites [10,27]. For example, chloroform, water and methanol was applied for freeze-dried carrot samples; after centrifugation, the aqueous phase (water/methanol) was used for analysing polar compounds, and the chloroform phase was evaporated and reconstituted with methanol for lipids analysis [28].

Cold extraction is recommended to avoid enzymatic activity. Some compounds can be used to stop the metabolomic activity, as long as global extraction is not compromised; for example, O-(carboxymethyl) hydroxylamine hemihydrochloride (OCMHA) was added to inhibit enzyme *alliinases* in garlic samples [14]. After extraction, centrifugation is applied to eliminate the solid residues and the proteins precipitated by the organic solvent.

In GC-based methods, the non-volatile compounds should be carefully removed or being chemically derivatized, which increases the complexity of the sample treatment adding time-consuming steps. A typical derivatization consist on two-steps process: methoxymation for ketone functional groups protection with methoxyamine, followed by silylation with reagents as BSTFA [27], MSTFA [29] and MTBSTFA [19], to reduce the polarity of the molecule by reacting with the active hydrogen of polar functional groups(-OH , -COOH, $-NH_2$, -SH and $–PO_4^{-3}$). In this way,

it was possible to detect simultaneously chemical families like amino acids, sugars, organic acids and some fatty acids, among other metabolites, in vegetable matrices [30] or human biofluids [31].

In volatolomics studies, GC analysis is the natural selection. Dynamic headspace purge and trap *(DHS-P&T)* has been used to extract the volatile fraction in olive oil samples. Volatiles were released from the sample by the use of a nitrogen stream and then retained on a reversed-phase sorbent cartridge [11]. In some cases, a solvent extraction from the trap cartridge is applied with a GC-suitable solvent such as n-hexane. Another alternative for volatile extraction is to establish an equilibrium between the vapour phase and the adsorbent in a closed space. Thus, headspace solid-phase microextraction (HS-SPME) has been satisfactorily used in seeds and whisky analysis [32,33]. The volume of sample, temperature, equilibrium time, the necessity of stirring, salting-out and the type of fibre are parameters commonly optimised in SPME analysis, and PDMS/DVB coating is suitable for volatile global screening [34]. Parallel to the SPME analysis, a more extensive range of compounds, including volatiles and semi-volatiles, could be extracted with ethyl acetate in whiskey samples [33].

Blood plasma, serum and urine are the common biofluids studied. The extraction of metabolites from urine is usually made by dilution with water and centrifugation followed by filtration for removal urine proteins or particulates. The dilution can be done before or after the centrifugation, and the degree of dilution uses to be in the range 1:1 to 1:3 V/V [35]. An attempt to normalise the dilution factor was carried out by the measuring of specific gravity by refractometry before the analysis [36]. Regarding serum and plasma, due to their high protein content, the sample preparation scheme involves a simple protein precipitation followed by centrifugation and reconstitution [37–39]. It is also possible to extract exogenous metabolites by using acidified methanol [18].

4 INSTRUMENTAL ANALYSIS

There is no universal analytical technique in untargeted metabolomics. The analysis of complex samples and the vast diversity of small molecules with diverse physico-chemical properties that constitute the metabolome entails the need to use a wide variety of analytical techniques. It is highly recommended to run more than one platform to enhance the compound coverage and to obtain comprehensive information. The two major analytical platforms to perform

259 untargeted metabolomics are nuclear magnetic resonance (NMR) spectroscopy and HRMS-based
260 techniques. NMR advantages are the robust structural elucidation capabilities, the non-destruction
261 of the sample and the detection of non-ionizable compounds, among others [40]. However, it is
262 not capable of reaching the sensitivity of HRMS-based techniques and is less suitable to be coupled
263 on-line to separation techniques. The hyphenation of separation techniques with HRMS reduces
264 the complexity of the mass spectral data, enhancing the sensitivity of the detection, providing
265 useful information about the physicochemical characteristics of the analysed compounds. In
266 untargeted metabolomics, GC and LC are the most used separation techniques and both can be
267 easily coupled to HRMS. The accurate-mass full-spectrum information provided by HRMS is
268 essential for the reliable identification of the compounds previously separated by chromatography.

269 From the 79 articles reviewed in this paper that perform untargeted metabolomics for food
270 related biomarkers, 68 used LC-HRMS, 7 GC-HRMS and only 4 a combination of both techniques.
271 In other cases, one of these platforms is complementary to other techniques as capillary
272 electrophoresis-HRMS, NMR o GC-MS (nominal mass analysers). This review focuses on the
273 combination of chromatography with HRMS.

274 **4.1 High Resolution Mass Spectrometry (HRMS)**

275 The progress in untargeted metabolomics has been mainly driven by the improvements of
276 the analytical techniques; the most important being MS technological innovations [41]. The
277 improvements have been mainly focused on the increase of mass resolving power and sensitivity,
278 as well as broadening the dynamic range and enhancing the acquisition rate [4,42]. Among the
279 HRMS analysers, time-of-flight (TOF) and the Orbitrap (OT) are the most used, while Fourier
280 transform ion cyclotron resonance (FT-ICR) is less applied due to its low acquisition rate, which
281 makes difficult the coupling with fast chromatographic separations, as well as to its higher
282 maintenance costs, making it a less affordable analyser.

283 HRMS can be also used as hybrid instruments combined with low resolution (LR) mass
284 analysers, such as ion trap (IT) or quadrupole (Q), allowing to work not only under full scan (FS)
285 mode but also under tandem mass mode (MS/MS or MS^n), improving the identification based on
286 the fragmentation patterns. TOF and hybrid Q-TOF instruments are the most used in untargeted
287 metabolomics applied to food related sciences (applied in 11 % and 68 % of the reviewed research
288 articles, respectively); followed by hybrid OT analysers (13 % using Q-OT, and 8 % using IT-

OT). These mass analyzers can achieve mass accuracy below 2 ppm (with internal calibration). Mass resolution expressed as full width at half maximum (FWHM) can reach values up to 80,000 and 1,000,000 for TOF-based and last generation OT-based instruments, respectively. However, the resolving power is dependent on the duty cycle for OT-based instruments while Q-TOF analyzers are able to acquire at a scan rate up to 100 Hz independently of the resolving power. This fact has made Q-TOFs better suited when the chromatographic peaks are narrow as in GC or fast LC separations [43]. Nevertheless, the innovation on OT instruments has allowed the recent introduction of GC-OT instruments into the market with an interesting potential in future untargeted metabolomics applications [44]. Hybrid MS analysers allow the simultaneous MS acquisition and MS/MS or MS^n in a single injection (i.e. FS and target MS^n analysis). Under these acquisition modes, one can obtain both (semi)-quantitative (from the FS) and structural (from the MS^n) information in a single injection. Data dependent and data independent acquisitions modes can be applied in analysis (DDA and DIA, respectively).

Under DDA, the instrument automatically switches from FS to MS/MS or MS^n of the preselected ions detected in the FS spectrum. This preselection is intensity dependent along with other predefined parameters and may negatively affect the DDA coverage specially for low abundance features [45]. Licha et al. satisfactorily applied Q-OT under DDA for analysis of mice plasma samples after application of ketogenic diet to study the metabolic profile and its relationship with tumour growth inhibition [46]. In the DDA, MS^2 criteria specified that the five most abundant ions from every scan cycle were isolated in the Q in a window of 0.8 m/z and subsequently fragmented. Tovar et al. implemented a DDA acquisition method in a Q-TOF instrument to study the effect of multifunctional diet in human metabolism where only the 4 most abundant ions from every precursor scan cycle were selected for fragmentation [47]. Nevertheless, there was the need to perform additional target MS/MS measurements for those potential markers that failed to be included in the previous DDA method.

DIA systematically performs the fragmentation of all precursor ions along the full *m/z* range (also called all-ion fragmentation (AIF) or MS^E among other commercial names) or within a selection of sequential mass windows (like SONAR or sequential window acquisition for all theoretical spectra (SWATH)). Although DIA covers the DDA limitation for low abundance ions, the resulting MS/MS spectra is a composite of fragment ions generated from all precursor ions.

Thus, it is required the aid of powerful algorithms to stablish the link between the precursor ion and the fragmentation pattern [45]. Hoyos Ossa et al. applied MS^E acquisition method for the origin discrimination of Colombian green coffee [48]. The fragmentation spectra obtained under MS^E were not enough informative to allow the identity of the markers. Therefore, additional target MS/MS analysis was made to confirm the structure of the compounds used in the model of discrimination by origin. More information about data acquisition in untargeted metabolomics can be found in the extensive review of Fenaille et al. [42].

It is worth noting the difficulties to optimize a methodology in untargeted metabolomics when the compounds that may be relevant are unknown, contrary to targeted approaches, such as profiling metabolomics, where they are known, and reference standards are commonly available. Therefore, the choice is usually based on the possibility of fragmenting the maximum number of ions as possible and thus being able to cover a wider range of compounds that could be potential markers. Guo et al. made a comparative study of FS, DDA and DIA (AIF mode) in LC-QTOF untargeted metabolomics with different LC separations using spiked human urine samples. The best results where for FS in terms of sensitivity and quantitative precision, higher quality of MS^2 spectra with DDA but better MS^2 spectral coverage with DIA [45].

4.2 Gas chromatography-high resolution mass spectrometry (GC-HRMS)

GC is ideal for the separation of thermally stable and volatile compounds (or volatile derivatives previous chemical derivatization). Capillary columns from non-polar stationary phases as 100 % dimethylpolisiloxane [31], to polar as 100 % polyethylenglicol [33] may be used. One of the most applied in untargeted metabolomics is the non-polar stationary phase 5 % dimethyl-95 % dimethylpolysiloxane [27,29,30,49,50] or similar [11,19,32]. GC-MS is well established in metabolomics [41] because of its advantages of high chromatographic resolution, sensitivity and separation reproducibility [51]. However, aqueous samples must be dried or subjected to solvent exchange before GC-MS analysis (which can entail volatile losses). As mentioned above, those compounds that are not naturally volatile must be carefully removed or being chemically derivatized.

Electron ionization (EI), a robust and hard-ionization technique, is the most commonly used in GC-based metabolomic studies [52] where useful spectral databases have been built over the years, such as NIST. The availability of these databases facilitates the rapid identification of

the markers by mass spectral matching, which makes it the main attractiveness of the GC-EI-MS, especially compared to the LC-MS based metabolomics [51].

The ionization source has notable impact on the mass analyser selected. Indeed, the significant in-source fragmentation makes a hybrid analyser less useful, and so GC-EI is commonly coupled to a single mass analyser as TOF working in FS acquisition. As illustrative example, the plasma metabolic profiles associated with meat and seafood consumption in Asian population [19] were studied by LC-QTOF and GC-EI-QTOF analysis (previous derivatization). While for highlighted markers from LC analysis, additional MS/MS acquisition were needed for structural elucidation, GC-EI-MS analysis was performed only in FS and markers were annotated by fragmentation spectra matching with NIST library, with final identity confirmation with reference standards. The use of GC-EI-MS with LR analysers (e.g. Q) under FS mode continues to be widespread, since the structural identification power of the fragmentation spectrum together with the libraries make the exact mass acquisition of HRMS less necessary, in addition to being clearly more economical and accessible instrument for most laboratories.

Chemical ionization (CI) is less applied compared to EI. CI is a soft-ionization technique able to preserve the precursor ion, limited commonly to targeted analysis, as it is strongly dependent on the reagent gas and pressure used for the ionization [53]. Stupak et al. performed additional target MS/MS analysis with positive CI where the precursor ion was not found for those potential markers of quality and authenticity of Scotch whiskey with excessive fragmentation in EI [33].

4.3 Liquid chromatography-high resolution mass spectrometry (LC-HRMS)

LC is the most employed separation technique especially for aqueous samples as biofluids or some food matrices. Furthermore, it does not usually require complex sample preparations and involves short run times compared to GC. Due to the diversity of stationary phases and the different mobile phases that can be used, versatility is one of the main advantages of this technique, allowing its applicability to the analysis from medium to highly polar, low volatility and/or thermolabile compounds. If the interest is to reach the maximum coverage, as in untargeted metabolomics, more than one separation mechanisms should be assayed.

Reversed phase LC (RPLC) and hydrophilic interaction chromatography (HILIC) are the most used stationary phases in untargeted metabolomics. Among RPLC, ideally suited for the

analysis of semi-polar/nonpolar analytes, C18 stationary phase is the most commonly used due to its well-known behaviour, good robustness and its ability to cover a wide range of compounds. It has been applied in the identification of biomarkers of intake [54,55], designation/geographical of origin [56,57], and the study of the effect of functional food in health [58,59], among others. Two complementary C18 RPLC strategies have been used to assess the changes in plasma metabolome by the consumption of an herbal supplement, one more geared towards the lipidic metabolome (lipidomics) and the other to general metabolome (metabolomics) [60]. C18 columns modified with polar endcapping (e.g. as HSS T3 from Waters) are becoming more popular, as they are able to support highly aqueous mobile phases (even 100 % water) expanding their applicability to more polar compounds compared to the traditional C18. Kozlowska et al. detected with this stationary phase some nitrogenous bases as tryptophan metabolites, organic acid and phase II metabolites in urine, usually difficult retain in conventional C18 [61].

HILIC seems to be the choice for polar to highly polar compounds, but insufficiently charged for ion-exchange chromatography. The stationary phase is a highly hydrophilic, such as silica or chemically modified silica (as amide) and the mobile phase is an organic solvent containing a small amount of water (at least 5 %) [62]. HILIC separations were applied for the analysis of polar lipids, in different life stages, of one of the most consumed seaweed for sushi (*Porphyra dioica*) [63] as well as for the assessment of garlic authenticity, detecting polar metabolites as phospholipids and small peptides and aminoacids [14]. This separation mode is more affected by the chromatographic conditions and matrix effects, and it is known to be less reproducible than RPLC regarding retention time. HILIC is commonly used simultaneously with RPLC to obtain a complementary information on those polar compounds that RPLC cannot separate. As example, this combination has been applied for discovery of consumption biomarkers [36,64]. and to study the effects of different diets on health [65,66]. Pérez-Miguez et al. highlighted the advantages of combining HILIC with RPLC (and even capillary electrophoresis) for the study of coffee roasting process showing a comparative of the metabolites identified by each strategy [67].

Electrospray ionization (ESI) is clearly the preferred approach in untargeted metabolomics based on LC-MS analysis. Indeed, all the LC-HRMS studies reviewed made use of ESI, and 77 % of them used both positive and negative ionization modes. In comparison with GC-EI-MS, LC-

ESI-MS is more affected by the instrument-to-instrument differences which makes troublesome the matching with mass spectral databases. This fact and the high quantity of non-reported compounds in LC databases may hamper the identification of the (bio)markers, being the main bottleneck of untargeted metabolomics studies based on LC-ESI-HRMS.

5 DATA PROCESSING

Huge amounts of data are generated in untargeted metabolomics using chromatographic techniques coupled to HRMS (xC-HRMS). The objective of the data processing is to extract the information of the detected features from the xC-HRMS raw 3D data and obtain a 2D data matrix where they are characterized by *m/z* ratio, retention time (RT) and their relative intensities across the samples, which will be used for statistical analysis. The main steps are (i) *Peak picking and deconvolution*. It consists on the detection of each measured ion in a sample and the assignation to a feature (m/z and RT). The peak picking algorithm and deconvolution works with the extracted ion chromatograms attending to some parameter of maximum *m/z* error, interval of time (minimum and maximum time width to be considered a chromatographic peak) and the minimum height or intensity, signal to noise ratio (S/N), among other parameters. (ii) *Retention time alignment*. The matched peaks with similar retention times and *m/z* ratio across multiple samples are grouped in accordance to a window of m/z and RT, to be assigned as the same feature and subsequently aligned. This parameter is especially important in LC, as it tends to present more drift that can cause slights differences in retention times along the run. The grouped peaks are then integrated and a peak height or peak area is assigned to the feature in each sample. (iii) *Gap filling*: It is applied to correct and fill in the missing peaks (0 signal) or peaks not detected due to the restrictions of the first two steps (lower intensities or bad peak shape in some of the samples), but actually may be present, that can affect the power of subsequent statistical analysis.

At this point, a first data table is obtained and the quality of the features data have to be assured in order to refine the data matrix. Some methodologies as normalization, scaling and data transformation allow the removal of unwanted variabilities that occur due to both experimental (systematic human and instrumental errors during the analytical process) and biological (e.g. number and size of cells, concentration of biofluids…). These corrections can be grouped as method-driven (normalization based on internal/external standards and/or quality control samples) and data-driven (scaling and data transformation) [68]. The different approaches that can be

applied depending on the source of variability will not be discussed here. As a reference, the reviews from Dudzik et al., describing strategies for quality assurance in the hole untargeted metabolomic process [69], and from Li et al. about different refining methodologies [70] can be consulted. Nevertheless, it is worth noticing the need to include quality control samples (QCs, representative average sample formed by a pool of all samples analysed) in the metabolomic run (e.g. every 5 or 10 samples), to monitor the instrumental analysis, and for validating the features in the data matrix [71]. This surveillance could be applied for example: 1) to filter those features absent in a certain number of QC samples [39,72]; 2) correct intensity drifts caused by variations during the analysis, a common method is to apply the locally estimated smoothing function (LOESS) [26,73]; and 3) to determine the repeatability of each feature along the QCs, removing from the data matrix those with high relative standard deviation (% RSD) [31,74].

There is a wide range of informatics tools to perform this important part of the untargeted metabolomic process for xC-HRMS data. Whether they are free or commercial tools, the processing is mainly the same, although it may differ in how the steps are carried out, some of them working with in-house made algorithms. In the literature reviewed, the most employed tools were open-source software as XCMS (R package or Online) [17,18,75,76], MZmine [16,77] and MetAlign [78]; and commercial software as Mass Profiler Pro (Agilent Technologies Inc.) [79,80] , Progenesis QI (Non Linear Dynamics, Waters) [15,25,72], MarkerLynx (Waters) [81,82], SIEVE (Thermo Scientific) [37] and Compound Discoverer (Thermo Scientific) [83]. Some of them, as the open source MS-DIAL [84] or Compound Discoverer (Thermo Scientific) among others commercial tools, not only perform the abovementioned processes, but also the extraction and deconvolution of DDA and DIA spectral data and annotation by comparison of the deconvoluted MS/MS spectra with in house and/or public data bases, which is especially important for conventional DIA spectral data interpretation [49,85,86]. Because the GC-HRMS technique is less used in this area, most of the listed tools were developed for LC-HRMS data. However, their application to the GC-HRMS data appears to be equally powerful. The tools used were basically proprietary software as Chroma TOF (LECO) [27] or MetaboScape (Bruker Daltonik) [50], and freeware as MzMine [11,19] or MS-DIAL [29]. There are other tools that are gathering strength due to the good results obtained in this type of data, such as PARAFAC2 based Deconvolution and Identification System (PARADISe) [87].

6 STATISTICAL ANALYSIS

Although the aim when analysing data from foodomics studies seems quite simple: find the differences in the metabolite profiles related to the experimental design, the complexity and size of the data, the elevated number of metabolites and the natural biological variation of individuals make challenging this exploration. Multivariate data analysis is a powerful tool to explore correlations or co-variations in such datasets. This can be done with (supervised) or without (unsupervised) a priori knowledge about the experimental design [88]. Different tools have evolved during the past few years, but the most often used chemometric method in unsupervised analysis includes principal component analysis (PCA).

Many software options, free and commercial, are available for univariate and/or multivariate statistical tests. Among commercial software, SIMCA P+ (Soft independent modelling by class analogy) (Umetrics, Sweden) [12,23,25,28,47,89–91] is one of the most used, as well as its light version EzInfo, (U-Metrics, Sweden) [11,92]. Regarding the free software, MetaboAnalyst, which also provides a companion R package (MetaboAnalystR) to complement the web-based application, is a suitable option [57,86].

There are several aspects to consider before facing the modelling of metabolomics data where the number of variables largely exceeds the number of objects. Data cleaning by one or successive pre-filtration steps should be able to reduce the number of features and eliminate irrelevant signals while avoiding or minimizing relevant chemical information loss. The most commonly used are: i) removal of variables that exhibit a poor stability, meaning relative standard deviation (%RSD or CV%) on peak area across the QCs [91]; ii) removal of variables not present in a minimum number of the samples in one group; iii) removal of those variables that have zeros in a determined number of the samples (if it applies) (retain features with “nonzero values”) [93]; iv) removal of variables that show a low fold change or no significant difference among sample groups or among blank runs and any of the sample groups [75,78,94]. Multiple univariate analysis tools as pairwise t-test, ANOVA, etc are available at this point to determine significant differences. With one or more of these pre-filtration steps, a more robust dataset is obtained with still sufficient markers to enable a meaningful analysis. An improvement in the clusterization of the samples is also observed, with a tight clustering of the QC samples and an increase in the explained variance [10,83].

Once dataset is pre-filtered, PCA can be applied as a first step for interrogating the data in order to observe trends, grouping and/or outliers. PCA obtains new uncorrelated variables preforming linear combinations of the original ones, called principal components (PCs), according to common patterns and maximizing the variance in data. In this way, the dimensionality of the data is reduced while still preserving information from the original data set. The first PC represents the largest variation in the data set. The second PC, orthogonal to the first, covers as much of the residual variation as possible, and so on. Objects far apart in the score plot are different with respect to what patterns the model describes and objects in close proximity exhibit similar variations (**see Figure 2A**) [95].

In some cases, PCA is enough to determine if the classes can be predicted from the variables (discriminatory PCA) and to identify which ones are important in predicting class membership. PCA allowed the identification of markers potentially useful for the detection fraudulent use of chicken "dead on arrival" instead of normally slaughtered ones [10]. It was also successfully applied when discriminating between three different studied legumes in order to fight against food fraud [26]. PCA has also been used as exploratory tool previous to supervised analysis for gaining an in-depth understanding of the inherent differences among samples. In this line, PCA was able to suggest that metabolomic changes during milk fermentation by *L. helveticus H9* were more obvious at the fermentation phases (0–8 h), as PCA scores of earlier time points scattered away from those of the later time points (beyond 10 h) and this information was useful for further supervised analysis [92].

Unsupervised hierarchical cluster analysis (HCA), with a heatmap plot, can be used also as exploratory method to observe clusters, analyse and visualize the metabolome differences and/or to confirm the classification performed by PCA [26,56,78]. As an example, HCA could distinguish 5 main groups of metabolites among the 282 serum metabolites after the intake of milk and yogurt; 236 metabolites increased postprandially and 46 features decreased postprandially [72].

Supervised techniques can be very helpful for highlighting sample/group differences when PCA results are masked by high levels of spectral noise, strong batch effects, or high within group variation among other reasons. Partial least squares-discriminant analysis (PLS-DA) is a supervised method that uses multiple linear regression to find the direction of maximum

529 covariance between a data set and class labels. PLS-DA highlights the separation between groups
530 of observations and identifies variables that have most of the class separating information. As an
531 example, the clustering of malt and blended whiskies previously observed in PCA was
532 subsequently highlighted by the PLS-DA indicating that highly significant differences exist among
533 the two Scotch Whisky categories [33]. A variant of PLS-DA is orthogonal partial least squares-
534 discriminant analysis (OPLS-DA), where the variation in the data is divided in between-classes
535 and within-classes (predictive and non-predictive) that are forced to be described, by the first and
536 second OPLS-DA component. Although it does not alter the performance of the classification
537 model of a PLS-DA, it has an easier interpretation [96]. As an example, OPLS-DA was used to
538 develop a model enable to differentiate between no red meat intake and high red meat intake, in
539 serum samples [15].

540 However, there is no guarantee that the main variation extracted by the PCA is reflecting
541 the hypothesis put forward. PCA analysis failed to separate samples based on the production
542 system but highlighted the potential effect of the production year on a carrot metabolome study
543 (**Figure 2A**). The data was then subjected to OPLS-DA and the model was refined (**Figure 2B**).
544 Variables that contributed to the classification of samples based on production year were
545 investigated and removed from the datasets. This was crucial to improve predictive ability,
546 specificity and sensitivity values of the models [28].

547 Different methods exist to perform the selection of markers. From the PCA it can be done
548 using loading plot, the backbone of the PCA model. From the loading plot of PCA it was possible
549 to find out which metabolites mainly contribute to the separation of licorice samples from three
550 different origins and species [78]. From PLS-DA, Variable Influence on the Projection (VIP)
551 values > 1-2 generally represent those metabolites carrying the most relevant information for class
552 discrimination. From the OPLS-DA, a combination of VIP and p(corr) derived from the S-plot, is
553 a good strategy to identify metabolites with the highest influence on the group separation. VIP >
554 1.0 and p(corr) > 0.5 cut-off allowed the selection of most relevant metabolites detected in liver of
555 Wistar rats for the separation of the high-cholesterol (HC) and high-cholesterol enriched with
556 onion (HCO) feeding groups [90].

557 There are still few studies that only use univariate analysis for discrimination where a wide
558 number of different tests can be found [10]. However, a combination of outputs coming from

559 univariate and multivariate analysis is the most satisfactory and complete strategy for selection of
560 markers [29,97]. Regarding univariate analysis normally used, they can be divided among pair
561 tests (one-way ANOVA, Student's t-test, etc) and non-pair tests (Kruskal-Wallis, the Mann-
562 Whitney U-test, Welch t test, etc) depending on the normality of the data [29,36,57,79]. These
563 tests should be followed by a False Discovery Rate calculation p-FDR < 0.05 (q value set at 0.01)
564 normally applying Benjamini-Hochberg procedure to rectify p-values in order to correct for
565 multiple hypothesis testing and reduce the false positives than are susceptible to occurred when
566 the number of variables largely exceeds de number of objects [36,47,91].

567 As an example, combination of VIP>1, p(corr)≥0.5, with fold-change≥1.5 and p-value
568 <0.05 (one-way ANOVA) was used for biomarkers selection of discriminant macrophages
569 metabolites between control and high-dose group of Panax ginseng group [97]. Two-way ANOVA
570 analysis is a potent approach that can be added to discover the metabolites affected by two factors.
571 As an example, the level of *Lonicerae Japonicae Flos* and the administration days (time) were the
572 two factors that affected the metabolism of the rat [74]. The quality of the models is generally
573 evaluated by the goodness-of-fit parameter (R2X), the proportion of the variance of the response
574 variable that is explained by the model (R2Y) and the predictive ability parameter Q2. R2X, R2Y
575 and Q2 values close to 1 indicate an excellent model, and thus values higher than 0.5 indicate good
576 quality of PLS-DA and OPLS-DA models. However, it is remarkable that, contrary to PCA, these
577 supervised methods tend to overfit models and can generate excellent class separation even with
578 random data. For this reason, results of these types of tests should be critically checked and
579 properly cross-validated using procedures in which some of the samples are left out and their
580 classification have to be predicted. In order to test for possible overfitting and to confirm that Q2
581 values are stable and relevant, permutation tests are used. As an example, 7-fold full cross-
582 validation and permutation test on the responses (500 random permutations) were performed, in
583 order to avoid over-fitting and prove the robustness of the obtained models [81]. Some authors
584 works claim that findings need to be further verified using a higher number of samples, other
585 statistical tools and/or other analytical tools [56]. Apart from the classification rates obtained by
586 internal cross-validation (automatically performed by software like SIMCA P+), external
587 validations using samples that had not been used for the construction of the models is not such a
588 frequent practice but definitely adds value to the developed model [25,28]. Sales et al even
589 performed a reduction of variables until 15 to create and validate a model that could be used as

starting point for classification of future olive oil samples by quality following a simpler targeted analysis [11]. Chatterjee et al developed an LC-MS/MS assay with the set of 34 markers identified for rapid authentication of shrimps species and it was tested with unknown shrimp samples from the market [25]. In the field of biomarkers of food intake (BFI), an independent separate controlled, single-blinded, cross-over meal study was carried out to validate the candidate biomarkers of meat intake identified in a previous study resulting in a set of six better validated candidate markers that were further used to predict beef intake [16].

Another important aspect of the validation of biomarkers is the biological plausibility of such identified makers. Additionally, in the field of BFIs, examination of dose-response has become an essential prerequisite to demonstrate the use of biomarkers in dietary assessment for further applications in nutritional epidemiology. Subsequent confirmation and validation of biomarkers in intervention, independent studies, other cohorts, less-controlled, also adds evidence to the output [16,36,54].

7 (BIO)MARKERS IDENTIFICATION

Structural characterization and elucidation of potential markers highlighted in the statistical data analysis is commonly a challenge in metabolomics and can become the bottleneck of the overall metabolomics process. In HRMS, accurate mass measurement is the gold-standard for identification procedure and it is essential for facing this process. Q-TOF and OT-based HRMS analysers are more and more popular because of their high specificity, high resolution and low exact mass deviation [57]. The current methods and tools available for annotation of metabolites in untargeted metabolomics studies applying LC-MS platforms have been recently reviewed [98].

Chemical Analysis Working Group, within Metabolomics Standards Initiative (MSI), proposed four levels of confidence in metabolite identification: Level I is for identified/confirmed compounds, when their identity is validated using authentic standards and subsequent MS analysis; Level II is for putatively annotated compounds (e.g. without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries)); Level III is for putatively characterized compound classes (e.g. based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class); Level IV is for unknown compounds —

although unidentified or unclassified, these metabolites can still be differentiated and quantified based upon spectral data [99]. More recently Schymanski et al. have reported a similar system but including five-levels of confidence [100]. It has been updated including ion mobility separation as an additional parameter for more reliable identifications [101].

The first step in the identification workflow is to recognise the (quasi-)molecular ion in the accurate mass spectrum (typically, protonated or deprotonated molecule in LC-MS), where the presence of adducts must be also taken into account. Some tools, like CAMERA for XCMS or Progenesis QI, allow componentization, which means that different signals from the same metabolite are grouped together offering greater confidence to the annotation. Then, the most likely elemental composition is calculated according to the mass error and isotope pattern. After that, fragment ion information based on MS/MS or DIA experiments is used to establish the fragmentation pathways and discard possible chemical structures. To this aim, the use of offline/online and commercially/freely/in-house available spectral databases, are of great help [102]. The most used databases in the reviewed literature are: METLIN repository database (https://metlin.scripps.edu), Human Metabolome Database (HMDB) (http://www.hmdb.ca), Kyoto Encyclopedia of Genes and Genomes (https://www.genome.jp/kegg/), FooDB (https://foodb.ca), Chemspider (http://www.chemspider.com/), PubChem Compound database, LIPID Metabolites and Pathways Strategy LIPID MAPS (https://www.lipidmaps.org/), Chemical Entities of Biological Interest (ChEBI), SIRIUS, CSI:Finger ID. In-silico fragmentation tools are also useful in this process, emphasizing MetFrag (https://msbi.ipb-halle.de/MetFrag/), MetFusion, MassBank (http://www.massbank.jp), FooDB (http://foodb.ca) and Competitive Fragmentation Modeling for metabolite identification (CFM-ID) (http://cfmid.wishartlab.com/), among others.

Several software programs are available to automatize and simplify this challenging process: Progenesis QI, Compound Discoverer, MS-DIAL, CSI: FingerID, and MyCompoundID. However, the expertise and knowledge of the analyst on mass spectrometry and fragmentation rules is crucial to avoid false identifications. The injection of a reference standard, if commercially available, is the last step to assure the identity of the marker. When not available, the synthesis of the candidate compound may be required for full confirmation of the identity. Once the markers are identified, quantitative methods using standard substances can be developed to confirm that the specific markers do accurately reflect the differences between the classes. As an example of this identification workflow, chlorogenic acid was highlighted as potential biomarker for

Colombian coffees discrimination [48] according with the annotation performed after comparison with Metlin. Additional MS/MS experiments allowed the fragmentation evaluation for structure confirmation with the aid of in silico tools like Mefrag (**Figure 3**).

The availability of NIST library is a clear advantage for marker identification in metabolomics studies based on GC-EI-MS analysis. Although this library works mainly in nominal mass, a first step in the identification of metabolites is possible. It can be supported by the isotopic pattern, exact mass for parent ion (if exists) and fragments, and Kovats retention index. Soft ionization sources as positive chemical ionization (PCI) or atmospheric pressure chemical ionization (APCI) enable obtaining highly diagnostic molecular ions and/or protonated molecules of compounds which are extensively fragmented under EI conditions [11,33]. Compound identification by GC−MS can also be complemented using FiehnLib library and the Golm Metabolome Database.

Unfortunately, despite the efforts invested in biomarker identification, this final goal is not always achieved. However, the minimal requirements of reporting for unknown metabolites (retention time, prominent ion and fragment ion) can still be fulfilled [99]. This was the case of the study by Chatterjee et al., in which some markers only yielded one fragment ion, thus decreasing the reliability of the identification [25].

8 APPLICATIONS

In this section, we outline a selection of untargeted metabolomic studies that made use of LC and/or GC coupled to HRMS in the field of food processing, including authenticity, quality and safety[4,103]; the discovery of biomarkers of food intake [104–106]; and assessment of effects of food and diet on health [107–109].

8.1 Food Processing

The growing demands by our society, authorities and scientists to advance knowledge about the food consumed has led to the development of robust analytical methodologies to improve the quality and safety of food products and prevent food frauds. Several applications have been developed in the last years related to food authenticity. Some works were directed towards the identification of markers for characterization of food samples by its geographical origin in honey [57], garlic [14] or Adzuki Bean [30]; and for authentication of Protected Designation of Origin (PDO) of Grana Padano cheeses [56] and Colombian coffees [48]. Moreover, patterns of different

agricultural practices were assessed in carrots [28], different varieties of legumes [26], potato [110], *Vaccinium* fruits [94] and almonds [111], as well as different species identity, geographical origin and production method of commercially prawn and shrimps [25]. The characterization of food products by untargeted approach was also determinant to prevent fraudulent practices, such as the production of adulterated fruit juices in citrus [89], the production of counterfeited Scotch Whisky [33], and dead on arrival instead of the typical slathered poultry meat [10]. The characterization of organic culture practices against traditional cropping systems in wheat grains [23] has been also evaluated.

The characteristics of food appreciated by customers, including appearance, texture, flavour, aroma and nutritional composition, are also crucial in food quality, and are often dependent on subtle changes in the food's metabolome [4]. Regarding the appearance of food, it was established a relation between the colour of sea buckthorn and its chemical properties, having the red ones a better quality [13]. In other experience, the aim was to discover biomarkers related to the taste of food as it was the case with quality assessment of green tea [12] and olive oil [11]. Another important factor of food quality is to find markers related to storage time since the food quality worsens. Thus, significant differences in the metabolites composition of chilled chicken meat were found in accordance with conservation period [24].

Regarding food safety, there is a demand of robust markers for prevention of bad practices and possible errors in the food supply chain. [4]. In this line, new markers of egg ageing were found by an untargeted metabolic approach [83]. Regarding bad agricultural practices, a strategy was developed to discriminate green tea samples in concerning their contamination levels [75]. To ensure food safety and quality, it is important food traceability, which means continuous monitoring of the foods products through the entire supply chain, enabling the correction of mistakes. The role of HRMS-untargeted metabolomics in this context is the identification of characteristics markers of each stage of the process. In this way, metabolomics was found a powerful tool to identify different patterns between fresh tiger nut milk and milk processed by ultra-high temperature treatment [91]. It was also employed for the investigation of potential markers of three different species of licore plants (*Glycyrrhiza species*), which are sweetening and flavouring agents in food and beer industries [78].

8.2 Food Intake

A significant challenge in nutritional research is the measurement of dietary intake, which must be both accurate and applied to large numbers of people [112]. Traditionally, FFQ, 24 h recall or other dietary assessment tools, have been the standard tools for dietary assessment. Unfortunately, these approaches are subjected to errors such as underreporting, recall errors and difficulty in assessment of portion sizes, which generate biased and inaccurate results and associations. The scenery in nutritional epidemiology changed with the emergence of high-throughput metabolomics techniques enabling the discovery of novel biomarkers of food intake (BFIs) that represent objective measures of dietary and specific food intake.

In the literature, several examples can be found on intervention studies that employ untargeted chromatography coupled to HRMS-based metabolomics for the discovery of food-derived metabolites in banana [55], pea [54], fermented dairy products [72], different varieties of tomatoes [79] and tomato juice [77]. Biomarkers related to coffee consumption habits in various European countries were researched in a cross-sectional study within the European Prospective Investigation on Cancer and Nutrition (EPIC) [113]. In another observational study within the Singapore Prospective Study Program (SP2), patterns of meat and seafood consumption were assessed based on plasma metabolic profiles [19]. Regarding meat consumption, a great interest exists in finding indicators of red and processed meat intake since its consumption is associated with the development of chronic diseases [15,16,20,36].

The identification of biomarkers related to the intake of supplements suspected of having a benefit for human health is also another field of recent research. Several interventional studies with different bioactive foods and supplements have been performed: bioactive garlic [38], kiwi wine [27], beetroot juice [61], *angelica keiskei* [60], green coffee bean extract (GCBE) [81] and *amalaki rasayana* [39]. The metabolic patterns related with food enriched with some bioactive compound have been also investigated, as for example, flavan-3-ol-enriched dark chocolate, compared with standard dark chocolate and white chocolate [17] and apple juice enriched with four groups of polyphenols [37].

A better understanding of the relation between dietary patterns and metabolic profiles is crucial for improving the recommendations of health authorities about what diet is better for a better quality of life. New Nordic diet (NDD), which was designed to be balanced and healthy,

was compared to average Danish diet (ADD) in a long intervention study, identifying potential metabolic patterns that indicate potential health benefits of the NDD [114]. On the other hand, a detailed dietary assessed method was employed in the *Supplémentation en Vitamines et Minéraux AntioXidants* (SU.VI.MAX) cohort with the aim of performing a cross-sectional study and look at the difference in the plasma metabolic profiles according to their adherence to the French dietary recommendations [18].

As supported by several studies, HRMS-based untargeted metabolomics is a powerful approach in the discovery of new BIFs. Per definition, metabolomics fingerprinting is a data-driven approach, what means that a new hypothesis is forged from the biomarkers discovered. Therefore, all BIFs discovered by untargeted approach are tagged as "putative" since its necessary a proper validation process to confirm the association of robust BIFs to a specific food or diet. BIFs discovered in intervention studies used to be confirmed by the use of independent cohort studies (cross-sectional), as for example, Karlsruhe Metabolomics and Nutrition (KarMeN) [55] and EPIC study cohorts [20,36]. Other strategies include the use of dose-response for validation and independent study for confirmation [54]. In the case of potential BIFs identified only in cohort studies, these do not assess a correlation with the food consumed but rather an association and should be confirmed with an interventional study to validate them [112]. Nevertheless, as there is not an established standard methodology for validation of BFIs, L.O. Dragsted et al. proposed validation criteria based on analytical and biological aspects [115].

8.3 Food and Health Effects

Since metabolomics can provide a complete picture of the general dietary intake and reflect the current biological status of an individual, another goal of untargeted metabolomics in the nutrition field is to study the complex relationships between nutritional exposure and the positives or negatives effects on health/disease state [105,116]. The information obtained not only allows an accurate monitoring of a diet and lifestyle but may also help to design strategies to manipulate the physiological state with the ultimate goal to improve the individual health thought personalized dietary interventions [108,109].

Untargeted metabolomics approaches based on chromatography-HRMS have been applied to determine how a whole diet can affect the health state and to identify the molecular mechanism involved [41]. To this aim, both interventional, with human or mice/rat models, and observational

769 studies have been carried out. The objective was to determine the changes occurred in the
770 metabolism under a specific diet [66] or the differences obtained between 5 diets in different mice
771 tissues [117]. Showalter et al. performed a multiplatform untargeted metabolomics study [49],
772 finding significant metabolic alterations that suggest that the physiology of lungs can be altered by
773 obesity. In other studies, the goal was to determine the relationship between the diet and a specific
774 disease. For example, the use of multifunctional diet in order to reduce cardiometabolic risk factors
775 [47], the adherence to a healthy Nordic diet of a Swedish prospective cohort and the risk of future
776 type 2 diabetes [65]; the potential of ketogenic diet as an auxiliary cancer therapy with tumor
777 Xenograft mouse models [46]; and the correlation of diet with microbiota and metabolism of
778 inflammatory bowel disease human patients [118]. Given the diversity and complexity of diet
779 constituents, some studies were focused on one diet constituent and the effects on health, as fish
780 or coffee intake and type 2 diabetes risk [119,120]; or the health detriment due to the consumption
781 of heated soybean oil [121] or sweetened beverages [73].

782 There is a growing interest in nutraceuticals or diet supplements, especially in the so called
783 'functional foods', food products to which a health benefit is attributed (naturally or artificially
784 added) besides its own nutritional contribution [122,123]. Nevertheless, a wide variety of food
785 products are potentially beneficial for health and it might be difficult to determine if they can be
786 classified as 'functional'. For this reason, a notable number of untargeted metabolomic approaches
787 have been performed to determine the impact on health or disease of specific products considered
788 as functional foods: ginseng [82,97], herbal traditional medicines [59,74],wholegrain rye bread
789 [80], walnuts [85], lettuce [58] and onion [90]. The controlled trial study of fish oil
790 supplementation during pregnancy, ended in the detection of several altered metabolic pathways
791 significantly associated with a reduced risk of asthma by age 5 [124]. Likewise, the effects of diets
792 supplements was assessed, as selenium impact on metabolic disorders [93] or the use of
793 xenoestrogens in combination with cancer therapy [76].

794 Moreover, it is important the characterization of food products for their validation as
795 functional food and in order to enhance their potential. An untargeted lipidomic approach was
796 applied for the discovery of potentially high valuable polar lipids of *Porphyra dioica*, algae
797 commonly used for *sushi* preparation [63]. HRMS-based metabolomics was also used to study the
798 process of probiotic food product process as the dynamics of skim milk fermentation by *L.*

helveticus H9 strain [92] and the distinction between the biofilm and planktonic state B. bifidum strain [29].

It is worth noticing that the term functional food is usually applied to food products that have naturally or artificially substances known for their benefits on health, such as essential fatty acids, flavonoids, vitamins, polyphenols, etc. [122]. For this reason, determining whether a food is a functional product is often carried out through target analysis of the compounds that are known to have a beneficial effect. There is abundant bibliography available on targeted metabolomics in this field, which however does not fall within the scope of this review. It should be noted that target and untargeted metabolomics approaches can be combined, as for example to study the effects of white-blue light and dark in growth of cacao cell suspensions [86].

9 FUTURE PROSPECTS AND CONCLUSIONS

The use of MS-based approaches for untargeted metabolomics for investigation of food (bio)markers is still far from reaching its maximum potential. HRMS will surely be dominant in the near future, and the continuous improvements in instrumentation will be translated to enhanced capabilities of the developed strategies. For example, to maximize the metabolome coverage, it is necessary to acquire MS data in complementary chromatographic and ionization modes, but also MS/MS data, which can be acquired under DDA and/or DIA modes with sequential mass windows (e.g., SWATH, SONAR).

Regarding DDA, however, the metabolite coverage is not usually enough, and many features may lack MS/MS data. Technological evolution has allowed improvements in this acquisition mode and increasing the acquisition speed, which together with new software developments make possible to perform automated and iterative DDA in the newest instruments [125]. This strategy automates iterative exclusion and inclusion lists to reduce the fragmentation of redundant features coming from the background and allows performing exhaustive precursor selection obtaining more relevant MS^2 spectra. Such lists are automatically imported into the DDA method before the first ddMS^2 acquisition of the sample and are updated prior the next ddMS^2 run, bypassing precursors already fragmented to the exclusion list. MS^2 spectra are acquired for compounds remaining on the inclusion list. This approach enables to cover a wider range of compounds (including the lower-abundance ones) that were lost by the traditional DDA methods.

However, in order to not increase significantly the acquisition time, such strategy is only applied to the QC samples, as it requires multiple reinjection until reaching the complete coverage of the compounds. In this way, the samples are acquired in FS mode, and the iterative DDA is only applied to a reduced number of QC samples for future compound characterization.

In relation to DIA, the incorporation of ion mobility spectrometry (IMS) to HRMS has allowed a new DIA mode. An example is High Definition MS^E ($HDMS^E$). As occurs in conventional MS^E, two functions are acquired at low and high collision energy, but after ion mobility separation. In this way, the precursor and the product ions are recorded with the same drift time (translated into CCS, $Å^2$). This opens the possibility to filter the fragmentation spectra (obtained from all the ions fragmented in the scan cycle) by the drift time of a target ion and to obtain cleaner spectra without interfering fragments of co-eluting ions. Thus, the visualization of only the products ions from a specific precursor is feasible, enhancing the purity of the MS^2 spectrum with the inherent benefits of DIA acquisitions regarding available MS/MS data for all future biomarkers. The potential of this technique has been recently evaluated for orange dietary biomarkers discovery [64] and implemented for comparison of different polar lipids extraction methods to be used in evaluation of botanical origin, with potatoes as a case of study [110]. In both studies, data processing was performed using Progenesis QI (Waters), a unique software, able of performing the processing of 4D data obtained with xC-IMS-HRMS instruments.

In terms of ionization techniques, the recent atmospheric pressure CI source (APCI) is an attractive alternative to EI in GC-HRMS analysis. APCI enables a soft ionization ensuring the preservation of the (pseudo)-molecular ion, which is of great interest when the molecular ion is absent from the highly fragmented EI spectrum, which would imply a reduction in the selectivity and sensitivity. As APCI works under atmospheric pressure, the same mass analyser can be shared by both LC and GC instruments, since the vacuum does not need to be broken as occurs with EI and CI sources [126]. However, due to the novelty of this technique, there is a lack of spectral databases under this ionization source in comparison with EI. Only two articles using GC-APCI-HRMS have been found, both related to olive oil. Sales et al. studied the volatile composition of olive oil to develop a classification model for quality assessment [11], and Olmo-García et al. applied this technique for olive oil origin discrimination [50].

All in all, the combination of gas and liquid chromatography with high-resolution mass spectrometry, together with technological advances in instrumentation, both in chromatography (e.g. new stationary phases, format and particle size) and HRMS (e.g. resolution power, acquisition speed, MS2 acquisition modes) have been crucial to explain the impulse of untargeted metabolomics in the last few years. In particular, this approach has driven the expansion of knowledge on food processing, intake and the effects of food in health. The hyphenation of modern chromatography and HRMS allows a highly efficient separation combined with the acquisition of sensitive and high-quality structural compound information, facilitating the detection and identification of metabolites in complex biological samples, such as food matrices or biofluids. For this reason, this hyphenation has become one of the most used techniques in untargeted metabolomics studies in the field of food and nutrition. The implementation of chromatography-HRMS techniques, together with correct study designs and appropriate sample treatments, as well as the use of upgraded data treatment programs and powerful statistical tools, has notably enhanced the capabilities of untargeted metabolomics in the food field.

The increasing demand for more exhaustive control over food processing, in terms of authenticity, quality and safety, can be met, addressing needs such as the characterization of food products by geographical origin or production method, and the detection of adulteration or bad practices. Regarding nutrition, the application of untargeted metabolomics using chromatography-HRMS has revealed potential biomarkers related to the intake of food products and diets. Moreover, this approach can help to understand the complex relationships between nutritional exposure and physiological state, by the study of the effects of diet, or potentially beneficial food products, on the metabolism, as well as to evaluate the benefits to health.

ACKNOWLEDGMENTS

L Lacalle-Bergeron acknowledges the financial support of Universitat Jaume I, Spain for his pre-doctoral grant (UJI 19I001/03). T. Portolés acknowledges Ramon y Cajal Program from the Ministry of Economy and Competitiveness, Spain (RYC-2017-22525) for funding her research. D. Izquierdo-Sandoval acknowledges the Ministry of Science, Innovation and

Universities of Spain for funding his research through the FPU pre-doctoral program (FPU19/01839). The authors acknowledge the financial support of University Jaume I (UJI-B2020-37) and Generalitat Valenciana, as research group of excellence PROMETEO/2019/040.

FIGURE CAPTIONS

Figure 1. General overview and schematic content of the untargeted metabolomics workflow based on xC-HRMS analysis

Figure 2. Figure constructed from Cubero-Leon et al. [28]. (A) PCA score plot where the first and the second principal components (t1 and t2) are shown. Each harvested year is represented with a different symbol. In picture legend 1: year 2005; 2: year 2006; 3: year 2007; 4: year 2008. (B) Score plot of OPLS-DA of model 10. The first predictive component (t1) and the first orthogonal component (to1) are shown. R2Y: explained variation. Ellipse Hotelling's T2 (95%). Organic samples (filled circles), conventional samples (filled squares).

Figure 3. Figure from Hoyos-Ossa et al. [48]. MS/MS spectra at different collision energies of tentative marker chlorogenic acid and fragmentation explanation obtained with the aid of in silico fragmentation tool (MetFrag).

LITERATURE

[1] M. Herrero, C. Simó, V. García-Cañas, E. Ibáñez, A. Cifuentes, Foodomics: MS-based strategies in modern food science and nutrition, Mass Spectrom. Rev. 31 (2012) 49–69. https://doi.org/10.1002/mas.20335.

[2] J.-L. Wolfender, G. Marti, A. Thomas, S. Bertrand, Current approaches and challenges for the metabolite profiling of complex natural extracts, J. Chromatogr. A. 1382 (2015) 136–164. https://doi.org/10.1016/j.chroma.2014.10.091.

[3] P. Maruvada, J.W. Lampe, D.S. Wishart, D. Barupal, D.N. Chester, D. Dodd, Y. Djoumbou-Feunang, P.C. Dorrestein, L.O. Dragsted, J. Draper, L.C. Duffy, J.T. Dwyer, N.J. Emenaker, O. Fiehn, R.E. Gerszten, F. B Hu, R.W. Karp, D.M. Klurfeld, M.R. Laughlin, A.R. Little, C.J. Lynch, S.C. Moore, H.L. Nicastro, D.M. O'Brien, J.M. Ordovás, S.K. Osganian, M. Playdon, R. Prentice, D. Raftery, N. Reisdorph, H.M. Roche, S.A. Ross, S. Sang, A. Scalbert, P.R. Srinivas, S.H. Zeisel, Perspective: Dietary Biomarkers of Intake and Exposure - Exploration with Omics Approaches, Adv. Nutr. 11 (2020) 200–215. https://doi.org/10.1093/advances/nmz075.

[4] M. Castro-Puyana, R. Pérez-Míguez, L. Montero, M. Herrero, Application of mass spectrometry-based metabolomics approaches for food safety, quality and traceability, TrAC Trends Anal. Chem. 93 (2017) 102–118. https://doi.org/10.1016/j.trac.2017.05.004.

[5] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass spectrometry-based metabolomics, Mass Spectrom. Rev. 26 (2007) 51–78. https://doi.org/10.1002/mas.20108.

[6] A.E. Lytou, E.Z. Panagou, G.-J.E. Nychas, Volatilomics for food quality and authentication, Curr. Opin. Food Sci. 28 (2019) 88–95. https://doi.org/10.1016/j.cofs.2019.10.003.

[7] F. Vivanco, M.G. Barderas, C.M. Laborde, M. Posada, F. De La Cuesta, I. Zubiri, G. Alvarez-Llamas, Metabolomic profiling for identification of novel potential biomarkers in cardiovascular diseases, J. Biomed. Biotechnol. 2011 (2011) 1–9. https://doi.org/10.1155/2011/790132.

[8] H.G. Gika, G.A. Theodoridis, R.S. Plumb, I.D. Wilson, Current practice of liquid

chromatography–mass spectrometry in metabolomics and metabonomics, J. Pharm. Biomed. Anal. 87 (2014) 12–25. https://doi.org/10.1016/j.jpba.2013.06.032.

[9] W.B. Dunn, I.D. Wilson, A.W. Nicholls, D. Broadhurst, The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans, Bioanalysis. 4 (2012) 2249–2264. https://doi.org/10.4155/bio.12.204.

[10] K.L. Sidwick, A.E. Johnson, C.D. Adam, L. Pereira, D.F. Thompson, Use of Liquid Chromatography Quadrupole Time-of-Flight Mass Spectrometry and Metabonomic Profiling To Differentiate between Normally Slaughtered and Dead on Arrival Poultry Meat, Anal. Chem. 89 (2017) 12131–12136. https://doi.org/10.1021/acs.analchem.7b02749.

[11] C. Sales, M.I. Cervera, R. Gil, T. Portolés, E. Pitarch, J. Beltran, Quality classification of Spanish olive oils by untargeted gas chromatography coupled to hybrid quadrupole-time of flight mass spectrometry with atmospheric pressure chemical ionization and metabolomics-based statistical approach, Food Chem. 216 (2017) 365–373. https://doi.org/10.1016/j.foodchem.2016.08.033.

[12] J. Jing, Y. Shi, Q. Zhang, J. Wang, J. Ruan, Prediction of Chinese green tea ranking by metabolite profiling using ultra-performance liquid chromatography–quadrupole time-of-flight mass spectrometry (UPLC–Q-TOF/MS), Food Chem. 221 (2017) 311–316. https://doi.org/10.1016/j.foodchem.2016.10.068.

[13] C. He, G. Zhang, J. Zhang, Y. Zeng, J. Liu, Integrated analysis of multiomic data reveals the role of the antioxidant network in the quality of sea buckthorn berry, FASEB J. 31 (2017) 1929–1938. https://doi.org/10.1096/fj.201600974R.

[14] V. Hrbek, M. Rektorisova, H. Chmelarova, J. Ovesna, J. Hajslova, Authenticity assessment of garlic using a metabolomic approach based on high resolution mass spectrometry, J. Food Compos. Anal. 67 (2018) 19–28. https://doi.org/10.1016/j.jfca.2017.12.020.

[15] D. Carrizo, O.P. Chevallier, J. V. Woodside, S.F. Brennan, M.M. Cantwell, G. Cuskelly, C.T. Elliott, Untargeted metabolomic analysis of human serum samples associated with

different levels of red meat consumption: A possible indicator of type 2 diabetes?, Food Chem. 221 (2017) 214–221. https://doi.org/10.1016/j.foodchem.2016.10.056.

[16] C. Cuparencu, Å. Rinnan, L.O. Dragsted, Combined Markers to Assess Meat Intake—Human Metabolomic Studies of Discovery and Validation, Mol. Nutr. Food Res. 63 (2019) 1900106. https://doi.org/10.1002/mnfr.201900106.

[17] L.M. Ostertag, M. Philo, I.J. Colquhoun, H.S. Tapp, S. Saha, G.G. Duthie, E.K. Kemsley, B. de Roos, P.A. Kroon, G. Le Gall, Acute Consumption of Flavan-3-ol-Enriched Dark Chocolate Affects Human Endogenous Metabolism, J. Proteome Res. 16 (2017) 2516–2526. https://doi.org/10.1021/acs.jproteome.7b00089.

[18] L. Lécuyer, C. Dalle, P. Micheau, M. Pétéra, D. Centeno, B. Lyan, M. Lagree, P. Galan, S. Hercberg, A. Rossary, A. Demidem, M.-P. Vasson, V. Partula, M. Deschasaux, B. Srour, P. Latino-Martel, N. Druesne-Pecollo, E. Kesse-Guyot, S. Durand, E. Pujos-Guillot, C. Manach, M. Touvier, Untargeted plasma metabolomic profiles associated with overall diet in women from the SU.VI.MAX cohort, Eur. J. Nutr. (2020). https://doi.org/10.1007/s00394-020-02177-5.

[19] Y. Lu, L. Zou, J. Su, E. Tai, C. Whitton, R. van Dam, C. Ong, Meat and Seafood Consumption in Relation to Plasma Metabolic Profiles in a Chinese Population: A Combined Untargeted and Targeted Metabolomics Study, Nutrients. 9 (2017) 683. https://doi.org/10.3390/nu9070683.

[20] R. Wedekind, P. Keski-Rahkonen, N. Robinot, V. Viallon, P. Ferrari, E. Engel, M.-C. Boutron-Ruault, Y. Mahamat-Saleh, F.R. Mancini, T. Kühn, T. Johnson, H. Boeing, M. Bergmann, A. Karakatsani, A. Trichopoulou, H. Peppa, C. Agnoli, M. Santucci de Magistris, D. Palli, C. Sacerdote, R. Tumino, M.J. Gunter, I. Huybrechts, A. Scalbert, Syringol metabolites as new biomarkers for smoked meat intake, Am. J. Clin. Nutr. 110 (2019) 1424–1433. https://doi.org/10.1093/ajcn/nqz222.

[21] J. Pezzatti, J. Boccard, S. Codesido, Y. Gagnebin, A. Joshi, D. Picard, V. González-Ruiz, S. Rudaz, Implementation of liquid chromatography–high resolution mass spectrometry methods for untargeted metabolomic analyses of biological samples: A tutorial, Anal. Chim. Acta. 1105 (2020) 28–44. https://doi.org/10.1016/j.aca.2019.12.062.

[22] D. Vuckovic, Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry, Anal. Bioanal. Chem. 403 (2012) 1523–1548. https://doi.org/10.1007/s00216-012-6039-y.

[23] Y. Zhang, S. Cao, Z. Zhang, X. Meng, C. Hsiaoping, C. Yin, H. Jiang, S. Wang, Nutritional quality and health risks of wheat grains from organic and conventional cropping systems, Food Chem. 308 (2020) 125584. https://doi.org/10.1016/j.foodchem.2019.125584.

[24] D. Wen, Y. Liu, Q. Yu, Metabolomic approach to measuring quality of chilled chicken meat during storage, Poult. Sci. 99 (2020) 2543–2554. https://doi.org/10.1016/j.psj.2019.11.070.

[25] N.S. Chatterjee, O.P. Chevallier, E. Wielogorska, C. Black, C.T. Elliott, Simultaneous authentication of species identity and geographical origin of shrimps: Untargeted metabolomics to recurrent biomarker ions, J. Chromatogr. A. 1599 (2019) 75–84. https://doi.org/10.1016/j.chroma.2019.04.001.

[26] R. Llorach, C. Favari, D. Alonso, M. Garcia-Aloy, C. Andres-Lacueva, M. Urpi-Sarda, Comparative metabolite fingerprinting of legumes using LC-MS-based untargeted metabolomics, Food Res. Int. 126 (2019) 108666. https://doi.org/10.1016/j.foodres.2019.108666.

[27] Q. Zeng, H. Song, X. Xu, W. Mao, H. Xie, J. Liang, X. Chen, D. Chen, Y. Zhan, Health effects of kiwi wine on rats: an untargeted metabolic fingerprint study based on GC-MS/TOF, RSC Adv. 9 (2019) 13797–13807. https://doi.org/10.1039/C9RA02138H.

[28] E. Cubero-Leon, O. De Rudder, A. Maquet, Metabolomics for organic food authentication: Results from a long-term field study in carrots, Food Chem. 239 (2018) 760–770. https://doi.org/10.1016/j.foodchem.2017.06.161.

[29] F.A. Sadiq, B. Yan, J. Zhao, H. Zhang, W. Chen, Untargeted metabolomics reveals metabolic state of Bifidobacterium bifidum in the biofilm and planktonic states, LWT. 118 (2020) 108772. https://doi.org/10.1016/j.lwt.2019.108772.

[30] T.J. Kim, J.G. Park, S.K. Ahn, K.W. Kim, J. Choi, H.Y. Kim, S.-H. Ha, W.D. Seo, J.K.

Kim, Discrimination of Adzuki Bean (Vigna angularis) Geographical Origin by Targeted and Non-Targeted Metabolite Profiling with Gas Chromatography Time-of-Flight Mass Spectrometry, Metabolites. 10 (2020) 112. https://doi.org/10.3390/metabo10030112.

[31] M. Goettel, R. Niessner, D. Mueller, M. Scherer, G. Scherer, N. Pluym, Metabolomic fingerprinting in various body fluids of a diet- controlled clinical smoking cessation study using a validated GCTOF- MS metabolomics platform, J. Proteome Res. 16 (2017) 3491–3503. https://doi.org/10.1021/acs.jproteome.7b00128.

[32] A. Erban, I. Fehrle, F. Martinez-Seidel, F. Brigante, A.L. Más, V. Baroni, D. Wunderlin, J. Kopka, Discovery of food identity markers by metabolomics and machine learning technology, Sci. Rep. 9 (2019) 9697. https://doi.org/10.1038/s41598-019-46113-y.

[33] M. Stupak, I. Goodall, M. Tomaniova, J. Pulkrabova, J. Hajslova, A novel approach to assess the quality and authenticity of Scotch Whisky based on gas chromatography coupled to high resolution mass spectrometry, Anal. Chim. Acta. 1042 (2018) 60–70. https://doi.org/10.1016/j.aca.2018.09.017.

[34] É.A. Souza-Silva, R. Jiang, A. Rodríguez-Lafuente, E. Gionfriddo, J. Pawliszyn, A critical review of the state of the art of solid-phase microextraction of complex matrices I. Environmental analysis, TrAC Trends Anal. Chem. 71 (2015) 224–235. https://doi.org/10.1016/j.trac.2015.04.016.

[35] E.J. Want, I.D. Wilson, H. Gika, G. Theodoridis, R.S. Plumb, J. Shockcor, E. Holmes, J.K. Nicholson, Global metabolic profiling procedures for urine using UPLC–MS, Nat. Protoc. 5 (2010) 1005–1018. https://doi.org/10.1038/nprot.2010.50.

[36] W. Cheung, P. Keski-Rahkonen, N. Assi, P. Ferrari, H. Freisling, S. Rinaldi, N. Slimani, R. Zamora-Ros, M. Rundle, G. Frost, H. Gibbons, E. Carr, L. Brennan, A.J. Cross, V. Pala, S. Panico, C. Sacerdote, D. Palli, R. Tumino, T. Kühn, R. Kaaks, H. Boeing, A. Floegel, F. Mancini, M.-C. Boutron-Ruault, L. Baglietto, A. Trichopoulou, A. Naska, P. Orfanos, A. Scalbert, A metabolomic study of biomarkers of meat and fish intake, Am. J. Clin. Nutr. 105 (2017) 600–608. https://doi.org/10.3945/ajcn.116.146639.

[37] K. Trošt, M.M. Ulaszewska, J. Stanstrup, D. Albanese, C. De Filippo, K.M. Tuohy, F.

Natella, C. Scaccini, F. Mattivi, Host: Microbiome co-metabolic processing of dietary polyphenols – An acute, single blinded, cross-over study with different doses of apple polyphenols in healthy subjects, Food Res. Int. 112 (2018) 108–128. https://doi.org/10.1016/j.foodres.2018.06.016.

[38] Á. Fernández-Ochoa, I. Borrás-Linares, A. Baños, J.D. García-López, E. Guillamón, C. Nuñez-Lechado, R. Quirantes-Piné, A. Segura-Carretero, A fingerprinting metabolomic approach reveals deregulation of endogenous metabolites after the intake of a bioactive garlic supplement, J. Funct. Foods. 49 (2018) 137–145. https://doi.org/10.1016/j.jff.2018.08.003.

[39] V. Kumar, A.A. Kumar, V. Joseph, V.M. Dan, A. Jaleel, T.R.S. Kumar, C.C. Kartha, Untargeted metabolomics reveals alterations in metabolites of lipid metabolism and immune pathways in the serum of rats after long-term oral administration of Amalaki rasayana, Mol. Cell. Biochem. 463 (2020) 147–160. https://doi.org/10.1007/s11010-019-03637-1.

[40] R.A. Silva, T.C.S. Pereira, A.R. Souza, P.R. Ribeiro, 1H NMR-based metabolite profiling for biomarker identification, Clin. Chim. Acta. 502 (2020) 269–279. https://doi.org/10.1016/j.cca.2019.11.015.

[41] X. Zhang, Q. Li, Z. Xu, J. Dou, Mass spectrometry-based metabolomics in health and medical science: a systematic review, RSC Adv. 10 (2020) 3092–3104. https://doi.org/10.1039/C9RA08985C.

[42] F. Fenaille, P. Barbier Saint-Hilaire, K. Rousseau, C. Junot, Data acquisition workflows in liquid chromatography coupled to high resolution mass spectrometry-based metabolomics: Where do we stand?, J. Chromatogr. A. 1526 (2017) 1–12. https://doi.org/10.1016/j.chroma.2017.10.043.

[43] C. Junot, F. Fenaille, B. Colsch, F. Bécher, High resolution mass spectrometry based techniques at the crossroads of metabolic pathways, Mass Spectrom. Rev. 33 (2014) 471–500. https://doi.org/10.1002/mas.21401.

[44] D. Stettin, R.X. Poulin, G. Pohnert, Metabolomics Benefits from Orbitrap GC–MS—

Comparison of Low- and High-Resolution GC–MS, Metabolites. 10 (2020) 143. https://doi.org/10.3390/metabo10040143.

[45] J. Guo, T. Huan, Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography–Mass Spectrometry Based Untargeted Metabolomics, Anal. Chem. 92 (2020) 8072–8080. https://doi.org/10.1021/acs.analchem.9b05135.

[46] D. Licha, S. Vidali, S. Aminzadeh-Gohari, O. Alka, L. Breitkreuz, O. Kohlbacher, R.J. Reischl, R.G. Feichtinger, B. Kofler, C.G. Huber, Untargeted Metabolomics Reveals Molecular Effects of Ketogenic Diet on Healthy and Tumor Xenograft Mouse Models, Int. J. Mol. Sci. 20 (2019) 3873. https://doi.org/10.3390/ijms20163873.

[47] J. Tovar, V.D. de Mello, A. Nilsson, M. Johansson, J. Paananen, M. Lehtonen, K. Hanhineva, I. Björck, Reduction in cardiometabolic risk factors by a multifunctional diet is mediated via several branches of metabolism as evidenced by nontargeted metabolite profiling approach, Mol. Nutr. Food Res. 61 (2017) 1600552. https://doi.org/10.1002/mnfr.201600552.

[48] D.E. Hoyos Ossa, R. Gil-Solsona, G.A. Peñuela, J.V. Sancho, F.J. Hernández, Assessment of protected designation of origin for Colombian coffees based on HRMS-based metabolomics, Food Chem. 250 (2018) 89–97. https://doi.org/10.1016/j.foodchem.2018.01.038.

[49] M.R. Showalter, E.B. Nonnecke, A.L. Linderholm, T. Cajka, M.R. Sa, B. Lönnerdal, N.J. Kenyon, O. Fiehn, Obesogenic diets alter metabolism in mice, PLoS One. 13 (2018) e0190632. https://doi.org/10.1371/journal.pone.0190632.

[50] L. Olmo-García, K. Wendt, N. Kessler, A. Bajoub, A. Fernández-Gutiérrez, C. Baessmann, A. Carrasco-Pancorbo, Exploring the Capability of LC-MS and GC-MS Multi-Class Methods to Discriminate Virgin Olive Oils from Different Geographical Indications and to Identify Potential Origin Markers, Eur. J. Lipid Sci. Technol. 121 (2019) ejlt.201800336. https://doi.org/10.1002/ejlt.201800336.

[51] M. Khodadadi, M. Pourfarzam, A review of strategies for untargeted urinary metabolomic

analysis using gas chromatography–mass spectrometry, Metabolomics. 16 (2020) 1–14. https://doi.org/10.1007/s11306-020-01687-x.

[52] D.J. Beale, F.R. Pinu, K.A. Kouremenos, M.M. Poojary, V.K. Narayana, B.A. Boughton, K. Kanojia, S. Dayalan, O.A.H. Jones, D.A. Dias, Review of recent developments in GC–MS approaches to metabolomics-based research, Springer US, 2018. https://doi.org/10.1007/s11306-018-1449-2.

[53] B.B. Misra, M. Olivier, High Resolution GC-Orbitrap-MS Metabolomics Using Both Electron Ionization and Chemical Ionization for Analysis of Human Plasma, J. Proteome Res. (2020) acs.jproteome.9b00774. https://doi.org/10.1021/acs.jproteome.9b00774.

[54] P. S.C. Sri Harsha, R. Abdul Wahab, C. Cuparencu, L. Dragsted, L. Brennan, A Metabolomics Approach to the Identification of Urinary Biomarkers of Pea Intake, Nutrients. 10 (2018) 1911. https://doi.org/10.3390/nu10121911.

[55] N. Vázquez-Manjarrez, C.H. Weinert, M.M. Ulaszewska, C.I. Mack, P. Micheau, M. Pétéra, S. Durand, E. Pujos-Guillot, B. Egert, F. Mattivi, A. Bub, L.O. Dragsted, S.E. Kulling, C. Manach, Discovery and Validation of Banana Intake Biomarkers Using Untargeted Metabolomics in Human Intervention and Cross-sectional Studies, J. Nutr. 149 (2019) 1701–1713. https://doi.org/10.1093/jn/nxz125.

[56] G. Rocchetti, L. Lucini, A. Gallo, F. Masoero, M. Trevisan, G. Giuberti, Untargeted metabolomics reveals differences in chemical fingerprints between PDO and non-PDO Grana Padano cheeses, Food Res. Int. 113 (2018) 407–413. https://doi.org/10.1016/j.foodres.2018.07.029.

[57] Y. Li, Y. Jin, S. Yang, W. Zhang, J. Zhang, W. Zhao, L. Chen, Y. Wen, Y. Zhang, K. Lu, Y. Zhang, J. Zhou, S. Yang, Strategy for comparative untargeted metabolomics reveals honey markers of different floral and geographic origins using ultrahigh-performance liquid chromatography-hybrid quadrupole-orbitrap mass spectrometry, J. Chromatogr. A. 1499 (2017) 78–89. https://doi.org/10.1016/j.chroma.2017.03.071.

[58] H. Ismail, A.L. Gillespie, D. Calderwood, H. Iqbal, C. Gallagher, O.P. Chevallier, C.T. Elliott, X. Pan, B. Mirza, B.D. Green, The Health Promoting Bioactivities of Lactuca

sativa can be Enhanced by Genetic Modulation of Plant Secondary Metabolites, Metabolites. 9 (2019) 97. https://doi.org/10.3390/metabo9050097.

[59] M. Wei, Z. Liu, Y. Liu, S. Li, M. Hu, K. Yue, T. Liu, Y. He, Z. Pi, Z. Liu, F. Song, Urinary and plasmatic metabolomics strategy to explore the holistic mechanism of lignans in S. chinensis in treating Alzheimer's disease using UPLC-Q-TOF-MS, Food Funct. 10 (2019) 5656–5668. https://doi.org/10.1039/C9FO00677J.

[60] H.-A. Oh, H. Lee, S. Park, Y. Lim, O. Kwon, J.Y. Kim, D. Kim, B.H. Jung, Analysis of plasma metabolic profiling and evaluation of the effect of the intake of Angelica keiskei using metabolomics and lipidomics, J. Ethnopharmacol. 243 (2019) 112058. https://doi.org/10.1016/j.jep.2019.112058.

[61] L. Kozlowska, O. Mizera, A. Mroz, An Untargeted Metabolomics Approach to Investigate the Metabolic Effect of Beetroot Juice Supplementation in Fencers—A Preliminary Study, Metabolites. 10 (2020) 100. https://doi.org/10.3390/metabo10030100.

[62] D.K. Trivedi, R.K. Iles, Do not just do it, do it right: urinary metabolomics -establishing clinically relevant baselines, Biomed. Chromatogr. 28 (2014) 1491–1501. https://doi.org/10.1002/bmc.3219.

[63] E. da Costa, V. Azevedo, T. Melo, A. Rego, D. V. Evtuguin, P. Domingues, R. Calado, R. Pereira, M. Abreu, M. Domingues, High-Resolution Lipidomics of the Early Life Stages of the Red Seaweed Porphyra dioica, Molecules. 23 (2018) 187. https://doi.org/10.3390/molecules23010187.

[64] L. Lacalle-Bergeron, T. Portolés, F.J. López, J.V. Sancho, C. Ortega-Azorín, E.M. Asensio, O. Coltell, D. Corella, Ultra-Performance Liquid Chromatography-Ion Mobility Separation-Quadruple Time-of-Flight MS (UHPLC-IMS-QTOF MS) Metabolomics for Short-Term Biomarker Discovery of Orange Intake: A Randomized, Controlled Crossover Study, Nutrients. 12 (2020) 1916. https://doi.org/10.3390/nu12071916.

[65] L. Shi, C. Brunius, I. Johansson, I.A. Bergdahl, B. Lindahl, K. Hanhineva, R. Landberg, Plasma metabolites associated with healthy Nordic dietary indexes and risk of type 2 diabetes—a nested case-control study in a Swedish population, Am. J. Clin. Nutr. 108

(2018) 564–575. https://doi.org/10.1093/ajcn/nqy145.

[66] S. Heischmann, L.B. Gano, K. Quinn, L.-P. Liang, J. Klepacki, U. Christians, N. Reisdorph, M. Patel, Regulation of kynurenine metabolism by a ketogenic diet, J. Lipid Res. 59 (2018) 958–966. https://doi.org/10.1194/jlr.M079251.

[67] R. Pérez-Míguez, M. Castro-Puyana, E. Sánchez-López, M. Plaza, M.L. Marina, Untargeted HILIC-MS-Based Metabolomics Approach to Evaluate Coffee Roasting Process: Contributing to an Integrated Metabolomics Multiplatform, Molecules. 25 (2020) 887. https://doi.org/10.3390/molecules25040887.

[68] A.M. De Livera, D.A. Dias, D. De Souza, T. Rupasinghe, J. Pyke, D. Tull, U. Roessner, M. McConville, T.P. Speed, Normalizing and integrating metabolomics data, Anal. Chem. 84 (2012) 10768–10776. https://doi.org/10.1021/ac302748b.

[69] D. Dudzik, C. Barbas-Bernardos, A. García, C. Barbas, Quality assurance procedures for mass spectrometry untargeted metabolomics. a review, J. Pharm. Biomed. Anal. 147 (2018) 149–173. https://doi.org/10.1016/j.jpba.2017.07.044.

[70] B. Li, J. Tang, Q. Yang, X. Cui, S. Li, S. Chen, Q. Cao, W. Xue, N. Chen, F. Zhu, Performance evaluation and online realization of data-driven normalization methods used in LC/MS based untargeted metabolomics analysis, Sci. Rep. 6 (2016) 1–13. https://doi.org/10.1038/srep38881.

[71] E. Gorrochategui, J. Jaumot, S. Lacorte, R. Tauler, Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: Overview and workflow, TrAC Trends Anal. Chem. 82 (2016) 425–442. https://doi.org/10.1016/j.trac.2016.07.004.

[72] G. Pimentel, K.J. Burton, U. von Ah, U. Bütikofer, F.P. Pralong, N. Vionnet, R. Portmann, G. Vergères, Metabolic Footprinting of Fermented Milk Consumption in Serum of Healthy Men, J. Nutr. 148 (2018) 851–860. https://doi.org/10.1093/jn/nxy053.

[73] W. Perng, L. Tang, P.X.K. Song, M. Goran, M.M. Tellez Rojo, A. Cantoral, K.E. Peterson, Urate and Nonanoate Mark the Relationship between Sugar-Sweetened Beverage Intake and Blood Pressure in Adolescent Girls: A Metabolomics Analysis in the ELEMENT Cohort, Metabolites. 9 (2019) 100. https://doi.org/10.3390/metabo9050100.

[74] L. Li, S. Ma, D. Wang, L. Chen, X. Wang, Plasma metabolomics analysis of endogenous and exogenous metabolites in the rat after administration of Lonicerae Japonicae Flos, Biomed. Chromatogr. 34 (2020). https://doi.org/10.1002/bmc.4773.

[75] G. Delaporte, M. Cladière, D. Jouan-Rimbaud Bouveresse, V. Camel, Untargeted food contaminant detection using UHPLC-HRMS combined with multivariate analysis: Feasibility study on tea, Food Chem. 277 (2019) 54–62. https://doi.org/10.1016/j.foodchem.2018.10.089.

[76] B. Warth, P. Raffeiner, A. Granados, T. Huan, M. Fang, E.M. Forsberg, H.P. Benton, L. Goetz, C.H. Johnson, G. Siuzdak, Metabolomics Reveals that Dietary Xenoestrogens Alter Cellular Metabolism Induced by Palbociclib/Letrozole Combination Cancer Therapy, Cell Chem. Biol. 25 (2018) 291-300.e3. https://doi.org/10.1016/j.chembiol.2017.12.010.

[77] Y. Hövelmann, A. Jagels, R. Schmid, F. Hübner, H.-U. Humpf, Identification of potential human urinary biomarkers for tomato juice intake by mass spectrometry-based metabolomics, Eur. J. Nutr. 59 (2020) 685–697. https://doi.org/10.1007/s00394-019-01935-4.

[78] G. Rizzato, E. Scalabrin, M. Radaelli, G. Capodaglio, O. Piccolo, A new exploration of licorice metabolome, Food Chem. 221 (2017) 959–968. https://doi.org/10.1016/j.foodchem.2016.11.068.

[79] M.J. Cichon, K.M. Riedl, L. Wan, J.M. Thomas-Ahner, D.M. Francis, S.K. Clinton, S.J. Schwartz, Plasma Metabolomics Reveals Steroidal Alkaloids as Novel Biomarkers of Tomato Intake in Mice, Mol. Nutr. Food Res. 61 (2017) 1700241. https://doi.org/10.1002/mnfr.201700241.

[80] P. Keski-Rahkonen, M. Kolehmainen, J. Lappi, V. Micard, J. Jokkala, N. Rosa-Sibakov, J. Pihlajamäki, P. V. Kirjavainen, H. Mykkänen, K. Poutanen, M.J. Gunter, A. Scalbert, K. Hanhineva, Decreased plasma serotonin and other metabolite changes in healthy adults after consumption of wholegrain rye: an untargeted metabolomics study, Am. J. Clin. Nutr. 109 (2019) 1630–1639. https://doi.org/10.1093/ajcn/nqy394.

[81] G. Peron, D. Santarossa, D. Voinovich, S. Dall'Acqua, S. Sut, Urine metabolomics shows an induction of fatty acids metabolism in healthy adult volunteers after supplementation with green coffee (Coffea robusta L.) bean extract, Phytomedicine. 38 (2018) 74–83. https://doi.org/10.1016/j.phymed.2017.11.002.

[82] Z. Qin, C. Jia, D. Liao, X. Chen, X. Li, Comparison of Serum Metabolite Changes of Radiated Mice Administered with Panax quinquefolium from Different Cultivation Regions Using UPLC-Q/TOF-MS Based Metabolomic Approach, Molecules. 23 (2018) 1014. https://doi.org/10.3390/molecules23051014.

[83] D. Cavanna, D. Catellani, C. Dall'Asta, M. Suman, Egg product freshness evaluation: A metabolomic approach, J. Mass Spectrom. 53 (2018) 849–861. https://doi.org/10.1002/jms.4256.

[84] H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, M. Arita, MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, Nat. Methods. 12 (2015) 523–526. https://doi.org/10.1038/nmeth.3393.

[85] M. Nakanishi, A. Matz, C. Klemashevich, D.W. Rosenberg, Dietary Walnut Supplementation Alters Mucosal Metabolite Profiles During DSS-Induced Colonic Ulceration, Nutrients. 11 (2019) 1118. https://doi.org/10.3390/nu11051118.

[86] A.M. Gallego, L.F. Rojas, H.A. Rodriguez, C. Mora, L. Atehortúa, A.I. Urrea, M.J. Guiltinan, S.N. Maximova, E. Gaquerel, M. Zuluaga, N. Pabón-Mora, Metabolomic profile of cacao cell suspensions growing in blue light/dark conditions with potential in food biotechnology, Plant Cell, Tissue Organ Cult. 139 (2019) 275–294. https://doi.org/10.1007/s11240-019-01679-3.

[87] L.G. Johnsen, P.B. Skou, B. Khakimov, R. Bro, Gas chromatography – mass spectrometry data processing made easy, J. Chromatogr. A. 1503 (2017) 57–64. https://doi.org/10.1016/j.chroma.2017.04.052.

[88] T. Skov, A.H. Honoré, H.M. Jensen, T. Næs, S.B. Engelsen, Chemometrics in foodomics: Handling data structures from multiple analytical platforms, TrAC Trends Anal. Chem. 60

(2014) 71–79. https://doi.org/10.1016/j.trac.2014.05.004.

[89] Z. Jandrić, M. Islam, D.K. Singh, A. Cannavan, Authentication of Indian citrus fruit/fruit juices by untargeted and targeted metabolomics, Food Control. 72 (2017) 181–188. https://doi.org/10.1016/j.foodcont.2015.10.044.

[90] D. González-Peña, D. Dudzik, A. García, B. Ancos, C. Barbas, C. Sánchez-Moreno, Metabolomic Fingerprinting in the Comprehensive Study of Liver Changes Associated with Onion Supplementation in Hypercholesterolemic Wistar Rats, Int. J. Mol. Sci. 18 (2017) 267. https://doi.org/10.3390/ijms18020267.

[91] J. Rubert, A. Monforte, K. Hurkova, G. Pérez-Martínez, J. Blesa, J.L. Navarro, M. Stranka, J.M. Soriano, J. Hajslova, Untargeted metabolomics of fresh and heat treatment Tiger nut (Cyperus esculentus L.) milks reveals further insight into food quality and nutrition, J. Chromatogr. A. 1514 (2017) 80–87. https://doi.org/10.1016/j.chroma.2017.07.071.

[92] Z. Mi, L. Kwok, J. Xue, Y. Wang, H. Zhang, Y. Chen, Fermentation dynamics of Lactobacillus helveticus H9 revealed by ultra-performance liquid chromatography quadrupole time-of-flight mass spectrometry, Int. J. Food Sci. Technol. 53 (2018) 1442–1451. https://doi.org/10.1111/ijfs.13723.

[93] X. Hu, J.D. Chandler, M.L. Orr, L. Hao, K. Liu, K. Uppal, Y.-M. Go, D.P. Jones, Selenium Supplementation Alters Hepatic Energy and Fatty Acid Metabolism in Mice, J. Nutr. 148 (2018) 675–684. https://doi.org/10.1093/jn/nxy036.

[94] K. Hurkova, L. Uttl, J. Rubert, K. Navratilova, V. Kocourek, M. Stranska-Zachariasova, F. Paprstein, J. Hajslova, Cranberries versus lingonberries: A challenging authentication of similar Vaccinium fruit, Food Chem. 284 (2019) 162–170. https://doi.org/10.1016/j.foodchem.2019.01.014.

[95] T. Skov, S.B. Engelsen, Chemometrics, Mass Spectrometry, and Foodomics, in: Foodomics, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2013: pp. 507–538. https://doi.org/10.1002/9781118537282.ch19.

[96] M. Bevilacqua, R. Bro, F. Marini, Å. Rinnan, M.A. Rasmussen, T. Skov, Recent

chemometrics advances for foodomics, TrAC Trends Anal. Chem. 96 (2017) 42–51. https://doi.org/10.1016/j.trac.2017.08.011.

[97] J. Hao, H. Hu, J. Liu, X. Wang, X. Liu, J. Wang, M. Niu, Y. Zhao, X. Xiao, Integrated Metabolomics and Network Pharmacology Study on Immunoregulation Mechanisms of Panax ginseng through Macrophages, Evidence-Based Complement. Altern. Med. 2019 (2019) 1–14. https://doi.org/10.1155/2019/3630260.

[98] W.J. Nash, W.B. Dunn, From mass to metabolite in human untargeted metabolomics: Recent advances in annotation of metabolites applying liquid chromatography-mass spectrometry data, TrAC Trends Anal. Chem. 120 (2019) 115324. https://doi.org/10.1016/j.trac.2018.11.022.

[99] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W.M. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A.N. Lane, J.C. Lindon, P. Marriott, A.W. Nicholls, M.D. Reily, J.J. Thaden, M.R. Viant, Proposed minimum reporting standards for chemical analysis, Metabolomics. 3 (2007) 211–221. https://doi.org/10.1007/s11306-007-0082-2.

[100] 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.

[101] A. Celma, J. V. Sancho, E.L. Schymanski, D. Fabregat-Safont, M. Ibáñez, J. Goshawk, G. Barknowitz, F. Hernández, L. Bijlsma, Improving Target and Suspect Screening High-Resolution Mass Spectrometry Workflows in Environmental Analysis by Ion Mobility Separation, Environ. Sci. Technol. (2020). https://doi.org/10.1021/acs.est.0c05713.

[102] M. Vinaixa, E.L. Schymanski, S. Neumann, M. Navarro, R.M. Salek, O. Yanes, Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects, TrAC Trends Anal. Chem. 78 (2016) 23–35. https://doi.org/10.1016/j.trac.2015.09.005.

[103] K. Böhme, P. Calo-Mata, J. Barros-Velázquez, I. Ortea, Recent applications of omics-

based technologies to main topics in food authentication, TrAC Trends Anal. Chem. 110 (2019) 221–232. https://doi.org/10.1016/j.trac.2018.11.005.

[104] L. Brennan, F.B. Hu, Metabolomics-Based Dietary Biomarkers in Nutritional Epidemiology—Current Status and Future Opportunities, Mol. Nutr. Food Res. 63 (2019) 1–5. https://doi.org/10.1002/mnfr.201701064.

[105] M. Guasch-Ferré, S.N. Bhupathiraju, F.B. Hu, Use of Metabolomics in Improving Assessment of Dietary Intake, Clin. Chem. 64 (2018) 82–98. https://doi.org/10.1373/clinchem.2017.272344.

[106] C. Collins, A.E. McNamara, L. Brennan, Role of metabolomics in identification of biomarkers related to food intake, Proc. Nutr. Soc. 78 (2019) 189–196. https://doi.org/10.1017/S002966511900048X.

[107] V. Özdemir, E. Kolker, Precision Nutrition 4.0: A Big Data and Ethics Foresight Analysis-Convergence of Agrigenomics, Nutrigenomics, Nutriproteomics, and Nutrimetabolomics, Omi. A J. Integr. Biol. 20 (2016) 69–75. https://doi.org/10.1089/omi.2015.0193.

[108] D. Braconi, G. Bernardini, L. Millucci, A. Santucci, Foodomics for human health: current status and perspectives, Expert Rev. Proteomics. 15 (2018) 153–164. https://doi.org/10.1080/14789450.2018.1421072.

[109] G. Mancano, M. Mora-Ortiz, S.P. Claus, Recent developments in nutrimetabolomics: from food characterisation to disease prevention, Curr. Opin. Food Sci. 22 (2018) 145–152. https://doi.org/10.1016/j.cofs.2018.03.012.

[110] C. Claassen, J. Kuballa, S. Rohn, Polar Lipids in Starch-Rich Commodities to be Analyzed with LC-MS-Based Metabolomics—Optimization of Ionization Parameters and High-Throughput Extraction Protocols, Metabolites. 9 (2019) 167. https://doi.org/10.3390/metabo9080167.

[111] R. Gil Solsona, C. Boix, M. Ibáñez, J. V. Sancho, The classification of almonds (Prunus dulcis) by country and variety using UHPLC-HRMS-based untargeted metabolomics, Food Addit. Contam. Part A. 35 (2018) 395–403.

https://doi.org/10.1080/19440049.2017.1416679.

[112] A. O'Gorman, L. Brennan, The role of metabolomics in determination of new dietary biomarkers, Proc. Nutr. Soc. 76 (2017) 295–302. https://doi.org/10.1017/S0029665116002974.

[113] J.A. Rothwell, P. Keski-Rahkonen, N. Robinot, N. Assi, C. Casagrande, M. Jenab, P. Ferrari, M.C. Boutron-Ruault, Y. Mahamat-Saleh, F.R. Mancini, H. Boeing, V. Katzke, T. Kühn, K. Niforou, A. Trichopoulou, E. Valanou, V. Krogh, A. Mattiello, D. Palli, C. Sacerdote, R. Tumino, A. Scalbert, A Metabolomic Study of Biomarkers of Habitual Coffee Intake in Four European Countries, Mol. Nutr. Food Res. 63 (2019) 1–10. https://doi.org/10.1002/mnfr.201900659.

[114] E. Acar, G. Gürdeniz, B. Khakimov, F. Savorani, S.K. Korndal, T.M. Larsen, S.B. Engelsen, A. Astrup, L.O. Dragsted, Biomarkers of Individual Foods, and Separation of Diets Using Untargeted LC-MS-based Plasma Metabolomics in a Randomized Controlled Trial, Mol. Nutr. Food Res. 63 (2019) 1800215. https://doi.org/10.1002/mnfr.201800215.

[115] L.O. Dragsted, Q. Gao, A. Scalbert, G. Vergères, M. Kolehmainen, C. Manach, L. Brennan, L.A. Afman, D.S. Wishart, C. Andres Lacueva, M. Garcia-Aloy, H. Verhagen, E.J.M. Feskens, G. Praticò, Validation of biomarkers of food intake-Critical assessment of candidate biomarkers, Genes Nutr. 13 (2018) 1–14. https://doi.org/10.1186/s12263-018-0603-9.

[116] C. Picó, F. Serra, A.M. Rodríguez, J. Keijer, A. Palou, Biomarkers of Nutrition and Health: New Tools for New Approaches, Nutrients. 11 (2019) 1092. https://doi.org/10.3390/nu11051092.

[117] A. Wells, W.T. Barrington, S. Dearth, A. May, D.W. Threadgill, S.R. Campagna, B.H. Voy, Tissue Level Diet and Sex-by-Diet Interactions Reveal Unique Metabolite and Clustering Profiles Using Untargeted Liquid Chromatography-Mass Spectrometry on Adipose, Skeletal Muscle, and Liver Tissue in C57BL6/J Mice, J. Proteome Res. 17 (2018) 1077–1090. https://doi.org/10.1021/acs.jproteome.7b00750.

[118] Y.J. Weng, H.Y. Gan, X. Li, Y. Huang, Z.C. Li, H.M. Deng, S.Z. Chen, Y. Zhou, L.S.

Wang, Y.P. Han, Y.F. Tan, Y.J. Song, Z.M. Du, Y.Y. Liu, Y. Wang, N. Qin, Y. Bai, R.F. Yang, Y.J. Bi, F.C. Zhi, Correlation of diet, microbiota and metabolite networks in inflammatory bowel disease, J. Dig. Dis. 20 (2019) 447–459. https://doi.org/10.1111/1751-2980.12795.

[119] L. Shi, C. Brunius, I.A. Bergdahl, I. Johansson, O. Rolandsson, C. Donat Vargas, H. Kiviranta, K. Hanhineva, A. Åkesson, R. Landberg, Joint Analysis of Metabolite Markers of Fish Intake and Persistent Organic Pollutants in Relation to Type 2 Diabetes Risk in Swedish Adults, J. Nutr. 149 (2019) 1413–1423. https://doi.org/10.1093/jn/nxz068.

[120] L. Shi, C. Brunius, I. Johansson, I.A. Bergdahl, O. Rolandsson, B. Guelpen, A. Winkvist, K. Hanhineva, R. Landberg, Plasma metabolite biomarkers of boiled and filtered coffee intake and their association with type 2 diabetes risk, J. Intern. Med. 287 (2020) 405–421. https://doi.org/10.1111/joim.13009.

[121] L. Wang, D. Yao, P.E. Urriola, A.R. Hanson, M. Saqui-Salces, B.J. Kerr, G.C. Shurson, C. Chen, Identification of activation of tryptophan–NAD+ pathway as a prominent metabolic response to thermally oxidized oil through metabolomics-guided biochemical analysis, J. Nutr. Biochem. 57 (2018) 255–267. https://doi.org/10.1016/j.jnutbio.2018.04.009.

[122] S.P. Claus, Development of personalized functional foods needs metabolic profiling, Curr. Opin. Clin. Nutr. Metab. Care. 17 (2014) 567–573. https://doi.org/10.1097/MCO.0000000000000107.

[123] A. Valdés, A. Cifuentes, C. León, Foodomics evaluation of bioactive compounds in foods, TrAC Trends Anal. Chem. 96 (2017) 2–13. https://doi.org/10.1016/j.trac.2017.06.004.

[124] D. Rago, M.A. Rasmussen, K.A. Lee-Sarwar, S.T. Weiss, J. Lasky-Su, J. Stokholm, K. Bønnelykke, B.L. Chawes, H. Bisgaard, Fish-oil supplementation in pregnancy, child metabolomics and asthma risk, EBioMedicine. 46 (2019) 399–410. https://doi.org/10.1016/j.ebiom.2019.07.057.

[125] Select Science, thermo scientific, The ultimate workflow for small molecule discovery, (n.d.).

[126] F. Hernández, M. Ibáñez, T. Portolés, M.I. Cervera, J. V. Sancho, F.J. López, Advancing towards universal screening for organic pollutants in waters, J. Hazard. Mater. 282 (2015) 86–95. https://doi.org/10.1016/j.jhazmat.2014.08.006.