Liquid Chromatographic Techniques in Food Sciences

Review

A Review on the Foodomics Based on Liquid Chromatography Mass Spectrometry

Yoshio Muguruma, Mari Nunome, and Koichi Inoue*

Graduate School of Pharmaceutical Sciences, Ritsumeikan University; 1–1–1 Nojihigashi, Kusatsu, Shiga 525–8577, Japan.

Received September 1, 2021

Due to the globalization of food production and distribution, the food chain has become increasingly complex, making it more difficult to evaluate unexpected food changes. Therefore, establishing sensitive, robust, and cost-effective analytical platforms to efficiently extract and analyze the food-chemicals in complex food matrices is essential, however, challenging. LC/MS-based metabolomics is the key to obtain a broad overview of human metabolism and understand novel food science. Various metabolomics approaches (e.g., targeted and/or untargeted) and sample preparation techniques in food analysis have their own advantages and limitations. Selecting an analytical platform that matches the characteristics of the analytes is important for food analysis. This review highlighted the recent trends and applications of metabolomics based on “foodomics” by LC-MS and provides the perspectives and insights into the methodology and various sample preparation techniques in food analysis.

Key words LC-MS; foodomics; metabolomics; sample preparation; food analysis

1. Introduction

At present, interest in food safety and quality has been increasing with regards to food intake in our daily lives. Daily food intake supplies energy and nutrition that supports the human body and mediates complementary components for human health and growth. Maintaining homeostasis for human health and preventing diseases through food intake is attractive to consumers and the food industry. Due to the globalization of food production and distribution, the food chain has become increasingly complex, making it more difficult to track unexpected changes in food products. Therefore, the identification and quantification of food-molecules should be important for these risk/safety assessments and regulatory controls. There are concerns regarding environmental and artificial interferences, such as food adulteration and contamination of food materials and by-products from microbial organisms, organic/inorganic chemical pollutants, and drugs. According to food attractive in worldwide, it would be well worth to study an evidence-based assessment and definition of food-molecules such as carbohydrate, pesticide, and others for public health risks related to food safety issues in the food market. Facilely, it is needed to develop and upload the analytical techniques regarding to food sciences for the regulatory assessment.

At present, there is also increasing attention on significant food products, such as contaminated, allergic and genetically modified (GM) materials, because they play important roles of regulatory assessments and monitoring evaluations. Food safety issue is supervised in the production, distribution, and consumption of processed and agricultural commodities all over the world. Recently, the presences of mycotoxins, allergenic additives, GM plants, multiple pesticide residues and cannabis in foods are studied by various evaluations. However, we must discuss new food challenges to reduce any possible risks from food directly. Therefore, these food materials and by-products should be strictly monitored and regulated, and an analytical evaluation threshold to ensure scientific food safety and quality is increasingly required. The different kinds of food ingredients remain the problem due to the lack of enriched technologies, and overcoming the challenge of establishing more sensitive, robust, and cost-effective analytical platforms to efficiently extract and analyze food components existing in complex food matrices is essential. Also, efficient analytical platforms to analyze huge amounts of data, food ingredient databases would be developed for foodomics.

Recently, a MS based metabolomics has been performed to elucidate the biological systems that mediate the metabolic variations of a hydrophilic and hydrophobic molecules in the body. This technology can be recognized with analysis of nutrients (e.g., amino acids, lipids, vitamins, and fatty acids), additives (e.g., colorants, preservatives, and fragrances), and food contaminants (e.g., pesticides, drugs, plasticizers, and surfactants). Specially, metabolomics platform has been successfully applied in various fields of food science, including food safety evaluation, quality management, aroma/flavor chemistry, and functional discovery. This review aims to present the recent trend and advance of LC-MS based metabolomics approach and application regarding to “Foodomics.” In addition, this review provides new perspectives and insights into the methodology, sample preparation techniques, and data processing in food analysis by LC-MS.
2. Foodomics

To study biological phenomena holistically, as an integrated interpretation of biomolecular features and variations in human health and diseases, is essential. It is important to analyze various genes and biomolecules, such as proteins and metabolites, to obtain comprehensive data. This concept is multi-“Omic” study/analysis that has advanced and applied with data framework. In 2009, Cifuentes showed that the multi-“Omic,” alias “Foodomics,” used to control food safety, quality, traceability, and hygiene management and to assess functional and nutritional factors in the field of food science (Fig. 1). For food research, a broad overview of omics studies is useful for investigating the actions and interplay of food components in the body and evaluating any unexpected contamination, degradation, expediency, multifunctionality, and undiscovered ability of foods. Foodomics covers various biological materials that constitute the human body, such as biopolymers (including nucleic acids, proteins, and polysaccharides) and their components (including nucleotides, peptides, amino acids, lipids, vitamins, and hormones) using genomics, transcriptomics, proteomics, peptidomics, lipidomics, and metabolomics for profiling, fingerprinting, authenticity, and biomarker classification in foods. These papers regarding to “Foodomics” have been increased based on PubMed research (Fig. 2).

The foodomics study has been widely used the MS technique (48% in total papers for foodomics) including HPLC, ultrahigh performance liquid chromatography (UHPLC), hydrophilic interaction liquid chromatography (HILIC), GC, and supercritical fluid chromatography (SFC). With the introduction of high-resolution MS, the combination of these LC techniques has a great potential to study food and nutrition components, ranging from small molecules to huge proteins with high sensitivity, which is important in foodomics study, such as proteomics and metabolomics. Furthermore, MS enables the simultaneous detection of various target analytes in com-

Fig. 1. Outline of “Foodomics”

Fig. 2. Reported Papers by Year Using PubMed with Keyword “Foodomics”
plex food mixtures with an individual m/z value. MS has been widely recognized as an important tool for proteomics, peptidomics, metabolomics, and lipidomics for food safety, quality, traceability, and functionality. MS-based proteomics technology measures a wide range of proteins, such as food allergens and enzymes in food materials/by-products using bottom-up or top-down approaches, reducing the allergic response and disease risks for susceptible individuals. The general flowchart for foodomics with MS is designed to achieve the final goal of selecting food materials for data processing. The key point of foodomics is the efficient extraction, purification, and analysis of food components existing in complex food matrices. Next section focused on LC-MS-based metabolomics, which aims at a low-molecular-weight compounds (metabolites with molecular weight <1500 Da).

3. Foodomics Based on LC-MS
LC-MS-based metabolomics and lipidomics platforms have been successfully applied in various fields of food sciences. Based on the coverage of these chemicals in food materials, the analytical approach can be mainly divided into untargeted or targeted LC-MS methods with their own advantages and limitations. Caja and Fiehn indicated in terms of quantifications and comparability of results between studies that fully untargeted/targeted approaches should be depended on the aim and scope of the underlying research questions based on analytical chemist’s decision. In addition, LC-MS platform provides several advantages such as column separation, reduced ion-suppression effects, sample preparation, and detection of many chemicals for untargeted/targeted approaches in food materials. Thus, based on LC-MS method, it is introduced to use each untargeted/targeted approach on foodomics.

In the recent past, for untargeted approach in foodomics, the quadrupole time-of-flight (Q-TOF) mass spectrometer, couples a TOF instrument with a quadrupole or high resolving power and mass accuracy Orbitrap instruments have been used. These instruments are framed by such hybrid system as Q-TOF where Q refers to a mass-resolving quadrupole to a radio frequency or hexapole collision cell and TOF or the linear quadrupole ion trap/Fourier transform ion cyclotron resonance (LIT/FT-ICR) that perform in the high resolution (>10000) and high mass accuracy (<5 ppm) regimes. Commonly, the Q-TOF instrument has been originally used for peptide sequencing in bioanlsysis but has worked out application in many areas of untargeted analytical methods in foodomics. However, for full scan-ions in which a single ion is monitored such as precursor ion and neutral loss scans, the sensitivity of the Q-TOF is lower than triple-quadrupole (QqQ) instrument because more ions are lost in TOF compared with a quadrupole. These limitations were overcome with the LIT/FT-ICR, which combines the speed, large trapping capacity, and MS3 capability with the high mass accuracy, resolving ability, sensitivity, and dynamic range. Recently, the LTQ-Orbitrap (LTQ stands for “linear trap quadrupole”) hybrid instrument (Thermo Electron Corporation) was utilized in foodomics. These untargeted metabolomics/lipidomics is a comprehensive analysis of all measurable analytes in a sample, including chemical unknowns to understand the broad insights into human metabolism through the detection of metabolic changes in biological pathways. Owing to the coverage, this approach is ideally suitable for screening analysis to discover an efficient marker when coupled with data analysis. For data analysis, such as metabolic fingerprinting and pattern analysis, a freely available MS spectral database, high-performance software to process large quantities of raw metabolomics data, and chemometric tools to extract information from big data are required to identify unknown compounds.

Targeted metabolomics with LC-QqQ MS/MS is the measurement of defined groups of metabolites. It is an acceptable and reproducible platform technology for detecting selected analytes based on multiple reaction monitoring (MRM) and plays an important role in routine analysis. Although this approach is limited to few metabolites because the number of commercially available standards and isotope-labeled reagents is insufficient, the strength of targeted metabolomics is to verify and validate the biological hypotheses based on its accuracy and robustness with a high quantity. Recently, a widely targeted metabolomics analysis has drawn attention to qualitative and quantitative methods in omics studies. This approach is large-scale targeted metabolomics with high accuracy of the targeted approach and high coverage of untargeted approaches. It explores hundreds of bioactive metabolites and provides in-depth breadth overall perspectives of a human metabolome to elucidate the mechanism. Comprehensive metabolomic data effectively determine the key metabolites and metabolic pathways related to food safety and quality. As an example of food quality change, long-term storage of tea causes quality deterioration, and sensory features are significantly changed. However, the association between sensory features and metabolic alterations remains unclear. Fan et al. investigated the sensory quality-related chemical changes in white peony tea during storage by a widely targeted metabolomic analysis to correlate non-volatile metabolite compositions with sensory traits, including six taste attributes (e.g., smoothness, sweetness, umami, astringency, thickness, and sourness). The decline of abundant flavonoids, tannins, amino acids, and the increase in total contents of phenolic acids and organic acids were observed to be related to the reduced astringency and umami, and increased browning of tea infusions with storage.

Lipidomics is a subsection of metabolomics, targeting lipid metabolism, which has an important role in the homeostatic maintenance of the human body, such as the digestion and absorption of dietary fat. It is known that the lipid profile of foods significantly changed during food processing. For example, roasting to improve the oxidative stability and odor of flaxseed oils before oil extraction can affect the lipid profile of oils. Zhang et al. applied a wide coverage of lipidomic approach to investigate the changes in lipid profile, including fatty acids, oxidized fatty acids, phospholipids, and triacylglycerols of unroasted and roasted flaxseed oil using dry air roasting. The results showed that the composition and content of lipids were significantly different between unroasted and roasted flaxseed oil; therefore, roasting had an obvious effect on lipids, particularly polyunsaturated fatty acids, phospholipids, and oxidized fatty acids. Hence, it is necessary to select a untargeted or targeted approach in terms of quantification and comparability of results, depending on the aim and scope of the underlying research questions.

To interpret the human metabolome, the metabolomics data require preprocessing (standardization, alignment, grouping, and annotation), statistical analysis (multivariable analysis and
machine learning technology), and metabolite identification to identify metabolic dysregulation in biological pathways. Although the development of open-source bioinformatic tools, such as Mass Spectrometry-Data Independent AnaLysis software (MS-DIAL), has significantly simplified the metabolic feature extraction, it remains a time-consuming and labor-intensive process of metabolomics workflow. Further bioinformatic algorithms to automatically combine metabolomic data generated from multiple analytical platforms are required.  

4. Sample Preparation for Foodomics

The amount of sample required for foodomics depends on the chemical concentration in food materials. Therefore, targeted food needs to be selected in accordance with the purpose of research from a categorical group. However, the selection of food samples has a wider range of choices than biological metabolomics. Moreover, there is also the possibility of degrading the detectability of target analytes because of the complexity of food matrices, which considerably hinders the selective extraction and decreases the sensitivity of the LC-MS method. Sample preparation is crucial in foodomics to achieve effective extraction, preconcentration of the targeted analytes, and adequate clean-up of matrix interferences prior to instrumental analysis.  

With the growing demand for foodomics, the sample preparation for pesticides, veterinary medicine, additives, and others has been expanded to develop useful and effective procedures. To eliminate other matrix components such as co-eluted hydrophilic and/or hydrophobic compounds from complex food matrices, various sample preparations, including liquid-liquid extraction (LLE), solid-phase extraction (SPE), the quick, easy, cheap, effective, rugged, and safe (QuEChERS) method, molecularly imprinted polymer (MIP), and derivatization is commonly used for different kinds of foods. For the targeted approach, we can select from various sample preparations based on the characteristics of the targeted analyte, whereas the untargeted approach represents a different challenge because a nonselective extraction is required to maximize the number of metabolites determined. Therefore, for the untargeted approach, an extraction method based on polar or nonpolar characteristics as one choice is applied. In any case, selecting proper sample preparation techniques that match the characteristics of analytes is important for food analysis.  

LLE is one of the most conventional techniques for polar or moderately polar analytes and is commonly used in many analytical applications to separate compounds or metal complexes based on their relative solubilities in two different immiscible solvents, such as water (polar) and an organic solvent (non-polar). To minimize the organic solvent used, the LPME method for both inorganic and organic analytes in different food matrices was implemented. This approach uses only microliters of solvents instead of several hundred milliliters in LLE. Recently, deep eutectic solvents (DESs) have been widely used for dispersive liquid-liquid microextraction (DLLME) of organic and inorganic analytes from aqueous and organic samples. Using DESs, a new approach for the LLE of analytes based on in situ formation of deep eutectic mixtures has been developed. During the extraction of targeted analytes in food matrices, selectivity can be easily achieved by varying the ratios of hydrogen bond donors and hydrogen bond acceptors (choline chloride, betaine, etc.) in their structures. Therefore, DES-based DLLME has gained considerable popularity among food researchers for pre-concentration in food matrices.  

SPE is the most frequently used and is applied in various fields, including food, environmental, pharmaceutical, and clinical research. It separates targeted analytes and interfering compounds from liquid samples using different solid adsorbents for separation and purification. Therefore, the choice of an optimal adsorbent based on the characteristics of the targeted analyte is significant to perform an appropriate SPE method because the adsorbents can affect the selectivity, affinity, and capacity. Recently, various SPE cartridges with separation modes such as reversed-phase and ion-exchange have been commercially available. Reversed-phase SPE cartridges such as C8 and C18, which involve a polar or moderately polar sample matrix and a nonpolar stationary phase, are recommended for extracting hydrophilic to lipophilic compounds as the first choice. Moreover, a cyclohexyl sorbent is used to extract moderately non-polar or neutral compounds and has a different selectivity than the C8 and C18 phases. In addition, a phenyl SPE cartridge extracts non-polar to moderately polar aromatic compounds by π–π interactions due to the electron density of the phenyl ring. However, SPE sorbents for the analysis of complicated food samples are still insufficient. Recent experiments showed that magnetic nanoparticles based SPE method was applied for sample preparation. Although they show high selectivity for targeted analytes, there are several problems regarding their high cost, laborious operation, low stability, and poor reusability.  

MIP has been used as a versatile and promising SPE sorbent material that can improve the efficiency and selectivity of SPE. MIP is a synthetic polymer of tailor-made artificial antibodies and receptors with specific molecular recognition sites for targeted analytes and their analogs. They are currently attracting attention owing to their high selectivity, high stability, low cost, and ease of preparation because they are synthetic polymers. In addition, magnetic MIP for selective recognition and extraction of chemicals was developed.  

However, the SPE procedure mentioned above is quite difficult and time-consuming to pretreat many samples. Therefore, a pipette-tip SPE method using a pipette tip as the SPE cartridge is expected to perform the SPE procedure simply and easily. Owing to the small amount of adsorbent, this procedure is very convenient, high-speed, and requires less sample volume, which significantly saves labor, time, and cost. Owing to the recent development of different types of micro and nano sorbents, including silica materials, carbon materials, organic polymers, MIP, and metal-organic frameworks for pipette-tip SPE, the scope of applications of a pipette-tip SPE method has expanded. Hroboňová and Brokešová reported the effectiveness of a pipette-tip SPE packed with coumarin-specific MIP, which showed greater selectivity and extraction efficiency of coumarins in wines compared with conventional SPE sorbents. Sun et al. reported that the superiority of pipette-tip SPE can give full play to complex food samples.  

Moreover, food studies employing the QuEChERS approach for the comprehensive isolation of analytes from different sample matrices have rapidly gained popularity in recent years. The QuEChERS was developed to recover pesticide residues from fruits and vegetables and has become a choice for food analysis through various modifications and enhancements over
the years. Generally, the QuEChERS procedure is based on an initial single-phase extraction in a tube with acetonitrile for the pretreatment of food samples using a dispersive SPE technique. A liquid–liquid partition is carried out by salting-out by adding anhydrous sodium chloride and magnesium sulfate, which removes the majority of the remaining matrix interfering. After centrifugation, the acetonitrile layer containing the analytes was collected. Although there are still some limitations of the QuEChERS method, the selectivity and clean-up are less efficient than SPE, and the purification is sometimes insufficient. The advantages of the QuEChERS method are its quick and simple procedure, easy handling, and low-error rate. In addition, an extensive range of QuEChERS available is also attractive for a wide selection of particles for the applicable analytical methods based on LC-MS-based foodomics.

Table 1 summarizes the recent and useful sample preparations based on LC-MS foodomics/exhaustive food analysis.

| Sample preparation | Analytes | Foods | LC-MS | Ref. |
|--------------------|----------|-------|-------|------|
| QuEChERS           | 500 Pesticide Residues | Hen Eggs | Waters Acquity UPLC I-Class/ Xevo TQD | Golge et al., 2021(77) |
| QuEChERS           | 302 Pesticides and contaminants | Catfish muscle | Shimadzu LC/Sciex QTRAP 6500 | Han and Sapozhnikova, 2020(80) |
| Dispersive micro-LLE | 76 Short/medium-chain chlorinated paraffins | Wines | Agilent 1290/6540 | Zhou et al., 2020(86) |
| Dispersive micro-SPE | 14 Endocrine-disrupting chemicals | Bottled mineral waters | Shimadzu LCMS-2020 | Vivas et al., 2021(80) |
| MIP-SPE            | 4 Phthalates | Tea | Agilent 1100/Sciex API 150EX | Alnaimat et al., 2020(80) |
| Freezing and two-step dSPE | 267 Pesticides and their metabolites | Beebred | Agilent 1260/Sciex QTRAP 6500 | Kiljanek et al., 2021(86) |
| Pipette-tip micro SPE | 12 Pesticides | Fruiting vegetables | Agilent 1200/6410 | Keikavousi Behbahani et al., 2021(86) |
| QuEChERS-DLLME    | 7 Neonicotinoid | Grains | Agilent 1322/6460 | Ma et al., 2020(84) |
| QuEChERS-enhanced matrix removal | 55 Phenolic compounds | Fatty matrices | Agilent 1290/6470 | Yin et al., 2021(86) |
| QuEChERS-magnetic SPE | 11 Estrogens | Pork and chicken muscles | Shimadzu LC/Sciex QTRAP 4000 | Xiong et al., 2020(86) |

5. Data Processing in Foodomics

The raw foodomics data is subjected to a data processing to build comfortable models for classification, pattern recognition, optimization, and prediction. Then, we must select the most appropriate approach to correct qualitative and/or quantitative data. However, data processing is likely to be a bottleneck because of the laborious and time-consuming procedures for foodomics. As an example of untargeted screening analysis, it is usually difficult to obtain retention time and MS/MS spectra information for the determination of unknown peaks if the open-source databases and helper libraries are currently unavailable. Therefore, identifying all unknown peaks detected from food extracts is likely impossible due in part to time limitations and insufficient MS information. The data processing has been carried out in three steps; (i) data preprocessing, (ii) univariate and multivariate analysis for the model building, and (iii) data assessment of the models. Usually, the raw data must be preprocessed before subjected to analytical software.

Data preprocessing is important to perform an appropriate statistical method, and however needs to pay attention because it is liable to change the essential results on metabolomics. It is used to adjust for differences in the offset between high and low abundant molecules as outliers. A scaling procedure for normalization divides each variable by a factor which is different for each variable, considering the degree of contribution. This often follows to eliminate systematic errors and ensure that different signals have a comparable influence on the result.

After the standardized data have been prepared, they are evaluated using statistical methods, including multivariate analysis, machine learning, and deep learning with artificial intelligence (AI) technology. A multivariable analysis is a commonsense way to extract relevant qualitative or quantitative information from complex data and to analyze group assignment based on similarities and/or differences in food analysis. Overall multivariate strategies are divided into nonsupervised and supervised approaches. Nonsupervised methods like principal component analysis (PCA) are often used to exploratory for pattern recognition, while supervised method like orthogonal partial least-squares-discriminant analysis (OPLS-DA) is used to class discrimination. PCA is performed to reduce the multidimensional data to a two- or three-dimensional data constructing new principal components (PCs) and explores the overall variations of dataset to visualize the total metabolic differences among the different samples. Later, OPLS-DA is applied to develop the classification model by incorporating class information into the model and finds directions to maximize class separability while aiming to minimize the dispersion between each class. Meanwhile, S-plot and variable importance in projection (VIP) help to extract and visualize the OPLS-DA predictive variables which contribute to class discrimination. However, the supervised procedure often tends to be overfitting, and thus it is recommended to confirm the model reliability by further cross-validations and other model performance criteria. Recently, a state-of-the-art computer science operating with AI technology have become a very widespread area known as machine learning and deep learning, and it has been introduced in food science and engineering. The most prominent feature is to learn the potential characteristics of big data and to improve the model performance by training. Frequently used supervised algorithms are random forest and support vector machines. The supervised learning is defined
that it uses labeled datasets to train algorithms for classification and prediction of outcomes accurately. For the food quality prediction, Kaya et al. applied the machine learning to the failure tolerance of the failed sensors to detect smell or flavor caused by physical aging and environmental factors. This approach showed a promising clear advantage compared to the conventional approaches and enhanced the assessment of food quality. The combination of deep learning and multi-source data fusion would enable a more comprehensive assessment of food. Therefore, deep learning is a promising tool in food quality and safety, and a classification/regression in deep learning is applicable to the field of food science in the future.

6. Conclusion

To evaluate the unexpected contamination, degradation, multifunctionality, and undiscovered ability of food materials and by-products, the foodomics approach has attracted more attention for ensuring food safety, food authenticity and quality, and food traceability, providing reliable data in many areas of nutrition, food science, and health. To meet the current requirements of food research, the development of new methodologies and improved sample preparation for various food matrices is a broad topic and continues to be challenging. Recent advances in high-throughput qualitative and quantitative analysis by LC-MS/MS significantly impacted the metabolomics and/or lipidomics workflow based on foodomics. LC-MS-based metabolomics and/or lipidomics techniques have been extensively explored to characterize food ingredients, allowing various analytical platforms for untargeted and targeted approaches. Future outlooks in metabolomics and lipidomics can be expected (1) to merge the targeted and untargeted approaches for an efficient metabolomics strategy, (2) to optimize suitable sample extraction of analytes from complex food matrices, (3) streamline sample preparation including automatization and/or semi-automatization, (4) to develop bioinformatic algorithms for data processing to automatically combine the huge amount of metabolic data, and (5) to validate and enrich databases on food constituents and libraries of MS/MS spectra for metabolite identification. In this review, LC-MS based metabolomics and lipidomics methodologies were demonstrated, and recent advances and applications in sample preparation techniques, such as LLE, SPE, and QuEChERS, were reviewed. Selecting the proper sample preparation techniques that match the characteristics of targeted analytes is important for food analysis, requiring a time-saving, high-throughput, and cost-effective method. Therefore, LC-MS-based metabolomics procedures should be further developed to ensure analytical data quality management and validate a standardized protocol for accurate, precise, and reliable quantification in food analysis.

Conflict of Interest The authors declare no conflict of interest.
