UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN

"Scienze degli alimenti"

CICLO XXXII

INNOVATIVE ANALYTICAL METHODS MERGED WITH CHEMOMETRICS FOR FOOD INTEGRITY AND SENSOMICS

Coordinatore:

Chiar.mo Prof. Chiara Dall'Asta

Tutori:

Chiar.mo Prof. Chiara Dall'Asta

Dott. Michele Suman

Dottorando:

Daniele Cavanna

Anni 2016/2019

Table of contents

Preface

Fraud in the food supply chain is becoming increasingly common due to the huge profits associated with this type of criminal activity; moreover, recently, an increasing interest of consumers for safety, authenticity and quality of food components has been detected, following highly mediatized crises such as the melamine scandal in China in 2007 and Horsegate in UK in 2013.

For these reasons, the general attention on the assurance of products quality and authenticity along the entire supply chain is dramatically increased in the last decade, pushing the scientific community to strengthen their efforts and develop new methods able to guarantee the food integrity.

The first point that should be clarified is the meaning of some common words used in the food frauds field.

Food integrity, according to a recent document published jointly by FAO and WHO, is "*the status of a food product where it is authentic and not altered or modified with respect to expected characteristics including, safety, quality, and nutrition*" (Codex Alimentarius, 2018), while in the same document it is declared that a *food fraud* is "*any deliberate action of businesses or individuals to deceive others in regards to the integrity of food to gain undue advantage.*" (Codex Alimentarius, 2018).

The most diffused fraudulent activity is the adulteration of the products; according to a recently published CEN document, the *adulteration* is "*a type of food fraud which includes the intentional addition of a foreign or inferior substance or element; especially to prepare for sale by replacing more valuable with less valuable or inert ingredient*" (CWA 17369:2019).

In the same document it is declared that possible adulterations are:

- Dilution: the *"process of mixing a liquid ingredient (solute) with high value with a liquid of lower value"* (CWA 17369:2019).
- Substitution: the *"process of replacing a nutrient, an ingredient or part of a food (often one with high value), with another nutrient, ingredient or part of food (often one with lower value)"* (CWA 17369:2019).
- Concealment: the "*process of hiding the low quality of food ingredients or products*" (CWA 17369:2019).
- Unapproved enhancement: the *"process of adding unknown and undeclared compounds to food products in order to enhance their quality attributes"* (CWA 17369:2019).

Together with the adulteration problems, the other main issue that affects all the food chains is the fraudulent declaration of raw materials geographical origins.

All these practices are Economically Motivated Adulterations, that are *"the intentional substitution or addition of a substance in a product for the purpose of increasing the apparent value of the product or reducing the cost of its production, for economic gain"* (Codex Alimentarius, 2018).

Industries are usually at the end of the production chain and are dramatically interested in the protection of the quality and authenticity of the purchased raw materials.

For this reason, from an industrial point of view, the development of analytical methods that are able to solve these topics should lead to a better brand protection and to a more safe and a more valuable finished product.

In recent years, rapid screening methods and confirmatory techniques were used for food frauds detection; what is clear is that methodologies focused only on a specific list of compounds are not enough because adulterants or contaminants not included on the target list would not be identified.

For this reason, the development of analytical methods able to work also in the "non-targeted" field is a crucial requirement for a better food fraud detection and the high resolution mass spectrometry is the preferred technique in this field: following a metabolomic approach it is possible to identify some compounds that can be considered specific markers of the target fraud.

However, at the moment, common and validated protocols and methods for the detection of unknown compounds are not available and rugged databases do not exist.

Furthermore, also rapid methods are important for food companies, taking into account that mass spectrometric or similar approaches require days of analysis but sometimes the fraudulent material should be stopped before its introduction in the production plant.

This PhD work was executed within this scenario: after a proposal of protocols standardization for non targeted analyses, the feasibility of these approaches for food frauds detection was explored in different raw materials, using both rapid and confirmatory techniques. An attention was also devoted on the industrial implications of the results obtained.

This research was developed in strictly cooperation with a food company (Barilla G. & R. Fratelli S.p.A.) and this is the reason why, for each section, also the industrial impacts of the results will be studied.

In addition, in a parallel direction devoted not to food authenticity/safety issues and instead oriented to maximize added value of food products in front of consumers preferences, in the last part of this project a similar analytical approach has been conceived within the framework of Sensomics Science, trying to identify active sensory molecules responsible of consumers' liking.

Bibliographic references

Codex Alimentarius (2018). – Discussion paper on food integrity and food authenticity - Joint FAO/WHO food standards programme. Codex committee on food import and export inspection and certification systems. Twenty-Fourth Session. Brisbane, Australia, 22 - 26October 2018. CX/FICS 18/24/7. Available at: http://www.fao.org/faowho-codexalimentarius/shproxy/en/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252FMeetings%252FCX-733-24%252FWorking%2BDocuments%252Ffc24_07e.pdf.

CWA 17369:2019 - Authenticity and fraud in the feed and food chain - Concepts, terms, and definitions –published 2019-01-23

Introduction

Frauds in food commodities are largely diffused in the world and nowadays are performed at different levels of the supply chain. These illegal actions affect not only the quality of the products but sometimes represent also a problem for consumers' safety. For this reason, the demand of both rapid and confirmatory analytical methods able to detect frauds increased in these years and several scientific publications are focused on this topic. However, at the moment, it is not really known what are the common approaches used in the industrial production facilities to fight these issues.

This first part of this introductive chapter presents the analytical tools suggested in the scientific literature to detect food frauds for eleven different commodities, carrying out simultaneously a comparison between researchers activities and what is actually done by the food companies.

A selection of the more relevant scientific methods published up to now is presented and compared with the information about the industrial behavior, obtained through several interviews with experts coming from different industries and food chains.

What emerged from this comparison is that the academic and the food industry worlds are not aligned so far. In agreement with literature data, several reliable analytical methods are currently available but they are not applied (or even not known) into industrial quality control laboratories. In some circumstances "academic suggestions" are not applicable in the routine analyses due to a combined mix of necessary costs and skills. These two entities should interact better, in order to develop robust techniques that could merge the scientific novelty with the reliable application in the food production plants.

The second part of this chapter is focused on the recent developments in non-targeted analyses.

There is an increasing availability of solutions and applications based on high resolution mass spectrometry (HRMS), allowing parallel non-targeted approaches, novel compounds identification and retrospective data analysis. For these types of methods sample-handling must be minimal to allow the inclusion of as many as possible chemical categories. However data-handling of such methods is much more demanding, together with the potential requirement to integrate multiplatform data as well as conducting data fusion. To allow the processing of massive amounts of information based on the separation techniques and mass spectrometry approaches employed, effective software tools capable of rapid data mining procedures must be employed and metabolomics based approaches does appear to be the correct way forward.

To verify the relevance of modelling results, appropriate model validation is essential for non-targeted approaches, confirming the significance of the chemical markers identified.

Therefore, the second part of this introductive chapter is devoted to review and assess the current state of the art with regards non-targeted mass spectrometry in food fraud detection within many food matrices and to propose a harmonized workflow for all such applications.

This review was accepted for publication in "*Trends in Food Science and Technology"*. For additional details see section "Author".

Strategies for fighting food fraud: scenario comparison of solutions reported in scientific literature and real approaches implemented in industrial facilities

Michele Suman[1], Francesca Lambertini[1], Daniele Cavanna[1,2], Cecilia Loffi[1,2]

[1]Barilla G. R. F.lli SpA, Advanced Laboratory Research, via Mantova 166, 43122, Parma, Italy
[2]Department of Food and Drug, University of Parma, Parco Area delle Scienze 95/A, 43124, Parma, Italy

Introduction

Recently, consumers have increased attention on the safety, authenticity and quality of food commodities (Borras, et al., 2015), and the development of reliable methods to detect fraud could better preserve the quality of production, with an advantage for the industry and the final consumer, who would have more confidence in the quality and safety of the product.

Current governance of food supply chains, through certifications, inspections and audit controls, has historically been developed with a focus on food safety.

Criminal actions not only involve the adulteration of raw materials or finished products, but also the fraudulent declaration of the geographical origin of commodities (Bontempo, Camin, Paolini, Micheloni, & Laursen, 2016).

Understanding the root causes of food integrity issues, behaviours that drive certain decisions and activities, and the data that could help in the identification and prevention of fraud is therefore very important. There is a close link between criminal activities, geographic regions and specific agrifood resources and businesses: in this sense, criminals become real entrepreneurs, looking mainly to increase profits on black market. Often, the risks are much higher when small-medium enterprises could be involved in some way; local complicity by a "corrupted political environment" that acts in order to facilitate illegal actions also cannot be excluded (Ponzi, 2017, Suppl. Mat.).

Speculation and unjustifiable low prices and promotional initiatives indirectly favour the work of illegal fraudulent networks (Elliott, 2019, Suppl. Mat.)

Both confirmatory techniques and rapid methods are used to detect food fraud, and sometimes the union of these two approaches increases the number of illicit samples identified (Black, Haughey, Chevalier, Galvin-King, & Elliott, 2016).

For this reason, one of the areas that appeared to stimulate interactions between researchers and industrial experts is the challenge of providing faster and more affordable methodologies for the identification of food fraud in some specific raw materials and commodities.

However, sometimes it seems that the new analytical methods developed and presented in the peer-reviewed journals are not applied or even not known by the quality control laboratory industry.

In specific circumstances, their lack of application can be justified by limited applicability only to specific fields: for instance, the use of molecular markers limits applicability (e.g. geographical origin, organic origin), and it requires extensive up to date databases (Cubero-Leon, De Rudder, & Maquet, 2018, Suppl. Mat.).

The present study stemmed from the activities included in the framework of the FoodIntegrity European Funded EU-FP7 Project (Ensuring the Integrity of the European food chain).

Specifically, FoodIntegrity was a 5-year/12 million Euro project, recently completed (2014-2018), with associated partners including regulators, consumers, academia and food industry. The main aims and key activities of the project were: (i) to provide Europe with a state of the art and integrated capability for detecting fraud and assuring the integrity of the food chain, (ii) to provide a sustainable body of expertise, (iii) to share data and knowledge, (iv) to develop methods and systems for industry, (v) to develop new methods of analysis, (vi) to develop early warning systems, (vii) to understand consumer behaviour for export.

The FoodIntegrity consortium focused dedicated work-packages to the three most important (mainly in terms of combined volumes and added values) food chains represented by extra virgin olive oil, fish and spirit drinks (not considered in the this paper, because this would have otherwise required an entire work exclusively dedicated to each topic).

Then, the consortium devoted specific attention to select, identify and evaluate the industrial impact of all the other eleven relevant food chains that can be affected by food fraud issues: coffee, cocoa & chocolate, spices, eggs & egg products, meat, honey, cereals & flours, tomatoes and derivatives, fruits & vinegar, cheese-milk derivatives, mushrooms.

The aim of this paper is to provide relevant information about the existence in scientific literature of applications of both rapid and confirmatory analytical strategies for fighting fraud and adulterations along different food chains; this information is then combined with feedback directly collected from many industry representatives about their internal approaches to prevent food fraud.

Therefore, what could be done and what actually is done represents the final goal of this work: a comparison between the analytical approaches suggested in scientific papers to detect food fraud and the current scenario within food companies is presented through an analysis of these eleven different types of raw materials / commodities.

Approach to data collection

In order to highlight new analytical strategies to detect food fraud, research on scientific literature was carried out with a focus on potential applicability and technology transfer, taking into account the most relevant papers published in the last 20 years, with some exceptions due to specific suggestions of the FoodIntegrity consortium. Mainly Elsevier-ScienceDirect and CAS-SciFinder databases were searched: a preliminary literature screening was executed combining the keywords "food fraud", "food adulteration", "industrial production" and "quality control" for each commodity. The results from extracted literature were then further refined using the keyword "rapid method", coming to a final number of approximately 1900 records. In parallel, another refinement was made using the keyword "confirmatory method". In this case the final number of extracted records reached 100.

Finally, a subsequent selection of the most relevant papers was evaluated within the Food Integrity Consortium, following the suggestions of different experts, mainly coming from the industrial integration work-package.

The critical literature revision was combined with a rational discussion (single interviews were conducted, face to face or by dedicated conference calls) with 22 industrial experts from the different selected food chains: senior scientists and managers that cover correlated roles within Quality Assurance, Vendor Assurance, Purchasing, R&D departments. In addition, information collected from literature and interviews was compared with outcomes derived from official panel discussions at two main events in 2015: Food Integrity EXPO UK Science Innovation Network Workshop (Milan, Lodi, September 2015) & Food Integrity Recent Advances in Food Analysis Open Days 2015 (Prague, November 2015) [more details, information and documentation about these events can be found on the FoodIntegrity website www.foodintegrity.eu)].

The involved industrial experts were encouraged to consider vulnerabilities to specific supply chains, with a view to protecting their businesses by using information and intelligence, also providing information about their risk assessments and assurance actions.

Commodity "Journey"

Each food chain is analysed separately as follows. There is an introduction with the list of the most diffused frauds for each commodity, with examples of rapid and confirmatory approaches presented in literature to manage these problems and, lastly, the industrial approach to the same topics. A comprehensive overview is reported in the summary tables, with indication for each method of whether it is for rapid (R) or confirmatory (C) purposes.

1. Coffee *(Table 1)*

1.1 Most diffused fraud

Due to its pleasant taste and flavour, coffee is one of the most common and valuable beverages. It is largely subjected to adulterations because of the high economic benefits that fraudster can gain through the addition of low-quality substitutes.

The most diffused coffee fraud concerns un declared blending of Arabica beans with cheaper Robusta beans, the fraudulent addition of husks, stems, wheat, barley and other components, and the false declaration of the geographical origin of beans (Morin , Jamin, Guyader, & Thomas, 2018, Suppl. Mat.).

1.2 Rapid approaches presented in literature

According to information reported in literature, most of the substitutes and additions can be commonly detected with Infrared and UV- spectroscopies: Near Infrared and UV visible spectroscopies, for example, can be used for the detection of added barley (Ebrahimi-Najafabadi, et al., 2012) and/or to discriminate between Arabica and Robusta coffee beans (Pizarro, Esteban-Diez, & Gonzalez-Saiz, 2007,Suppl.Mat.) (Dankowska, Domagała, & Kowalewski, 2017). UV, together with the DRIFT (Diffuse Reflectance Infrared Fourier Transform Spectroscopy) approach, can also be used to detect foreign materials (Souto, et al., 2015) (Reis, Franca, & Oliveira, 2013, Suppl. Mat.) (i.e. husks, stems, etc.), as well as DSC (Differential Scanning Calorimetry) and FTIR techniques (Brondi, Torres, Garcia, Jerusa, & Trevisan, 2017).

In all these rapid methods, chemometric models as PCA (Principal Component Analyisis) or PLS (Partial Least Squares) are performed, after data collection, to clearly detect the adulterated samples.

1.3 Confirmatory approaches presented in literature

Moving to the confirmatory approaches, HPAEC-PAD (High Performance Anion-Exchange chromatography with Pulsed Amperometric Detection) analysis of carbohydrates is commonly used for the detection of fraudulent addition of soybeans, wheat or other foreign substances (Domingues, et al., 2014) (Pauli, et al., 2014, Suppl. Mat.).

The same goal can also be reached using UHPLC-HRMS (Ultra high-Performance Liquid Chromatography – High Resolution Mass Spectrometry) techniques followed by chemometric evaluations (Cai, Ting, & Jin-Ian, 2016).

On the other hand, thanks to SPME (Solid Phase Micro Extraction), coffee purity can also be evaluated through the extracted aroma-related compounds from coffee samples, using GC-MS (Gas Chromatography – Mass Spectrometry) or GC x GC (Toledo, Hantao , Ho, Augusto, & Anderson,

2014, Suppl. Mat.) (Mondello, et al., 2004, Suppl.Mat.); the same analytical strategy is also sometimes used to detect fraudulent barley addition (Oliveira , Oliveira, Franca, & Augusti, 2009, Suppl.Mat.).

Arabica and Robusta beans can be discriminated thanks to DNA extractions and PCR amplification (Spaniolas, Tsachaki, Bennett, & Tucker, 2008, Suppl.Mat.) or with ^{1}H-NMR analysis (Medina, et al., 2017).

The geographical origin of coffee beans is evaluated using NMR spectroscopy, which after data processing can discriminate different countries (Consonni, Cagliani, & Cogliati, NMR based geographical characterization of roasted coffee, 2012, Suppl. Mat.) (Arana, et al., 2015).

1.4 Industrial approach and needs

Industrial experts declared that NIR and FTIR techniques are used as rapid methods only to discriminate between Arabica and Robusta beans, and in parallel, they indicated that they also use sensory analyses performed by a trained experts panel for this purpose. The detection of foreign components is executed evaluating the Carbohydrate profile, using HPAEC-PAD techniques; on the contrary, NMR spectroscopy approaches to discriminate the geographical origin of coffee beans is implemented only occasionally in the industry.

In the evaluation of other types of fraud not yet widely studied in literature, the interviewed experts highlighted that they are interested in analytical methods devoted not only to coffee beans but also to coffee beverages in their final forms present on the market (coffee, coffee based cold/hot drinks, etc.); moreover, they declared that potential fraud could occur relatively frequently at the coffee shop/bar level, in the form of a deliberate change/mix/blending of a declared 100% Arabica variety with a Robusta variety. Portable instruments should be developed to detect this type of fraud.

2. Cocoa & Chocolate *(Table 2)*

2.1 Most diffused fraud

A typical fraud in the cocoa food chain is related to false declarations about product composition: its supply chain is very fragmented, starting for instance from a very elevated number of small farmers in poor regions of the world; some adulterations may occur simply by increasing the apparent weight of the cocoa raw material deliveries thanks to the exploitation of cocoa shells (Rektorisova & Tomaniova, 2018, Suppl.Mat.). However, as suggested by the experts, these types of fraud may also employ other “exotic” substances, like stones, powder, small iron balls, wood pieces, etc.

Another crucial issue for this commodity is the false declaration of the geographical origin of the cocoa and/or the Theobroma varieties used.

Recently, the use of vegetable fats other than cocoa butter (CB) in chocolate has been subjected to monitoring: a recent European law allows the use of these vegetable fats up to a maximum of 5% in the final product.

In addition, adulteration of cocoa powder with carob flour is another recent type of fraud reported by industry experts, together with the undeclared presence of allergens in the final product.

2.2 Rapid approaches presented in literature

As concerns the false declaration of product composition, even if not all the adulterations can be immediately identified with the naked eye, several analytical techniques allow the detection of extraneous materials: the addition of unfermented cocoa beans, for example, is easily detected by FT-NIR spectroscopy followed by chemometric data evaluation (Ernest, Xing-yi, Wu, & Huang, 2014); this technique (with specific adjustments) also allows the detection of exogenous compounds in cocoa or chocolate formulations (Che Man, Syahariza, Mirghani, Jinap, & Bakar, 2005, Suppl.Mat.) (Liu, Zhao, Li, Chen, & Zhao, 2015, Suppl. Mat.).

The same techniques allow evaluation of the geographical origin of cocoa and discrimination of different varieties (Teye, Huang, Takrama, & Haiyang, 2014a, Suppl.Mat.) (Teye, Huang, Dai, & Chen, 2013, Suppl.Mat.) (Teye, Huang, Han, & Botchway, 2014b, Suppl. Mat.).

Moreover, with the aim to reach the same purposes, the evaluation of Pb and Cd total amounts and their bioavailability should be highlighted as another rapid discriminating technique (Mounicouy, Szpunary, Andreyz, Blakez, & Lobinsky, 2003, Suppl.Mat.).

With regard to Cocoa Butter Equivalents (CBE) in addition to Cocoa Butter (CB), FTIR or UV-Vis techniques followed by chemometric analysis can detect their addition up to 5% (Goodacre & Anklam, 2001, Suppl. Mat.) (Dankowska, Data fusion of fluorescence and UV spectroscopies improves the detection of cocoa butter adulteration, 2017), while NIR spectroscopy, combined with PCA and PLS-DA, allows a qualitative and quantitative detection of cocoa powder with carob flour (Quelal-Vasconez, Perez-Esteve, Arnauld-Bonachera, Barat, & Taulens, 2018).

2.3 Confirmatory approaches presented in literature

The polysaccharides fingerprint recorded with HPLC, with a subsequent PCA evaluation, can be used for the detection of exogenous compounds in cocoa or chocolate formulations (Yang, et al., 2015).

Other techniques that are applied for the same purposes are GC-MS or MALDI-TOF-MS (Matrix Assisted Laser Desorption Ionization – Time of Flight – Mass Spectrometry) (Caligiani, Cirlini, Palla, Ravaglia, & Arlorio, 2007, Suppl. Mat.) (Bonatto & Silva, 2015).

The largely diffused methods for the assessment of geographical origins are NMR (Caligiani, Palla, Acquotti, Marseglia, & Palla, 2014, Suppl.Mat.) and LC-TOF-MS (Hori, Kiriyama, & Tsumura, 2016), together with isotopic ratio evaluation (Diomande, et al., 2015).

Furthermore, discrimination of different Theobroma varieties is presented in literature from a biological point of view with a Polymerase Chain Reaction approach (Haase & Fischer, 2007, Suppl.Mat.).

With regard to Cocoa Butter Equivalents (CBE) in addition to Cocoa Butter (CB), the study of fatty acids compositions (Spangenberg & Dionisi, 2001, Suppl. Mat.) and the evaluation of Triacylglycerol profiles, obtained with chromatographic techniques (Dionisi, et al., 2004, Suppl. Mat.) (Buchgraber, Androni, & Anklam, 2007, Suppl. Mat.) (Buchgraber, Senaldi, Ulberth, & Anklam, 2004, Suppl. Mat.), are the most common approaches.

Ultimately, the emerging issue of the undeclared presence of allergens in the final product can be tested with a PCR method (Lopez-Calleja, et al., 2013).

2.4 Industrial approach and needs

Both small and large enterprises do not have rapid screening approaches for the assesment of products composition in place, and they base their preventive strategies completely on outsourcing the quality evaluation of the product to specialized external laboratories.

On the contrary, the interviewees confirmed that CBE addition to CB is evaluated internally with the study of fatty acids composition using chromatographic techniques.

Furthermore, for the assessment of the geographical origin of raw materials, only isotopic ratio studies are applied internally, while the evaluation of other risks, like the undeclared presence of allergens, is totally outsourced to third party companies.

3. Spices *(Table 3)*

3.1 Most diffused fraud

Spices are raw materials that have been subjected to many types of fraudulent activities over centuries and across different geographical regions, mainly affecting the business of EU industries through incoming risks from Asia and North Africa.

The complexity and heterogeneity of the production and supply chain determines physical points of medium/high risk at the local harvesting and storage centres: literally, farmers bring bags of spices to stores daily and lots cannot usually be directly correlated to each field of cultivation; this generates potential "mixing risks".

The most diffused adulterations are indeed addition of illegal dyes, blending of different spices and false declaration of geographical origin.

3.2 Rapid approaches presented in literature

As concerns the addition of illegal dyes, the presence of Sudan dyes in Paprika, Turmeric and Curry can be rapidly detected using synchronous fluorescence spectroscopy with multivariate classification (Di Anibal, Rodriguez, & Albertengo, 2015, Suppl. Mat.), with UV-Visible techniques (Di Anibal, Odena, Ruisánchez, & Callao, 2009) or with FT-MIR spectroscopy in combination with SIMCA analysis as a statistical tool (Horn, Esslinger, Pfister, Fauhl-Hassek, & Riedl, 2018).

The same fraud can be easily detected in Chili with NIRS (Near Infrared Spectroscopy) or Raman spectroscopy coupled with PCA or PLS-DA (Haughey, Galvin-King, Ho, Bell, & Elliott, 2015).

Furthermore, in Saffron the addition of dye can be detected with AOS (Artificial Olfactive Systems), usually known by the name Electronic Noses (Heidarbeigi, et al., 2015, Suppl. Mat.).

In order to directly manage the *in-situ* mixing of different spices, portable Electronic Noses and Ion Mobility Spectrometers were developed (Banach, Tiebe, & Hubert, 2012), while the differentiation between "True" Cinnamon (Cinnamon Verum) and other commonly marketed species can be rapidly performed with a DART-HRMS (Direct Analysis Real Time–High-Resolution Mass Spectrometry) approach, which allows the detection of specific markers like Coumarin (Avula, Smillie, Wang, Zweigenbaum, & Khan, 2015).

Adulteration of saffron with other less expensive plants can be detected by DRIFTS (Diffuse Reflectance Mid-Infrared Fourier Transform Spectroscopy), in the mid-infrared combined with chemometric studies (Petrakis & Polissiou, Assessing saffron (Crocus sativus L.) adulteration with plant-derived adulterants by diffuse reflectance infrared Fourier transform spectroscopy coupled with chemometrics, 2017).

3.3 Confirmatory approaches presented in literature

Liquid Chromatography coupled with Mass Spectrometry (LC-MS) analysis is the typical approach used to detect addition of illegal dyes (Li, et al., 2013, Suppl. Mat.), and NMR spectroscopy coupled with chemometric data processing is also applied in literature (Petrakis, Cagliani, Polissiou, & Consonni, 2015, Suppl. Mat.).

With regard to the fraudulent mixing of different spices, marker molecules for Saffron adulterations can be detected with HPLC-PDA/MS (High Performance Liquid Chromatography – Photodiode Array Detection / Mass Spectrometry) methods (Sabatino, et al., 2011, Suppl. Mat.), while NMR technique seems to be another interesting tool to detect this type of fraud: it allows, for example, discriminating between Protected Designation of Origin (PDO) spices and non-PDO spices (Cagliani, Culeddu, Chessa, & Consonni, 2015).

For the same purpose, a completely different method that can be used is the study of DNA barcoding, to detect fraudulent spices in PDO species (e.g. chilli addition to black pepper powder) (Parvathya, et al., 2014).

Ultimately, the assessment of the geographical origin can be performed using an ICP-MS approach, which following data processing can, for example, identify a geographical classification for Saffron (D'Archivio, Giannitto, Incani, & Nisi, 2014, Suppl. Mat.).

3.4 Industrial approach and needs

Taking into account the complexity and heterogeneity of the production chain, industry operates according to the "counter samples philosophy", which means that for each lot delivery a counter sample must first be sent from the storage centres: if the analysis proves its conformity, then the lot is purchased, and when it reaches its final destination it is checked again (to match its analytical parameters and values with those of the original counter sample previously analysed).

As concerns typical fraud, most of the methods presented in literature are well-known; nevertheless, analyses are frequently outsourced to external specialized laboratories.

However, the interviewed industrial managers highlighted that many risks stem not only from adulterations, but also from microbiological (e.g. salmonella due to contaminated water used in rinsing phases) or chemical aspects (e.g. high mycotoxins content or adoption of prohibited pesticide formulations), and for this reason they are expressed interest in methods that could solve these issues.

Another current topic is cross-contamination by allergens: in compliance with new pertinent EU legislation, industries are trying to set adequate standards specifically designed to avoid this risk with suppliers. Once again, an exchange of knowledge with research centres could help to pinpoint the best solution.

4. Eggs & Egg products *(Table 4)*

4.1 Most diffused fraud

Fraud can be potentially found in all different types of egg products: eggs, shelled eggs, egg powder. In some past cases, prohibited water additions were carried out to dilute liquid egg products for economic benefit.

At present, one of the most critical issues is the egg freshness, which makes a major contribution to the value of the product, because consumers may perceive variability in freshness as a lack of quality (Lin, Zhao, Sun, Chen, & Zhou, Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis, 2011).

Furthermore, another well-known type of fraud is the addition of incubated eggs to the fresh eggs, together with the addition of dyes.

The emerging risks related to this commodity are the use of conventional products fraudulently declared as organic and the introduction of Melamine (which appears to increase the protein content) in eggs.

4.2 Rapid approaches presented in literature

Scientific literature presents several rapid non-destructive methods to assess freshness of eggs and/or egg products: both NIR (Giunchi, Berardinelli, Ragni, Fabbri, & Silaghi, 2008, Suppl. Mat.) (Zhao, et al., 2010, Suppl. Mat.) (Lin, Zhao, Sun, Chen, & Zhou, Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis, 2011) and Vis-IR (Abdanan, Minaei, Hancock, & Karimi, 2014) (Liu, Ying, Ouyang, & Li, 2007, Suppl. Mat.) spectroscopies, coupled with chemometric data processing, can detect this type of fraud directly on the shelled egg. Moreover, AOS (Suman, Riani, & Dalcanale, MOS-based artifical olfactory system for the assessment of egg products freshness, 2007) or tests based on Hyperspectral Imaging are reported as other non-destructive methods for solving the same problem (Zhang, Pan, Tu, Zhan, & Tu, Non-destructive internal quality assessment of eggs using a synthesis of hyperspectral imaging and multivariate analysis, 2015), together with fast-GC electronic noses coupled with chemometric data analyses (Yimenu, Kim, & Kim, 2017); (Cavanna, Zanardi, Dall'Asta, & Suman, 2019).

Other suggested approaches for freshness evaluation use the intrinsic fluorescence of thick albumen and egg yolk (Karoui, Schoonheydt, Decuypere, Nicolaï, & Baedermaeker, 2007, Suppl. Mat.), the quantification of S-Ovalbumin (Huang, et al., 2012, Suppl. Mat.) or, again, a colorimetric test based on the reaction between albumen and 3.3',5.5'-tetramethylbenzidine (Rossi, Hidalgo, & Pompei, 2001, Suppl. Mat.).

With regard to Melamine detection, literature presents portable instruments that, thanks to a surface-enhanced Raman spectroscopy approach, can solve this problem (Cheng & Dong, 2011).

4.3 Confirmatory approaches presented in literature

With regard to the topic of freshness, scientific literature suggests that the assessment of albumen freshness, combining the results obtained with NIR and NMR spectroscopies, could be an effective method for resolving the problem (Kemps, et al., 2007, Suppl. Mat.).

Furthermore, the addition of dyes can be easily detected with LC-MS instruments (Liu, Hei, He, & Li, 2011), and this technique, if used for the evaluation of the carotenoids fingerprinting, also seems to be a promising tool for discriminating between organic and conventional eggs (van Ruth, et al., 2011).

Ultimately, chromatographic techniques are also relevant for the detection of Melamine (Rodriguez Mondal, Desmarcheller, Konings, Acheson-Shalom, & Delatour, 2010, Suppl. Mat.) (Deng, et al., 2010, Suppl. Mat.) (Xia, et al., 2009, Suppl. Mat.).

4.4 Industrial approach and needs

In this food chain, the application of paper trails and documental strategies to look for potential traceability incongruences is especially relevant (e.g. mass balance, etc…).

Prohibited addition of water can be easily checked by industries through dry matter testing.

Even though several rapid methods are available for the assessment of eggs freshness, industries do not generally implement these techniques in their routine controls, instead preferring to evaluate this type of fraud using the more common Lactic Acid quantification, as required by EU legislation (Reg. CE 853/2004).

Also, the illegal addition of incubated eggs is usually detected through enzymatic assays with a specific molecular marker target (3-hydroxybutyric acid) (Reg. CE 853/2004).

In agreement with the information presented in literature, the addition of dyes and false declaration of organic eggs are detected using chromatographic and mass spectrometric techniques.

On the contrary, due to a lack of expertise and relative high cost of the equipment, the industry usually outsources the analyses of Melamine, together with the detection of veterinary drugs, PCB and dioxins, to external certified laboratories.

5. Meat *(Table 5)*

5.1 Most diffused fraud

At present, adulteration of meat products is becoming a serious problem and fraud is diffused at different levels of the supply chain. In the recent years, several scandals regarding meat derived

products catalysed public opinion, in particular the partial replacement of beef with horse meat in some ready-to-eat products commercialized by famous brands.

Meat is often exposed to adulteration, the most common being: (i) false indication of the origin of meats and/or the animal feeding regime (for example in organic/PDO products) (Monahan, Schmidt, & Moloney, 2018, Suppl. Mat.), (ii) substitution of the species or replacement of meat with fat, (iii) missing declaration about a previous meat process (irradiation or thawing) (Leygonie, Britz, & Hoffman, 2012, Suppl. Mat.), (iv) possible presence of additives (Ballin, 2010) or (v) the use of anabolic hormones to increase muscle mass.

5.2 Rapid approaches presented in literature

Literature is full of examples of rapid methods to detect substitution of meat species: pork addition to meatballs or similar products, for example, can be rapidly detected with FTIR (Rohmana, Sismindari, Erwanto, & Che Man, 2011, Suppl. Mat.) or NIR spectroscopies (Kuswandi, Cendekiawan, Kristiningrum, & Ahmad, 2015, Suppl, Mat.) with multivariate data processing.

Adulterants can be also found in minced beef, and the same spectroscopic techniques mentioned above can detect them (Ropodi, Pavlidis, Mohareb, Panagou, & Nychas, 2015) (Kamruzzaman, Makino, Oshita, & Liu, 2015, Suppl. Mat.) (Morsy & Sun, 2013, Suppl. Mat.), potentially also in addition to UV-Vis spectroscopy (Alamprese, Casale, Sinelli, Lanteri, & Casiraghi, 2013, Suppl. Mat.).

Still in the spectroscopic field, MIR-ATR is a technique that, when coupled with chemometrics, can detect the addition of by-products like heart, liver and kidney in beef burgers (Zhao, Downey, & O'Donnell, 2014, Suppl. Mat.). Literature also reports an ELISA (Enzyme-Linked Immunosorbent Assay) test kit as a method used to detect undeclared species (Ayaz, Ayaz, & Erol, 2006, Suppl. Mat.).

Furthermore, contamination of samples with residues of horse meat can be checked by specific LFIA (Lateral Flow Immunoassay) devices (Masiri, et al., 2017, Suppl. Mat.) (Ha, et al., 2018).

A strictly correlated requirement is the development of methods able to perform meat speciation: IR spectroscopy (Al-Jowder, Kemsley, & Wilson, 1997, Suppl. Mat.) and RT-PCR (Yin, et al., 2016, Suppl. Mat.) (Natonek-Wiśniewska, Krzyścin, & Piestrzyńska-Kajtoch, 2013, Suppl. Mat.) can solve this problem, while literature also presents a rapid method that uses an optical thin-film biosensor chip with colour changes that can be perceived by the naked eye (Wang, Zhu, Chen, Xu, & Zhou, 2015).

With regard to lacking declaration of a previous meat process, literature once again highlights methods like NIR hyperspectral imaging and chemometrics (Barbin, Sun, & Su, 2013, Suppl. Mat.)

or the measurement of Aconitase activity (Škorpilova, Šimoniova, Rohlik, & Pipek, 2014) as possible methods for detecting this type of fraud.

A final point concerning rapid methods: literature suggests that, thanks to the analysis of muscles with DART-TOF-MS, chickens can be differentiated according to the diets they followed during life (Cajka, Danhelova, Zachariasova, Riddellova, & Hajslova, 2013). Therefore, illegal dispensation of feed with chicken bone meal can be detected.

5.3 Confirmatory approaches presented in literature

The confirmatory technique generally considered as the most useful for meat substitution is the RT-PCR (Ali, et al., 2012, Suppl. Mat.) (Soares S. , Amaral, Oliveira, & Mafra, 2013, Suppl. Mat.): for instance, several papers highlight the possibility to detect murine meat in mutton samples (Fang & Zhang, 2016, Suppl. Mat.), canine meat in burger formulations (Rahman, et al., 2015, Suppl. Mat.) and buffalo meat adulteration (Sakaridis, Ganopoulos, Argiriou, & Tsaftaris, 2013, Suppl. Mat.).

Exploring other confirmatory possibilities, chromatography also allows the detection of meat additions, with identification of some compounds, like Myoglobin, as markers of the adulteration (Giaretta, Di Giuseppe, Lippert, Parente, & Di Maro, 2013, Suppl. Mat.).

As concerns the possibility of performing meat speciation, the applied techniques are NMR spectroscopy (Jakes, et al., 2015, Suppl. Mat.) and the LC-MS approaches, which can detect specific biomarkers of different meat species (Montowska, Alexander, Tucker, & Barrett, 2015, Suppl. Mat.) (von Bargen, Brockmeyer, & Humpf , 2014, Suppl. Mat.) (Ruiz Orduna, Husby Yang, Ghosh, & Beaudry, 2015, Suppl. Mat.) (Prandi, et al., 2017, Suppl. Mat.) (Prandi, et al., 2018).

The use of non-declared irradiated meat can be detected with mass spectrometric approaches (LC-MS and GC-MS) (Chen, Fan, Song, Liu, & Zhang, 2013, Suppl. Mat.) (Li, Ha, Wang, & Li, 2010, Suppl. Mat.), and once again, NMR spectroscopy is also indicated as a powerful tool (Zanardi, et al., 2015).

Isotopic ratios evaluation is the typical approach used for the assessment of geographical origin of raw material, while the illegal use of anabolic hormones can be easily detected using HPLC-MS/MS, as reported in scientific papers since the start of the century (Draisci, Palleschi, Ferretti, Lucentini, & Cammarata, 2000).

5.4 Industrial approach and needs

With regard to the detection of meat substitution, the immunoassay approach is largely diffused in the industrial field, together with the more robust RT-PCR technique; in addition, some of the inteviewed experts also implement cromatographic methods to manage this issue.

On the contrary, for the meat speciation and for the detection of anabolic hormones, even knowing all the possibilities presented by the scientific community, industry prefers not to adopt them internally and, instead, outsources the analyses to certified laboratories.

In addition, the risk of lacking declarations of previous meat processes currently seems to be managed solely through paper-trails and audit protocols instead of analytical solutions, while the geographical origin of meat is evaluated outsourcing isotopic ratios analyses.

6. Honey *(Table 6)*

6.1 Most diffused fraud

Honey is a saturated solution of sugars, long used as a natural source of carbohydrates and is an important ingredient for food products and also for traditional medicine, thanks to its antimicrobial, anti-inflammatory and antioxidant effects.

The addition of exogenous sugars or other sweeteners is a well-known economically motivated fraud, despite

legislation dictating that nothing can be added to honey (EC. (2001). Council Directive 2001/110/EC of 20 December 2001 relating honey. Suppl. Mat.).

Another important parameter at high risk of fraud is the botanical origin of honey: monofloral honey varieties are perceived as better-quality products, and are the most appreciated by consumers with an increase in their market values.

In addition, also for this commodity, a critical point is the geographical origin declaration, together with the deliberate addition of enzymes (amylase, mainly), to keep an apparently high standard with a longer shelf-life, even starting with low quality raw materials.

6.2 Rapid approaches presented in literature

With regard to fraudulent additions to honey, some rheological parameters, such as viscosity, for example, are influenced by these additions (Yilmaz, et al., 2014, Suppl. Mat.); furthermore, NIR spectroscopy coupled with chemometrics is able to detect the presence of high fructose corn syrup (Chen, et al., 2011) and other sweeteners (Zhu, et al., 2010, Suppl.Mat.). The same results can be obtained with ATR-FTIR (Gallardo-Velázquez, Osorio-Revilla, Zuñiga-de Loa, & Rivera-Espinoza, 2009, Suppl. Mat.) and Raman spectroscopy (Li, Shan, Zhu, Zhang, & Ling, 2012, Suppl. Mat.).

Recently, the use of an electronic tongue, based on potential multistep pulse voltammetry, combined with PCA analysis is used to distinguish between pure honey and added syrup (Sobrino-Gregorio, Bataller, Soto, & Escriche, 2018), and a similar approach also discriminates different floral origins (Sousa, et al., 2014, Suppl. Mat.).

In addition, literature suggests that the evaluation of the diastase activity through a direct potentiometric measurement (Sak-Bosnar & Sakač, 2012) or with a spectrophotometric evaluation (Voldřich, Rajchl, Čížková, & Cuhra, 2009) could be interesting tools for the detection of enzymes addition.

6.3 Confirmatory approaches presented in literature

The saccharides profile obtained with liquid chromatography combined with chemometrics is a well-known confirmatory method for detecting added sugars (Wang, et al., 2015, Suppl. Mat.); as a valid alternative method, the same group of molecules can also be evaluated using HPAEC with pulsed amperometric detection (Morales, Corzo, & Sanz, 2008, Suppl. Mat.). High performance liquid chromatography combined with electrochemical detection (HPLC-ECD) is also efficient for detecting honey adulteration with various flavour enhancers (e.g. malthol, ethyl malthol, vanillin) (Liu C. et al., 2018).

Finally, $^{13}C/^{12}C$ isotope ratio evaluation (Çinar, Ekşi, & Coşkun, 2014) and NMR spectroscopy (Ribeiro, et al., 2014) are other common analytical approaches used for exogenous sweeteners detection in honey.

This last technique (Spiteri, et al., 2015) (Consonni, Cagliani, & Cogliati, 2013, Suppl. Mat.), together with ion chromatography (Fermo, Beretta , Maffei Facino , Gelmini , & Piazzalunga , 2013, Suppl. Mat.), DNA characterization (Soares S. , Amaral, Oliveira, & Mafra, 2015, Suppl. Mat.) and ICP techniques (Chua, Abdul-Rahaman, Sarmidi, & Aziz, 2012, Suppl. Mat.) are the most reported confirmatory approaches for botanical origin discrimination and geographical origin assessment.

6.4 Industrial approach and needs

Industrial experts declared that a first internal check is performed with the microscopy analysis on pollen, which can give preliminary but important information and ideas about possible risks of fraud. NMR spectroscopy and isotopic ratios evaluation are the most used techniques for the detection of sweetener addition and for geographical origin discrimination, but these analyses are usually outsourced to external laboratories, due to cost and expertise reasons.

On the contrary, the fraudulent presence of exogenous enzymes is controlled internally using immunoassay techniques.

7. Grains & Flour *(Table 7)*

7.1 Most diffused fraud

The most diffused fraud for this raw material is the addition of common wheat to durum wheat.

Durum wheat, indeed, is more expensive than common (bread) wheat and is considered of superior quality for the manufacture of pasta products; in Italian pasta, according to local legislation (Presidential Decree no. 187, 9 February 2001, Suppl. Mat.) a maximum of 3% of common wheat is permitted, because this could be the result of cross-contamination during agricultural practices.

Other grain adulterations are reported in literature, such as, for example, the contamination of spelt with soft wheat or the substitution of higher value species with lower value flours.

Another diffused fraud is adulteration with Melamine, used with the aim of increasing the apparent protein content in grains.

Recently, plant variety monitoring has become an emerging requirement for preserving the authenticity of some PDO breads.

7.2 Rapid approaches presented in literature

Concerning the detection of common wheat in durum wheat, NIR spectroscopy is suggested in literature as a useful technique (Cocchi, et al., 2006, Suppl. Mat.) (Vermeulen, Suman, Fernández Pierna, & Baeten, 2018).

In addition, Direct Analysis Real Time–High-Resolution Mass Spectrometry (DART–HRMS) seems able to discriminate different Triticum species (Miano et al., 2018), while ELISA kits can detect Melamine in wheat and flours (Garber & Brewer, 2010).

7.3 Confirmatory approaches presented in literature

Literature suggests several approaches for the detection of common wheat addition, most of them related to PCR techniques. Different possible molecular biology markers are identified for wheat (Bryan, Dixon, Gale, & Wiseman, 1998, Suppl. Mat.), flours (Casazza, et al., 2012), semolina and pasta (Terzi, Malnati, Barbanera, Stanca, & Faccioli, 2003, Suppl. Mat.), sometimes preferring the use of specific DNA microsatellite regions (Sonnante, et al., 2009, Suppl. Mat.). It was recently demonstrated that PCR assays can be also helpful in assessment of the authenticity of other grains. (Silletti, et al., 2018).

As regards the HPLC based technique for detection of the same fraud, several methods identify one or more compounds present only in the adulterated samples: for example, according to information presented by Barnwell *et al* more than twenty years ago, a peak related to a mixture of □/□-gliadins can be found only if common wheat is added to pasta (Barnwell, McCarthy, Lumley, & Griffin, 1994, Suppl. Mat.). Moreover, a more recent study suggests that a tryptic "Puroindoline a" peptide could be a suitable marker for common wheat adulteration at the flour level (Russo, et al., 2014, Suppl. Mat.).

Moving to another type of fraud, the contamination of spelt with soft wheat can be detected with an RFLP-LOC-CE (Restriction fragment length-lab on chip-capillary gel electrophoresis) method that employs the □-gliadin gene as the target (Mayer, et al., 2012).

In addition, adulteration with Melamine can be highlighted with robust chromatographic methods (Ehling, Tefera, & Ho, 2007, Suppl. Mat.), while plant varieties should be monitored, for example, using a DHPLC (Denaturing High Performance Liquid Chromatography) approach (Giancaspro, et al., 2016).

7.4 Industrial approach and needs

The grains industry over the last decade has mainly explored molecular diagnostic techniques potentialities (progressively substituting previous electrophoretic approaches) for the detection of common wheat adulterations in durum wheat. Both end-point and real-time approaches are indicated by the interviewed experts.

The substitution of higher value species with lower value flours is monitored internally with immunoassay tests, while the potential use of Melamine is evaluated with HPLC-MS methods.

8. Tomatoes and derivatives *(Table 8)*

8.1 Most diffused fraud

Tomatoes and derivatives should be monitored along the entire production chain to prevent adulteration, false geographical origin certifications and fraudulent organic farming declarations.

A common fraud of tomato sauce is the addition of sucrose, while another popular adulteration, which currently concerns different tomato-based foodstuffs, is the addition of Sudan dyes.

Another emerging requirement for this commodity is to be able to discriminate through different cultivars, in order to preserve PDO commodities and, more in general, to guarantee product quality.

8.2 Rapid approaches presented in literature

Scientific literature proposes several rapid methods with EN (Electronic Noses) and ET (Electronic Tongues) that can verify tomato freshness (Hong, Wang, & Qi, 2015, Suppl. Mat.) or the addition of overripe juices to fresh cherry tomato juices (Hong & Wang, 2014b, Suppl. Mat.).

Sucrose addition can be qualitatively and quantitatively detected by multispectral imaging combined with PLS (Partial Least Squares), least squares-support vector machines (LS-SVM), and back propagation neural network (BPNN) (Liu, Hao, Su, Chen, & Zheng, 2017), while spectroscopic techniques seem to be a rapid method for discriminating different tomato cultivars (Shao, et al., 2015, Suppl. Mat.).

Lastly, according to the work presented by Novotná *et al*, ambient mass spectrometry coupled with chemometrics seems to be a rapid method for discriminating organic and conventional samples as a function of their metabolomic fingerprint (Novotná, et al., 2012, Suppl. Mat.).

8.3 Confirmatory approaches presented in literature

Literature suggests that the addition of Sudan dyes can be detected with MS techniques such as APCI-TOF-MS. (Sciuto, et al., 2017, Suppl. Mat.); furthermore, the popular approaches for discrimination of different tomatoes cultivars include SSR (Simple Sequence Repeats) fingerprinting (Scarano, Rao, Masi, & Corrado, 2015), 5S rRNA regions (Sun, et al., 2014, Suppl. Mat.) or headspace GC-MS, with chemometric evaluation of the volatile metabolites (Figueira , Câmara, Pereira, & Câmara, 2014).

For the discrimination between organic and conventional farming, NMR profiling with linear discriminant analysis seems promising (Hohmann, Christoph, Wachter, & Holzgrabe, 2014), while the study of the mineral composition seems the most suitable approach reported in literature to trace the geographical origin of products (Bontempo, et al., 2011, Suppl. Mat.).

8.4 Industrial approach and needs

Instead of using rapid methods for adulteration detection, some virtuous enterprises have set standards to certify that the entire supply chain is monitored: the farmers and the retailers, documenting through paper trail approaches, must declare what they use during the entire year campaign; moreover, industries must also apply field testing and reward the achievements of quality goals defined with other partners.

As for false organic farming declarations, the common "indirect" approach adopted by the interviewed experts is pesticides analysis: the use of these compounds is not allowed if a cultivation is declared "organic", so a positive result for one or more pesticides should demonstrate the fraudulent "organic" product declaration.

Pesticide profiling is also a method used by the producers to obtain information about provenience from foreign countries such as Asia region, where some active ingredients that are banned in Europe are still permitted.

Continuing on this topic, recently some virtuous large enterprises have initiated successful collaborations with authorities to set up protocols, constraints and objective parameters pertinent to geographic characterization at regional levels, specifically based on IRMS (Isotope Ratio Mass Spectrometry) techniques, as suggested in literature.

9. Fruits & Vinegar *(Table 9)*

9.1 Most diffused fraud

For both fruit juices and vinegars, the most typical fraud is adulteration with exogenous compounds that have a lower price: for vinegar, for example, alcohol of different origins is largely used, whereas for fruit juices, water, cheaper fruit juices or industrial sweeteners are mixed with the natural products. Within this scenario, requirements for methods to protect the PDO productions (i.e. balsamic vinegar of Modena) have dramatically increased in the last decade.

9.2 Rapid approaches presented in literature

Scientific literature suggests that the use of spectroscopic techniques coupled with chemometric data treatments could be a rapid method for detecting product adulteration (Boggia, Casolino, Hysenaj, Oliveri, & Zunin, 2013, Suppl. Mat.) (Sáiz-Abajo, González-Sáiz , & Pizarro, 2005, Suppl. Mat.) (Vardin, Tay , Ozen, & Mauer, 2008, Suppl. Mat.).

In addition, the same techniques can discriminate different vinegars or juices (Snyder, Sweeney, Rodriguez-Saona, & Giusti, 2014) (Xie, Bu, & Peng, 2012, Suppl.Mat.) (Liu, et al., 2011, Suppl. Mat.), and the sensors field is also presented as a rapid method for classifying different vinegar brands (Zeng, Cao , Liu, Chen, & Ren, 2015).

9.3 Confirmatory approaches presented in literature

With regard to product adulteration, chromatographic techniques can detect specific markers of these activities (Abad-García, et al., 2014) or record metabolomic fingerprints that discriminate between "real" and "addicted" foodstuffs (Vaclavik, Schreiber, Lacina, Cajka, & Hajslova, 2012, Suppl.Mat.). Other techniques suggested by literature are the use of SP-LDI Mass Spectrometry or the use of the micellar electro-kinetic chromatography (MEKC): the first method detects exogenous compounds addicted to the balsamic vinegar (Guerreiro, de Oliveira , Ferreira, & Catharino, 2014) and the second one detects the amino acid L-Asparagine as marker of apple juice addition to pomegranate juice (Tezcan, Uzaşçı, Uyar, Oztekin, & Erim, 2013, Suppl. Mat.).

Other possible approaches that can be used for the detection of product dilutions or additions are RT-PCR techniques (Pardo, 2015) and IRMS and NMR measurements. (Thomas & Jamin, 2009, Suppl.Mat.) (Camin, et al., 2013)

These last two techniques are also used in literature to certify the geographical origin of PDO "Balsamic vinegar of Modena" (Werner & Roßmann, 2015) (Papotti, et al., 2015, Suppl.Mat.) (Consonni, Cagliani, Rinaldini, & Incerti, 2008a, Suppl. Mat.).

9.4 Industrial approach and needs

Industry is increasing the use of rapid NIR analysis for fingerprinting and calibration as official methods for parameters such as acidity, sugar content and alcoholic strength, which could represent a first sign of manipulation.

Moreover, chromatographic approaches (with PDA or MS detectors) are also applied internally in production plants for adulterations detection, while PCR, NMR and isotopic techniques are applied, but outsourced to external laboratories.

10. Cheese – Milk derivatives *(Table 10)*

10.1 Most diffused fraud

Milk and milk products are largely produced and consumed in Europe because of their high nutritional values; they are often subjected to fraud that can affect the entire chain (from raw milk to the final products).

One of the most common illegal activities is the adulteration of the finished products: powdered milk or several exogenous compounds (like urea, whey and hydrogen peroxide) could be added to fresh milk; undeclared mixtures of milks coming from different species could be used during the production of some PDO cheeses (like "Mozzarella di Bufala") or some exogenous fats can be added to butter.

Other types of fraud considered as crucial problems are dilution with water and Melamine additions, with the latter representing a direct risk for consumers' safety.

Moreover, with regard to the other commodities presented, authenticity and geographical origin of the products should be monitored to prevent false declarations, especially to protect some PDO products like "Parmigiano Reggiano" or "Mozzarella di bufala campana".

10.2 Rapid approaches presented in literature

Adulteration of finished products can be detected with IR spectroscopy, one of the rapid approaches largely presented in literature (Santos, Pereira-Filho, & Rodriguez-Saona, 2013, Suppl. Mat.) (Fadzlillah, Rohman, Ismail, Mustafa, & Khatib, 2013, Suppl. Mat.) (Jaiswal, Jha, Kaur, & Borah, 2017, Suppl. Mat.).

In addition, researchers also mention turbidimetrics (Scholl, Farris, & Mossoba, 2014, Suppl.Mat), DART-HRMS (Hrbek, Vaclavik, Elich, & Hajslova, 2014), Raman spectroscopy (Yazgan Karacaglar, Bulat, Boyaci, & Topcu, 2018) and MALDI-TOF-MS (Calvano, Monopoli, Loizzo, Faccia, & Zambonin, 2013b, Suppl. Mat.) methods as potential solutions to detect the common adulterations.

Water addition can be detected using DSC techniques (Tomaszewska-Gras, 2012, Suppl. Mat.) or through evaluation of electrical (Banach, Żywica, Szpendowski, & Kiełczewska, 2012) or physical parameters.

Melamine addition can be identified with Raman (Ma, et al., 2013, Suppl. Mat.) or ATR-FTIR (Limm, Karunathilaka, Yakes, & Mossoba, 2018), while PDO products and geographical origin can be controlled with IR spectroscopies (Woodcock, Fagan, O'Donnell, & Downey, 2008, Suppl. Mat.) (Gori, Maggio, Cerretani, Nocetti, & Caboni, 2012) or DART-HRMS (Hrbek, Vaclavik, Elich, & Hajslova, 2014).

10.3 Confirmatory approaches presented in literature

Chromatography (Schneider, Werkmeister, Becker, & Pischetsrieder, 2011, Suppl. Mat.), IRMS (Capici, Mimmo, Kerschbaumer, Cesco, & Scampicchio, 2015) and NMR (Fadzillah, et al., 2015, Suppl. Mat.) are the most commonly described approaches for adulteration detection, but other methods, such as the isoelectric focusing (Fuselli, et al., 2015, Suppl. Mat.) should also be considered.

Chromatographic methods are also used for Melamine identification (Ihunegbo, Tesfalidet, & Jiang, 2010, Suppl. Mat.), for PDO protection and for geographical origin evaluation (Ferreira & Caçote, 2003, Suppl. Mat.); in addition, for the last two types of fraud, NMR (Mazzei & Piccolo, 2012) (Brescia, Monfreda, Buccolieri, & Carrino, 2005, Suppl. Mat.) and Real Time-PCR (Rentsch , et al., 2013, Suppl. Mat.) are also valid possibilities presented in literature.

10.4 Industrial approach and needs

Similar to tomatoes, there are some virtuous enterprises in the dairy sector that are trying to monitor the entire production chain: farmers and retailers, documenting through paper trail approaches, must declare what they use during the entire year; moreover, industries must apply field testing and define quality goals to be reached with other partners.

As for the analytical approaches applied in plants, product adulterations and Melamine addition are commonly detected with NIR instruments, while water addition is controlled through evaluation of electrical or physical parameters.

In addition, LC-MS and immunoassay tests are the most commonly employed methods for quality control in production sites to preserve the authenticity of PDO products.

As concerns the geographical origin of products, producers know the strategic importance of this topic on the market: for this reason, over the last decade some virtuous large enterprises initiated successful collaborations with authorities to set up protocols, constraints and objective parameters for geographic characterization at regional levels.

11. Mushrooms

11.1 Most diffused fraud

Mushrooms come from every region of the world, growing in different periods and climate conditions throughout the year: therefore, the import/export flows are quite complex. The supply chain is usually very fragmented, starting from the single pickers (critical point of the chain) and arriving at the final distributors.

There is a high risk of commercialization of species which are different from the declared species, with an added value difference; many buyers look often only at prices and do not adequately check corresponding quality.

Another typical fraud is the undeclared treatment with sulphites to improve the brightness and appearance of mushrooms and increasing their shelf life.

11.2 Rapid approaches presented in literature

Very few scientific papers are focused on fraud risks in the mushroom chain; so far literature seems to be more interested in the evaluation of the nutritional parameters and edibility of mushrooms.

To manage authenticity issues, once again NIR spectroscopy seems to be the most effective rapid approach: according to information presented in literature, for example, this technique coupled with chemometric class-modelling seems able to confirm the authenticity of "*porcini* mushrooms" (Casale, et al., 2016).

11.3 Confirmatory approaches presented in literature

As far as we know, no confirmatory methods have been developed for fraud detection.

11.4 Industrial approach and needs

For mushrooms authenticity assessment, in general, producers have one or more skilled mycologist experts, who are devoted to checking all the entering raw materials, mainly through microscopy and botanical analysis, as alternatives to other more expensive analytical methods.

Often there is also a second counter-check by a mycologist sent from a control authority; in addition, these small-medium producers can count on a number of trained employees directly on the production lines, highly experienced in identifying potential risks during their routine activities.

Scientific literature demonstrated that some mushrooms species can retain radionuclides and toxic elements from the soil and the air (Loaiza, et al., 2012) (de Castro, Maihara, Silva, & Figueira, 2012); for this reason, mushrooms that come from regions of the world exposed to full-blown risk of radiation (e.g. recent years in Russia or Japan) undergo radioactivity controls with a final certification from an authority that allows the release of the products for food production or directly into the market.

Ultimately, producers highlighted that the risk of fraud increases when considering chopped mushrooms instead of whole mushrooms, but in this case the genetic speciation, which could easily detect this kind of adulteration, is commissioned to external laboratories only in a few, specific cases.

Producers final remarks

The open discussion with food industrial experts highlighted some interesting key points that could be helpful for the global scientific community: the vulnerability of "highly fragmented" chains (with a high number of steps, players and geographical areas) is perhaps the most crucial issue for food fraud, followed in most cases by the absence of implementation of constantly applied rapid screening procedures. This second point is in part the consequence of a limited level of awareness of risks and threats in the food fraud scenario within the stakeholder groups, in particular in small-medium enterprises.

The sensitivity of the food industry to investing money into analytical monitoring plans with a set of rapid and confirmatory analysis optimized for different chains needs to be amplified, concretely transmitting the perception that this investment is a sort of insurance to prevent much more serious economic-image damages that could arise after a public domain food fraud case.

At the same time, the food industry needs to invest into new interdisciplinary professional figures, who must combine different skills, from analytical chemistry to food technology, to quality assurance and marketing-purchasing.

The virtuous path and counterattack that industries would like to put in place are based on different strengths. Firstly, today many food tests that are still at the "research stage" should be implemented into a "routine scenario" for more effective and wider prevention and control, adopting appropriate certification standards and accreditations in the meantime; for this reason, the desire to expand and consolidate the scientific support network is increasing.

Ideally, production lines should be monitored continuously, performing multidirectional analytical approaches related to the sensory field, as well as to chemical & biomolecular fields, with both targeted and untargeted approaches; regular inspections along the entire supply chain should be performed, with the aim to map and identify any emerging fraud risk.

Other crucial points that industries highlighted are the need to better establish separate assessment & prevention models for safety and authenticity issues, and the desire to achieve a standardized database for food profiling, sharing data in the scientific community while still ensuring confidentiality channels.

Their suggestion on this topic is to make efforts on reducing the economic impact of "databases re-construction" year by year.

Finally, there is an increasing demand for the internal setting of precise purchasing technical specifications, characterized by more restrictive conditions than those directly related to the legislation in force, in order to discourage adulterations.

Conclusion

This paper presented a comparison of the analytical approaches suggested in scientific literature for fraud detection in several food chains with the real actions implemented by producers in their production plants.

Food producers declared to know a relevant number of fraud studied in literature, but sometimes they do not apply concrete methods in their facilities to identify all the corresponding sophistications/adulterations.

It seems that there are only few situations where the academic and industrial production worlds are aligned, and in most cases the methods applied by academic researchers are not implemented in the plants.

This relates to the fact that not all food producers stay updated on what scientific literature recommends; at the same time, some methods, even if known, are not designed for routine application to meet industry needs (due to skills/costs that are not always affordable): therefore, at present these are only efficient in a research laboratory.

In the future, there is an outlook for two worlds to continue to interact and "contaminate/cross fertilize" each other: therefore, food industries will have more sophisticated techniques available to them, and researchers will be encouraged to develop reliable methods that could also be applied to routine analyses, and not only used from a research standpoint.

Acknowledgement

This work was supported by the European Union, Seventh Framework Program (FP7/2007-2013), under Grant Agreement 613688 (Ensuring the Integrity of the European food chain _ FoodIntegrity).

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common Approach / Applied Research	Short method description / Basic Research	Reference
Barley or corn addition		Several NIR spectra recorded and analysed with PLS regression DRIFTS spectra acquisition followed by a PCA analysis SPME-GC-MS analysis of the headspace volatiles. PCA applied to the chromatographic data obtained Adulteration with corn detected with DSC and FTIR as spectroscopic techniques combined with PCA and PLS models as statistical tools FTIR and chemometric data analysis are used for the detection of multiple adulterants in ground and roasted coffees	(Ebrahimi-Najafabadi, et al., 2012) (R) (Reis, Franca, & Oliveira, 2013) (S.M.) (R) (Oliveira , Oliveira, Franca, & Augusti, 2009) (S.M.) (C) (Brondi, Torres, Garcia, Jerusa, & Trevisan, 2017) (R) (Reis, Botelho, Franca, & Oliveira, 2017) (R) (S.M.) (Reis, Franca, & Oliveira, 2016) (R) (S.M.)
Foreign materials (i.e. husks, sticks) addition	Carbohydrate profile using HPAEC-PAD	UV-Vis spectra recording and SPA-DA data processing DRIFTS spectra acquisition followed by a PCA analysis Fingerprint recording with direct infusion ESI-MS and PCA data elaboration Carbohydrate analysis with HPLC-HPAEC-PAD and post column derivatization HPLC-UV-Vis. Data comparison with PCA	(Souto, et al., 2015) (R) (Reis, Franca, & Oliveira, 2013)(S.M.) (R) (Aquino, et al., 2014) (S.M.) (R) (Domingues, et al., 2014) (C)
Not declared Arabica and Robusta mixture	FTIR/NIR screening Sensory analysis and expert panel tasting	NIR spectra recorded and analysed with PLS Regression with a specific data pre-processing method PCR-RFLP analysis Discrimination of different varieties in blends by means of UV-VIS and Fluorescence spectroscopies with PCA-MLR models as statistical tools	(Pizarro, Esteban-Diez, & Gonzalez-Saiz, 2007)(S.M.) (R) (Spaniolas, Tsachaki, Bennett, & Tucker, 2008) (S.M.) (C) (Dankowska, Domagała, & Kowalewski , 2017) (R)
Soybean, wheat and/or rice addition		Carbohydrates fingerprint recorded with HPAEC-PAD. PCA and LDA data analysis. Oligosaccharide profile recorded with UPLC-HRMS and PCA data elaboration.	(Pauli, et al., 2014) (S.M.) (C) (Cai, Ting, & Jin-Ian, 2016) (C)
Geographical origin		NMR spectroscopy and chemometric elaboration for geographical origin discrimination Adulteration of Colombian variety with grains from other countries checked by ^{1}H-NMR, ATR-MIR and NIR combined with classification models for species or origin discrimination	(Consonni, Cagliani, & Cogliati, 2012) (S.M.) (C) (Arana, et al., 2015) (C) (Medina, et al., 2017) (R)
Coffee purity		SPME-GC-MS followed by PLS regression for data analysis SPME- GC x GC analysis Analysis of four adulterants in Brazilian Arabica blends with ^{1}H NMR as fast technique	(Toledo, Hantao , Ho, Augusto, & Anderson, 2014) (S.M.) (C) (Mondello, et al., 2004) (S.M.) (C) (Ribeiro, Boralle, Redigolo Pezza, Pezza, & Toci, 2017) (C)

Table 1: Coffee (S.M.: reference can be found in the supplementary materials)

RISK	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
Addition of unfermented cocoa beans		FT-NIR spectra recorded and analysed with SVM and Si-PLS	(Ernest, Xing-yi, Wu, & Huang, 2014) (R)
Discrimination through different cocoa beans varieties and geographical origin fraudulent declaration	IRMS	FT-NIR, electronic tongue and data fusion obtained with these two techniques for accurate classification of cocoa beans	(Teye, Huang, Takrama, & Haiyang, 2014a) (S.M.)(Teye, Huang, Dai, & Chen, 2013) (S.M.) (Teye, Huang, Han, & Botchway, 2014b) (S.M.)(R)
		Total amount of Cd and Pb and their bioavailability correlated to the geographical origin of different cocoa powder samples	(Mounicouy, Szpunary, Andreyz, Blakez, & Lobinsky, 2003) (R) (S.M.)
		PCR and PCR-RFLP used to discriminate through different Theobroma species	(Haase & Fischer, 2007)(S.M.) (C)
		NMR coupled with chemometrics allows the geographical identification of cocoa beans analysing their metabolic profile	(Caligiani, Palla, Acquotti, Marseglia, & Palla, 2014)(S.M.) (C)
		LC-TOF-MS coupled to PLS-DA allows the geographical discrimination of cocoa beans using Polyphenols as markers	(Hori, Kiriyama, & Tsumura, 2016) (C)
		Evaluation of N and C isotope ratios in fermented cocoa bean samples. PCA and PLS-DA for geographical discrimination	(Diomande, et al., 2015) (C)
False quality composition declaration	Outsourced to external laboratories	Cocoa butter characterization for specific markers with ATR-MIR spectroscopy and multivariate analysis	(Maurer & Rodriguez-Saona, 2013) (S.M.) (R)
		GC-MS evaluation of chiral compounds (hydroxyl acids, catechins, amino acids) as quality markers of the cocoa beans	(Caligiani, Cirlini, Palla, Ravaglia, & Arlorio, 2007) (S.M.) (C)
		Chocolate samples classified according to the cocoa content thanks to MALDI-TOF-MS technique	(Bonatto & Silva, 2015) (C)
Exogenous compounds addition		FTIR with ATR, followed by PLS, is used to detect the presence of lard in chocolate formulation	(Che Man, Syahariza, Mirghani, Jinap, & Bakar, 2005)(S.M.) (R)
		NIR spectroscopy followed by PCR and PLS data elaboration for the detection of potato starch addition	(Liu, Zhao, Li, Chen, & Zhao, 2015) (S.M.) (R)
		Polysaccharides HPLC fingerprint profiling coupled with PCA allows the detection of exogenous plant materials	(Yang, et al., 2015) (C)
Cocoa Butter Equivalents (CBE) addition to Cocoa Butter (CB)	Free fatty acid composition and Triacylglycerol profile evaluation using Gas Chromatography	FTIR spectroscopy and chemometrics used to clearly identify CBE addiction from 2% to 5% in chocolate	(Goodacre & Anklam, 2001) (S.M.) (R)
		Detection of CBE addition to CB analyzing the triacylglycerol profile with HPLC-ELSD and subsequent statistical evaluation	(Dionisi, et al., 2004)(S.M.) (C)
		Fatty acids concentrations and their isotopic data are treated with PCA	(Spangenberg & Dionisi, 2001) (S.M.) (C)
		Triacylglycerol profiles of cocoa butters obtained with Gas Liquid Chromatography and subsequent statistical evaluation	(Buchgraber, Androni, & Anklam, 2007)(S.M.) (Buchgraber, Senaldi, Ulberth, & Anklam, 2004)(S.M.) (C)
		CBE Identification based on the analysis of sterol and triterpene alcohol degradation products formed during the flat processing	(Macarthur, Crews, & Brereton, 2000) (S.M.) (C)
		Content of Milk Fat in chocolate and the potential addition of Lauric fat determined by GLC of triglycerides	(Pontillon, 1995) (S.M.) (C)
		Adulteration with CBEs detected by Fluorescence and UV-VIS spectroscopies; data fusion of the two spectroscopies and creation of PCA and LDA models	(Dankowska, 2017) (C)
Addition of carob flour		Adulteration of powder with carob flour checked by NIR combined with PCA and PLS-DA for qualitative analysis and PLS regression for quantitative analysis	(Quelal-Vasconez, Perez-Esteve, Arnauld-Bonachera, Barat, & Taulens, 2018) (R)
Hazelnut traces detection		PCR based on the ITS marker facilitates the detection of hazelnut traces in chocolate, preventing ingestion of hidden allergens	(Lopez-Calleja, et al., 2013) (C)

Table 2: Cocoa & Chocolate (S.M.: reference can be found in the supplementary materials)

SPICE	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
Paprika	Geographical origin with stable isotopes analysis; Chromatography and spectroscopy investigations to assess the purity of the spices. All these analyses are outsourced to external laboratories	Synchronous fluorescence technique coupled with PLS-DA allows the detection of forbidden Sudan I dye addition	(Di Anibal, Rodriguez, & Albertengo, 2015) (S.M.) (R)
		UV-Vis spectroscopy coupled with different chemometric techniques for the detection of Sudan I, II, III and IV dyes addition	(Di Anibal, Odena, Ruisánchez, & Callao, 2009) (R)
		Different adulterants in various amounts (Sudan I, Sudan IV, lead chromate, lead oxide, silicon dioxide, polyvinyl chloride, gum arabic) detected by FT-MIR spectroscopy and OCSIMCA as statistical tool	(Horn, Esslinger, Pfister, Fauhl-Hassek, & Riedl, 2018) (R)
Chili		A comparison between NIRS and Raman spectroscopy to detect Sudan I addition. PCA or PLS-DA models are applied	(Haughey, Galvin-King, Ho, Bell, & Elliott, 2015) (R)
		Simultaneous determination of eight illegal dyes in chili products with LC-MS/MS analysis	(Li, et al., 2013) (S.M.) (C)
Turmeric		UV-Vis spectroscopy coupled with different chemometric techniques for the detection of Sudan I, II, III and IV dyes addition	(Di Anibal, Odena, Ruisánchez, & Callao, 2009) (R)
Saffron		EN system to detect the aroma fingerprints and subsequent PCA analysis to highlight dyes or other exogenous components	(Heidarbeigi, et al., 2015) (S.M.) (R)
		The amount of 12 selected elements detected with ICP-MS, after LDA elaboration, allows the Saffron geographical discrimination	(D'Archivio, Giannitto, Incani, & Nisi, 2014) (S.M.) (C)
		An HPLC/PDA/MS method is able to detect marker molecules that identify a Saffron adulteration	(Sabatino, et al., 2011)(S.M.) (C)
		NMR spectroscopy with PCA and subsequent OPLS-DA brings to a discrimination between PDO and not PDO Saffron	(Cagliani, Culeddu, Chessa, & Consonni, 2015) (C)
		NMR spectroscopy coupled with PLS-DA is able to detect Saffron adulteration with plant adulterants	(Petrakis, Cagliani, Polissiou, & Consonni, 2015)(S.M.) (C)
		Adulteration with other cheaper plants detected by DRIFTS in the mid-infrared combined with PLS-DA for the authentication of saffron and PLS regression models for the quantification of each adulterant	(Petrakis & Polissiou, 2017) (R)
Curry		UV-Vis spectroscopy coupled with different chemometric techniques for the detection of Sudan I, II, III and IV dyes addition	(Di Anibal, Odena, Ruisánchez, & Callao, 2009) (R)
Cinnamon		Differentiation between “true” Cinnamon and “other” Cinnamon species using DART-QToF-MS and PCA data elaboration	(Avula, Smillie, Wang, Zweigenbaum, & Khan, 2015) (R)
Black pepper		Use of DNA barcoding approach to detect Chilli bio adulteration of traded black pepper powder	(Parvathya, et al., 2014) (C)
Spice mixtures		Spice mixtures for sausages and saveloy discriminated using portable EN and IMS with LDA and PCA elaboration respectively	(Banach, Tiebe, & Hubert, 2012) (R)

Table 3: Spices (S.M.: reference can be found in the supplementary materials)

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
False Freshness and quality declarations	Lactic acid quantification	Tests based on hyperspectral imaging with a combination of analytical techniques to determine the internal quality of eggs NIR spectroscopy with different chemometric techniques for non-destructive freshness assessment on shell eggs MOS-artificial olfactory system. A correlation with the legal freshness parameters is demonstrated Eggs freshness evaluated using the intrinsic fluorophores of thick albumen and yolk A Vis or VIS-IR wavelengths range (400-1100 nm) transmittance method allows the evaluation of intact chicken eggs quality S-Ovalbumin is presented as a reference index to express commercial shell egg freshness as equivalent egg age A colorimetric test based on the reaction between albumen and 3,3',5,5'-tetramethylbenzidine is used for freshness evaluation Albumen freshness is evaluated combining the results obtained with Vis-NIR and NMR spectroscopy GC-E nose for freshness discrimination and for prediction of storage times in hen eggs and in egg products	(Zhang, Pan, Tu, Zhan, & Tu, Non-destructive internal quality assessment of eggs using a synthesis of hyperspectral imaging and multivariate analysis, 2015)(R) (Giunchi, Berardinelli, Ragni, Fabbri, & Silaghi, 2008) (S.M.) (Zhao, et al., Identification of egg's freshness using NIR and support vector data description, 2010) (S.M.) (Lin, Zhao, Sun, Chen, & Zhou, Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis, 2011) (R) (Suman, Riani, & Dalcanale, MOS-based artifical olfactory system for the assessment of egg products freshness, 2007) (R) (Karoui, Schoonheydt, Decuypere, Nicolaï, & Baedermaeker, 2007)(S.M.) (R) (Abdanan, Minaei, Hancock, & Karimi, 2014) (Liu, Ying, Ouyang, & Li, 2007)(S.M.) (R) (Huang, et al., Estimation of egg freshness using S-Ovalbumin as an indicator, 2012)(S.M.) (R) (Rossi, Hidalgo, & Pompei, 2001) (S.M.) (R) (Kemps, et al., 2007)(S.M.) (C) (Yimenu, Kim, & Kim, 2017) (Cavanna, Zanardi, Dall'Asta, & Suman, 2019) (R)
Melamine addition	Outsourced to external laboratories	Portable surface-enhanced Raman Spectroscopy is used for fast detection of Melamine contamination, also at trace levels HPLC-MS/MS methods with specific sample pre-treatments for the simultaneous detection Melamine and Cyanuric Acid GC-MS coupled with UPLC-MS/MS methods	(Cheng & Dong, 2011) (R) (Rodriguez Mondal, Desmarcheller, Konings, Acheson-Shalom, & Delatour, 2010) (S.M.) (Deng, et al., 2010)(S.M.) (C) (Xia, et al., 2009) (S.M.) (C)
Conventional eggs declared as organic	Carotenoid profiling	Carotenoids fingerprint obtained with HPLC-PDA and KNN elaboration	(van Ruth, et al., 2011)(C)
Dyes addition	HPLC-MS/MS approach	Simple sample extraction procedure and UHPLC-MS/MS analysis	(Liu, Hei, He, & Li, 2011) (C)

Table 4: Eggs & Eggproducts (S.M.: reference can be found in the supplementary materials)

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
Meat adulteration	Immunoassay techniques Molecular diagnostic techniques LC-MS	FTIR or NIR spectroscopies coupled with chemometric techniques (PCA, PLS) are able to detect pork addition in beef meatball A Real-Time PCR method allows the detection of pork addition in meatballs Multispectral images and chemometric data elaborations for the detection of minced beef substituted with pork and vice versa Detection of beef offal adulterants like heart, liver and kidney in beef burgers with MIR-ATR spectroscopy and chemometrics UV-Vis, NIR and MIR techniques with multivariate data analysis and data fusion for the detection of adulterants in minced beef Detection of species adulterations in meat products using enzyme-linked immunosorbent assay A real-time PCR system allows the detection of murine meat in mutton samples A RT- PCR approach based on SYBR Green dye used for the quantitative detection of pork meat in processed meat products PCR followed by an HRM analysis developed for the detection of buffalo meat adulterations A lab-on-a-chip PCR-RFLP approach allows the detection of canine DNA in burger formulations An UPLC method highlight the presence of undeclared meat addition in beef burgers using Myoglobin as marker Adulteration of different kind of meat with pork meat determined by parallel reaction monitoring (PRM) mass spectrometry and screening of five specific peptides by shotgun proteomic approach based on tryptic digests of certain protein Contamination of samples with residues of horse meat checked by highly specific lateral flow immunoassay for a rapid identification of raw and cooked horse meat in very low amounts Tandem mass spectrometry to detect beef and pork meat in complex and processed food matrices using two different marker peptides specific for beef and pork meat LC-MS to identify and quantify eight different meat species (duck, rabbit, chicken, turkey, buffalo, equine, deer and sheep) in Bolognese sauce, using a species-specific peptide marker for each species	(Rohmana, Sismindari, Erwanto, & Che Man, 2011) (S.M.) (Kuswandi, Cendekiawan, Kristiningrum, & Ahmad, 2015) (S.M.) (R) (Ali, et al., 2012) (S.M.) (C) (Ropodi, Pavlidis, Mohareb, Panagou, & Nychas, 2015) (R) (Zhao, Downey, & O'Donnell, 2014) (S.M.) (R) (Alamprese, Casale, Sinelli, Lanteri, & Casiraghi, 2013) (S.M.) (Morsy & Sun, 2013)(S.M.) (Kamruzzaman, Makino, Oshita, & Liu, 2015)(S.M.) (R) (Ayaz, Ayaz, & Erol, 2006) (S.M.) (R) (Fang & Zhang, 2016) (S.M.) (C) (Soares S. , Amaral, Oliveira, & Mafra, 2013) (S.M.) (C) (Sakaridis, Ganopoulos, Argiriou, & Tsaftaris, 2013) (S.M.) (C) (Rahman, et al., 2015) (S.M.) (C) (Giaretta, Di Giuseppe, Lippert, Parente, & Di Maro, 2013) (S.M.) (C) (Pan, Chen, Chen, Huang, & Han, 2018) (S.M.) (C) (Masiri, et al., 2017) (S.M.) (R) (Prandi, et al., Mass spectrometry quantification of beef and pork meat in highly processed food: Application on Bolognese sauce, 2017) (S.M.) (C) (Prandi, et al., 2018) (C)
False meat classification	Immunoassay techniques	A Sheep-PCR-Strip approach for the identification of raw and heat-treated mutton	(Yin, et al., 2016) (S.M.) (R)

	Molecular diagnostic techniques LC-MS	Rapid method for the identification of meat species using optical thin-film biosensor chip with colour changes With MIR spectroscopy it is possible to distinguish minced chicken, pork and turkey meats and between fresh and thawed samples A PCR method allows the identification of bovine, porcine, ovine and chicken components in a wide range of animal products Thanks to a LESA-MS method, specific peptide markers for different meets species have been identified HPLC-MS/MS method able to discriminate horse and pork meat in different matrices An HRMS method identifies peptide markers able to discriminate through different meat species A screening protocol using 60MHz NMR spectroscopy allows the discrimination between beef and horse meat Adulteration with other undeclared meats checked by DNA lateral flow assay for a rapid screening of unknown adulterants and ELISA assay for the quantification of specific adulterants	(Wang, Zhu, Chen, Xu, & Zhou, 2015) (R) (Al-Jowder, Kemsley, & Wilson, 1997) (S.M.) (R) (Natonek-Wiśniewska, Krzyścin, & Piestrzyńska-Kajtoch, 2013) (S.M.) (C) (Montowska, Alexander, Tucker, & Barrett, 2015) (S.M.) (C) (von Bargen, Brockmeyer, & Humpf , 2014) (S.M.) (C) (Ruiz Orduna, Husby, Yang, Ghosh, & Beaudry, 2015) (S.M.) (C) (Jakes, et al., 2015) (S.M.) (C) (Ha, et al., 2018) (R)
Thawed meat declared fresh		For chicken meat, the results are obtained with the measurement of the Aconitase activity A NIR hyperspectral imaging method coupled with PLS-DA discriminates between fresh and thawed meat in porks	(Škorpilova, Šimoniova, Rohlik, & Pipek, 2014) (R) (Barbin, Sun, & Su, 2013) (S.M.) (R)
Addition of blood agents to meat food	Immunoassays techniques	With an HPLC-MS/MS method, fibrin peptides are identified as markers, since they are released from the blood protein fibrinogen	(Grundy, et al., 2007) (C)
Use of non-declared irradiated meat		O-Tyrosine and m-Tyrosine are specific markers detected by HPLC_FLD and HPLC_MS/MS methods A metabolic or lipid profiling study of non-irradiated and irradiated beef using NMR and chemometrics Specific hydrocarbons are detected in irradiated chilled beef using an HS-SPME-GC_MS technique	(Chen, Fan, Song, Liu, & Zhang, 2013) (S.M.) (C) (Zanardi, et al., 2015) (Zanardi, et al., 2013)(S.M.) (C) (Li, Ha, Wang, & Li, 2010) (S.M.) (C)
Use of anabolyc hormones	Outsourced to external laboratoris	The HPLC-MS/MS method presented is able to detect hormones and their metabolites in bovine serum and urine	(Draisci, Palleschi, Ferretti, Lucentini, & Cammarata, 2000) (C)
Geographical origin	IRMS / ICP-MS	Isotope analysis of organic elements like carbon and nitrogen	(Nakashita, et al., 2008) (C)
Indirect control of feed frauds		Thanks to a DART-TOFMS analysis, chicken muscles are discriminated according to the feed eaten	(Cajka, Danhelova, Zachariasova, Riddellova, & Hajslova, 2013) (C)

Table 5: Meat (S.M.: reference can be found in the supplementary materials)

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
Sweeteners addition	NMR and/or stable isotope analysis performed by external laboratories	NIR spectroscopy coupled with chemometrics allows the detection of sweeteners materials	(Chen, et al., 2011) (Zhu, et al., 2010) (S.M.)(R)
		Using FT-Raman spectroscopy, the adulterations with beet and cane invert are highlighted	(Paradkar & Irudayaraj, 2002) (S.M.) (R)
		Measuring their dielectric constant, samples that contain sucrose syrup as adulterant are discriminated from the pure honey ones	(Guo, Liu, Zhu, & Wang, 2011) (S.M.) (R)
		Raman spectroscopy and chemometric data elaborations for the detection of HFCS and maltose syrup in honey	(Li, Shan, Zhu, Zhang, & Ling, 2012) (S.M.) (R)
		Sugars addition changes some rheological parameters (i.e. viscosity) in honey. A comparison with the results obtained with HPLC-RID is presented	(Yilmaz, et al., 2014)(S.M.) (R)
		ATR-FTIR spectroscopy and chemometrics for the detection of syrups and sugars in Mexican honeys	(Gallardo-Velázquez, Osorio-Revilla, Zuñiga-de Loa, & Rivera-Espinoza, 2009)(S.M.) (R)
		HPLC and GC simultaneously used for honey fingerprints recording. PCA elaboration and exogenous sugars detection	(Cotte, Casabianca, Chardon, Lheritier, & Grenier-Loustalot, 2003) (S.M.) (C)
		$^{13}C/^{12}C$ ratio evaluation for the detection of honey adulterated with high fructose corn syrup	(Croft, 1987)(S.M.) (Çinar, Ekşi, & Coşkun, 2014) (Padovan, De Jong, Rodrigues, & Marchini, 2003)(S.M.) (C)
		Starch addition detected thanks to an HPLC saccharide profile comparison. A target peak is identified only in the adulterated sample	(Wang, et al., 2015)(S.M.) (C)
		GC-MS analysis for the detection of Inulotriose as marker of high fructose inulin syrups addition	(Ruiz-Matute, Rodríguez-Sánchez, Sanz, & Martínez-Castro, 2010) (S.M.) (C)
		NMR spectroscopy used for the detection of high fructose corn syrup adulteration. Comparison with physicochemical data	(Ribeiro, et al., 2014) ((Spiteri, et al., 2015) (C)
		An HPAEC-PAD method is able to find honey adulteration with various industrial bee-feeding sugar syrups	(Cordella, Militão, Clément, Drajnudel, & Cabrol-Bass, 2005)(S.M.) (C)
		An HPAEC-PAD oligosaccharide analysis is able to detect corn syrups and HFCS addition	(Morales, Corzo, & Sanz, 2008) (S.M.) (C)
		SCIRA-MS and SNIF-NMR are evaluated as a combined way to characterize honey samples and to combat adulterations	(Cotte, et al., 2007) (S.M.) (C)
		Detection of honey adulterated with maltooligosaccharides with an UPLC-ELSD method	(Zhou, et al., 2014) (S.M.) (C)
		Adulteration with various flavours enhancers (maltol, ethyl malthol, vanillin) determined by high performance liquid chromatography with electrochemical detection (HPLC-ECD) method	(Liu C. , et al., 2018) (C)
		Adulteration of monofloral varieties with various syrups in different amounts checked with an electronic tongue combined with PCA analysis to distinguish between types of pure honey and syrup, and with PLS analysis to determine the level of the adulterants in each honey	(Sobrino-Gregorio, Bataller, Soto, & Escriche, 2018) (R)
		Adulteration of different varieties with sucrose checked with electron Impedance Spectroscopy (EIS) and FT-MIR	(Das, et al., 2017) (R)
		Using a potentiometric ET, monofloral honeys with a high variability in floral origin are correctly classified	(Sousa, et al., 2014) (S.M.) (R)

False botanical origin declaration		Thanks to NMR spectroscopy, analytical criteria are defined to check the authenticity of both mono- and multi-floral honey Pyrolysis-mass spectrometry is presented as a way for the rapid discrimination of honeys from different botanical origin Different DNA extraction methods with sample pre-treatments are compared in order to improve the botanical origin identification The distribution of some IC ions (i.e. Na^{+},Ca^{2+} etc.) in honey samples is identified as indicator of botanical origin and authenticity	(Spiteri, et al., 2015) (C) (Radovic, Goodacre, & Anklam, 2001b) (S.M.) (C) (Soares S. , Amaral, Oliveira, & Mafra, 2015) (S.M.) (C) (Fermo, Beretta , Maffei Facino , Gelmini , & Piazzalunga , 2013) (S.M.) (C)
Exogenous enzymes addition		Evaluation of the diastase activity in nine varieties by direct potentiometric measurement of triiodide ion and evaluation of sensor and analyte parameters by means of least squares fitting of potentiometric data Determination of the hydrolytic activity of diastase through Schade test to check the addition of foreign amylase	(Sak-Bosnar & Sakac, 2012) (C) (Voldrich, Rajchl, Cizkova, & Cuhra, 2009) (S.M.) (R)
False geographical origin declaration	NMR and/or stable isotope analysis performed by external laboratories	NIR fingerprints of Corsican honey samples are analyzed by a range of chemometric tools to confirm their claimed provenance ICP-AES and ICP-MS used for the elemental profile. K and Na are found to be good markers for honey geographical origin HS-SPME with a GCxGC-TOF-MS detection is employed for fast characterization of volatiles, in order to confirm honey authenticity Some specific aroma marker compounds related to the floral and geographical origin are detected with a dynamic headspace GC-MS method The saccharides content evaluation with NMR spectroscopy allows the geographical discrimination of honey from three different floral sources	(Woodcock, Downey, & O'Donnell, 2009) (S.M.) (R) (Chua, Abdul-Rahaman, Sarmidi, & Aziz, 2012) (S.M.) (de Alda-Garcilope, Gallego-Picó , Bravo-Yagüe , Garcinuño-Martínez , & Fernández-Hernando , 2012) (S.M.) (C) (Cajka, Hajslova, Pudil, & Riddellova, 2009) (C) (Radovic, et al., 2001a) (S.M.) (C) (Consonni, Cagliani, & Cogliati, 2013) (S.M.) (C)

Table 6: Honey (S.M.: reference can be found in the supplementary materials)

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
Adulteration of durum wheat and other cereals with common wheat	Immunoassay tests NIR spectroscopy	NIR spectroscopy coupled with chemometrics allows the quantification of common wheat in durum wheat	(Cocchi, et al., 2006) (S.M.) (R)
		The fraud is detected with the evaluation of C17:0/C21:0 alkylresorcinol homologue ratio in the samples	(Knödler, Most, Schieber, & Carle, 2010) (S.M.) (C)
		In pasta samples, thanks to an RP-HPLC method, a peak due to a mixture of γ/β-gliadins can be found only if common wheat is added	(Barnwell, McCarthy, Lumley, & Griffin, 1994) (S.M.) (C)
		PCR methods with specific markers are able to identify common wheat addition along the entire production chain: from wheat to pasta through flours and semolina.	(Bryan, Dixon, Gale, & Wiseman, 1998) (S.M.) (Terzi, Malnati, Barbanera, Stanca, & Faccioli, 2003) (S.M.) (Casazza, et al., 2012) (Sonnante, et al., 2009) (S.M.) (C)
		An UHPLC-MS/MS method is able to detect a peptide of Puroindoline-a (Pin-a) that is present only in common wheat	(Russo, et al., 2014) (S.M.) (C)
		Thanks to the polymorphism in the γ-gliadin gene, using RFLP-LOC-CE techniques, spelt flour adulterations with soft wheat can be detected	(Mayer, et al., 2012) (C)
		Water-soluble proteins are extracted and analysed with CZE in order to detect durum wheat adulterations	(Piergiovanni, 2007) (S.M.) (C)
		Determination of the authenticity of wheat and other cereals with PCR assays and DNA fingerprinting through tubulin-based polymorphism (TBP)	(Silletti, et al., 2018) (C)
		Addition/contamination of exogenous nitrogen-rich adulterants to raise the protein content in foodstuffs determined by quantitative analysis of both high and low levels of 14 nitrogen-rich adulterants in various food matrixes by means of LC-MS/MS and isotopic diluition	(Frank, Bessaire, Tasser, Goyion, & Delatour, 2017) (C)
		DART–HRMS to discriminate different Triticum species and to check the risk of substitution of higher value wheat species with lower value flours.	(Miano, et al., 2018) (R)
		Near infrared (NIR) hyperspectral imaging as a tool for discriminating between both common and durum wheat at the singulated kernel and bulk sample levels	(Vermeulen, Suman, Fernández Pierna, & Baeten, 2018) (R)
Melamine contamination	HPLC -MS	Two ELISA kits are presented as able to rapidly detect melamine in wheat products	(Garber & Brewer, 2010) (R)
		Thanks to a specific HPLC method, Melamine and other by-products are simultaneously detected in cereal flours	(Ehling, Tefera, & Ho, 2007) (S.M.) (C)
Varietal traceability of durum wheat		DHPLC is used for setting up a SNP-based method that allows the varietal traceability of the durum wheat "Timilia" cultivar, in order to preserve the authenticity of Pane Nero di Castelvetrano	(Giancaspro, et al., 2016) (C)

Table 7: Cereals and Flours (S.M.: reference can be found in the supplementary materials)

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
False tomatoes freshness declaration		The freshness of cherry tomatoes is evaluated using an EN coupled with supervised and semi-unsupervised chemometric approaches	(Hong, Wang, & Qi, 2015) (S.M.)(R)
		Data fusion of the results obtained with an Electronic Tongue and an Electronic Nose to track freshness of cherry tomatoes squeezed for juices	(Hong & Wang, 2014c) (S.M.) (R)
Product adulteration		Several data fusion approaches of the results obtained with Electronic Tongue and Electronic Nose instruments are presented in order to detect fresh cherry tomato juices adulterated with overripe tomato juices at different percentage	(Hong, Wang, & Qiu, 2014a) (Hong & Wang, 2014b) (S.M.)(R)
		Thanks to the evaluation of $\delta^{18}O_{water}$, it is possible to determine if the tomato passata is genuine or has been obtained by diluting tomato paste	(Bontempo, et al., 2014) (C)
		Sucrose adulteration qualitatively and quantitatively detected by multispectral imaging combined with partial least squares (PLS), least squares-support vector machines (LS-SVM), and back propagation neural network (BPNN) as chemometric tools	(Liu, Hao, Su, Chen, & Zheng, 2017) (R)
		Adulteration of different foodstuffs with Sudan dyes detected by screening technique based on APCI-TOF-MS	(Sciuto, et al., 2017) (S.M.) (C)
False tomatoes breed and cultivar declaration		A Vis-NIR and chemometric method for tomatoes bred discrimination by spaceflight mutagenesis from leafs or fruits	(Shao, et al., 2015) (S.M.) (R)
		A multispectral imaging method with chemometrics for discrimination and identification of tomato cultivars of Nepal	(Shrestha, Deleuran, Olesen, & Gislum, 2015) (R)
		Tomato varieties are distinguished thanks to the analysis of the 5S rRNA regions	(Sun, et al., 2014) (S.M.)(C)
		Several tomato cultivars are discriminated according to their volatile metabolomic expression, obtained with HS-SPME-GC-MS and chemometric data analysis	(Figueira , Câmara, Pereira, & Câmara, 2014) (C)
		With the SSR fingerprint, a molecular genetic analysis is able to identify false “San Marzano” PDO tomatoes	(Scarano, Rao, Masi, & Corrado, 2015) (C)
False organic farming tomatoes declaration	Pesticides analyses	The fraud is detected with a metabolomic fingerprint approach with DART-TOF-MS	(Novotná, et al., 2012) (S.M.) (R)
		The fraud is detected thanks to NMR spectroscopy and LDA data elaboration	(Hohmann, Christoph, Wachter, & Holzgrabe, 2014) (C)
False geographical origin declaration	IRMS	Tomatoes provenance is evaluated analysing the mineral composition of the samples with an ICP-oa-TOF-MS approach	(Fragni, Trifirò, & Nucci, 2015) (S.M.) (C)
		$^{87}Sr/^{86}Sr$ isotope ratio evaluation, obtained with TIMS, is able to discriminate between Chinese and Italian tomatoes	(Trincherini, Baffi, Barbero, Pizzoglio, & Spalla, 2014) (C)
		Isotope ratios and mineral composition (obtained with ICP, IRMS and IC techniques) are used to trace the geographical origin of tomatoes and derivatives (juice, passata and paste)	(Bontempo, et al., 2011) (S.M.)(C)

Table 8: Tomatoes & derivatives (S.M.: reference can be found in the supplementary materials)

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
False origin declaration	IRMS / NMR spectroscopy	Using a cataluminescence sensor with LDA, this method is able to distinguish through different brands of vinegar	(Zeng, Cao , Liu, Chen, & Ren, 2015) (R)
		The fingerprints obtained with a chemiluminescence sensor array, analysed with LDA, bring to a discrimination of different vinegars	(Niu & Liu, 2014) (S.M.)(R)
		Vis-NIR spectroscopy, with different chemometric techniques, identifies the varieties of rice vinegars	(Liu, et al., 2011) (S.M.)(R)
		UV spectroscopy coupled with PCA and ANN is used to distinguish vinegar samples	(Xie, Bu, & Peng, 2012) (S.M.) (R)
		An FTIR spectroscopy method is used to predict the percent composition of concord juice in a grape juice blend	(Snyder, Sweeney, Rodriguez-Saona, & Giusti, 2014) (R)
		NMR spectroscopy coupled with multivariate data analysis is used to classify BVM and TBVM	(Papotti, et al., 2015) (S.M.) (C)
		C, H and O isotope analyses with IRMS and ^{2}H-NMR are used for the authentication of balsamic vinegars	(Werner & Roßmann, 2015) (C)
		^{13}C-NMR spectroscopy is used for the evaluation of Glucose and Fructose isoforms in order to certify the TBVM	(Consonni, Cagliani, Rinaldini, & Incerti, 2008a) (S.M.) (C)
Product adulteration	IRMS / NMR spectroscopy Polyphenols and organic acid profiling with chemometric elaborations	NIR spectroscopy with OSC is presented as a way to detect alcohol or molasses vinegar addition to wine vinegar	(Sáiz-Abajo, González-Sáiz , & Pizarro, 2005) (S.M.)(R)
		Fluorescence spectroscopy coupled with ICA is used for the detection of grapefruit juice added to orange juice	(Ammari, Redjdal, & Rutledge, 2015) (R)
		A UV-Vis and chemometric method highlights the addition of filler juice and water to pomegranate juices	(Boggia, Casolino, Hysenaj, Oliveri, & Zunin, 2013) S.M.) (R)
		Thanks to FTIR spectroscopy and chemometrics, the grape juice addition to pomegranate juice concentrate is detected	(Vardin, Tay , Ozen, & Mauer, 2008) (S.M.) (R)
		An SP-LDI-MS method is able to detect apple, alcohol or wine vinegar addition to balsamic vinegar	(Guerreiro, de Oliveira , Ferreira, & Catharino, 2014) (C)
		The $(D/H)_{CH3}$ values, obtained with SNIF-NMR, detect the molasses spirit or synthetic vinegar addition to rice vinegar	(Hsieh, Li, Cheng, & Ma, 2013) (S.M.) (C)
		^{2}H/^{1}H ratio by SNIF-NMR and ^{13}C/^{12}C ratio by IRMS are used to detect adulterations of vinegar using synthetic Acetic Acid; the ^{18}O/^{16}O ratio of water by IRMS allows the differentiation of wine vinegars from dried grape vinegars	(Thomas & Jamin, 2009)(S.M.) (Camin, et al., 2013) (C)
		Sugars determination, obtained with CZE-UV technique, coupled with chemometric data elaboration is able to detect cheaper juices addition to high-value juices (i.e. orange, pineapple)	(Navarro-Pascual-Ahuir, Lerma-García, Simó-Alfonso, & Herrero-Martínez, 2015) (S.M.)(C)
		A MEKC-LIF method suggests that the amino acid L-Asn can be used as a powerful marker of the apple juice addition to pomegranate juice	(Tezcan, Uzaşçı, Uyar, Oztekin, & Erim, 2013) (S.M.) (C)
		Real-time PCR analyses are used for the detection of mandarin in orange juices	(Pardo, 2015) (Aldeguer, López-Andreo, Gabaldón, & Puyet, 2014) (S.M.) (C)
		Thanks to some HPLC-MS methods, the metabolomic fingerprints of different juices are recorded and elaborated with chemometric techniques, allowing authenticity certification and adulterations detection	(Vaclavik, Schreiber, Lacina, Cajka, & Hajslova, 2012) (S.M.) (Jandrić, et al., 2014)(C)
		Several Polyphenols are determined in different juices with HPLC-PDA detection. For each type of juice, specific markers able to highlight fraudulent blending have been identified	(Abad-García, et al., 2014) (Silva, et al., 2000) (S.M.)(C)
False organic farming declaration		ICP-MS with PCA and SIMCA chemometric techniques for the discrimination of organic and ordinary grape juice samples	(Borges, et al., 2016) (C)

Table 9: Fruits & Vinegar (S.M.: reference can be found in the supplementary materials)

FRAUD	INDUSTRY FIELD	RESEARCH FIELD	
	Common approach	Short method description	Reference
False product authenticity declaration	Immunoassay techniques HPLC/MS-MS	NIR and MIR spectroscopies, coupled with chemometric techniques, are used to assess cheese quality and authenticity DART-HRMS technique coupled with chemometrics is able to discriminate among milks obtained from various farm animal species (cow, goat and sheep) Three different grana cheeses are analysed with GC-O and PTR-MS. For each type of grana some specific volatile compounds have been identified The gamma-casein profile, obtained with isoelectric focusing techniques, is presented as a good marker for the certification of Grana Padano cheese An HPLC method is able to quantify the percentage of bovine, ovine and caprine milk mixtures in PDO cheeses The proteolytic profile, obtained with Urea-PAGE and HPLC, if subjected to chemometric evaluation, could highlight some potential markers of PDO cheeses authenticity HRMS-NMR spectroscopy directly identifies specific metabolites of “Mozzarella di Bufala Campana” Two RT-PCR methods are able to evaluate the milk fractions of cow, sheep and goat in dairy products	(Woodcock, Fagan, O’Donnell, & Downey, 2008) (S.M.) (Cevoli, et al., 2013) (S.M.) (R) (Hrbek, Vaclavik, Elich, & Hajslova, 2014) (R) (Boscaini, van Ruth, Biasioli, Gasperi, & Märk, 2003) (S.M.)(C) (Restani, Velona, Carpen, Duranti, & Galli, 1996) (S.M.) (C) (Ferreira & Caçote, 2003) (S.M.)(C) (Guerreiro, Barros, Fernandes, Pires, & Bardsley, 2013) (S.M.) (C) (Mazzei & Piccolo, 2012) (C) (Rentsch , et al., 2013) (S.M.) (C)
Geographical origin	IRMS analyses Characterization of volatile compounds Immunoassay tests	The geographic origin of Emmental Cheese is determined with GC with MS or FID detectors and with a mass spectrometry based electronic nose. All data are treated with Principal Component Analysis The data obtained with MIR with front-face fluorescence spectroscopy are treated with multivariate techniques in order to discriminate through Emmental cheeses coming from different European countries ATR-FTIR spectroscopy with chemometric analysis is able to discriminate through the PDO “Parmigiano Reggiano” cheese and other grana-type cheeses coming from Italy and Europe The analysis of volatile compounds with SIFT-MS clearly separates New Zealand “parmesan” from Parmigiano Reggiano or Grana Padano cheeses Emmental cheeses are classified according to their country of origin thanks to a group of properties (i.e. nitrogen fractions, free amino acids, peptide profile etc.) The evaluation of some stable isotope ratios and of the mineral composition allows the discrimination of Emmental cheeses according to their geographical origin Cheese water-soluble metabolites, analysed with NMR spectroscopy, are studied with a chemometric approach in order to separate Parmigiano Reggiano samples from other “Grana” samples coming from east Europe countries IC, ICP-AES, NMR and IRMS, together with chemometric methods, are used for the geographical characterization of “mozzarella di bufala” cheeses Statistical models based on the Isotopic and mineral composition are able to predict the origin of different type of European cheeses and to discriminate between Parmigiano Reggiano cheese and imitators	(Pillonel, Ampuero, Tabacchi, & Bosset, 2003b) (S.M.) (R) (Karoui, et al., 2004) (S.M.) (R) (Gori, Maggio, Cerretani, Nocetti, & Caboni, 2012) (R) (Langford, et al., 2012) (C) (Pillonel, Albrecht, Badertscher, Buetikofer, & Bosset, 2003a) (S.M.) (C) (Pillonel, et al., 2003c) (S.M.) (C) (Consonni & Cagliani, 2008b) (C) (Brescia, Monfreda, Buccolieri, & Carrino, 2005) (S.M.)(C) (Camin, et al., 2012) (Camin, et al., 2004)(S.M.)(C)

Product adulteration	NIR spectroscopy	A MALDI-TOF-MS method detects some peptides that are markers of powdered milk addition to fresh cow's milk	(Calvano, Monopoli, Loizzo, Faccia, & Zambonin, 2013b) (S.M.)(C)
		An ATR-MIR-microspectroscopy method with chemometric elaboration is able to detect the addition of whey, hydrogen peroxide, urea and other compounds to milk samples	(Santos, Pereira-Filho, & Rodriguez-Saona, 2013) (S.M.) (R)
		A turbidimetric method identifies soy, pea, rice and wheat proteins addition in milk powder	(Scholl, Farris, & Mossoba, 2014) (S.M.) (R)
		The phospholipids profile, obtained with MALDI-TOF-MS, is evaluated in order to detect low value cow milk addition to high value milk (from sheeps or goats)	(Calvano, De Ceglie , Aresta, Facchini, & Zambonin, 2013a) (S.M.) (C)
		DART-HRMS technique coupled with chemometrics is able to detect the addition of vegetable oils to soft cheese	(Hrbek, Vaclavik, Elich, & Hajslova, 2014) (R)
		Butter adulteration with mutton fat is detected with FTIR-ATR spectroscopy coupled with multivariate analysis	(Fadzlillah, Rohman, Ismail, Mustafa, & Khatib, 2013) (S.M.) (R)
		The TAG composition in 2 PDO cheeses (Mahon and Machego), obtained with GC techniques and treated with multiple regression equations, allows the detection of foreign fats during ripening.	(Fontecha, Mayo, Toledano, & Juárez, 2006) (S.M.) (C)
		IRMS is able to trace low price milk or powdered milk additions along the entire Stelvio cheese production chain	(Capici, Mimmo, Kerschbaumer, Cesco, & Scampicchio, 2015) (C)
		Not declared hen's egg white lysozyme addition to cheeses is detected with HPLC-FLD methods	(Pellegrino, & Tirelli, 2000) (S.M.) (Schneider, Werkmeister, Becker, & Pischetsrieder, 2011) (S.M.) (C)
		Cow milk addition to "Mozzarella di bufala" can be detected with the evaluation of the cheese microbiological quality and with a gel electrophoresis characterization	(Buzi, Pinto, Ramos, & Biondi, 2009) (S.M.)(C)
		The bovine whey addition to water buffalo ricotta cheese is detected with an isoelectric focusing technique	(Fuselli, et al., 2015) (S.M.)(C)
		The phosphorylated beta-casein f33-48 peptide is presented as a novel species-specific marker for the detection of "Mozzarella di bufala" adulterations	(Russo, et al., 2012) (S.M.)(C)
		A ^{13}C-NMR method is able to detect the adulteration of butter fat with synthetic triacylglycerols	(Picariello, et al., 2013) (S.M.)(C)
		Butter adulterated with lard is detected with NMR spectroscopy and chemometric data analysis	(Fadzillah, et al., 2015) (S.M.) (C)
		Adulteration with cheaper fats and oils detected with Raman and PCA analysis	(Yazgan Karacaglar, Bulat, Boyaci, & Topcu, 2018) (R)
		Adulteration with anionic detergents determined by FT-IR spectroscopy combined with PCA and SIMCA models	(Jaiswal, Jha, Kaur, & Borah, 2017) (S.M.) (R)
Melamine addition	NIR spectroscopy	Melamine in milk and milk powder is detected with surface-enhanced Raman spectroscopy	(Ma, et al., 2013) (S.M.) (R)
		A HPLC-UV method with HILIC stationary phase is able to detect melamine in milk powder	(Ihunegbo, Tesfalidet, & Jiang, 2010) (S.M.) (C)
		Adulteration with melamine as a surrogate contaminant detected by screening tools based on portable ATR-FTIR combined with a SIMCA classification model	(Limm, Karunathilaka, Yakes, & Mossoba, 2018) (R)
		Melamine adulteration detected by NIR combined with one-class partial least squares (OCPLS) as chemometric tool	(Chen , Tan, & Wu, 2017) (S.M.)(R)
Water addition	Physical tests (i.e. cryoscopy, conductivity etc.)	Statistical evaluations confirm that some electrical parameters (i.e. admittance, impedance etc.) change according to the percentage addition of water to milk	(Banach, Żywica, Szpendowski, & Kiełczewska, 2012) (R)
		Digital images of samples are recorded and then different colour parameters (i.e. red, blue, saturation, intensity etc.) are evaluated with chemometrics in order to detect water or NaOH addition to cow's milk	(Santos, Wentzell, & Pereira-Filho, 2012) (S.M.)(R)
		A differential scanning calorimetry method is presented as able to detect water content in butter	(Tomaszewska-Gras, 2012) (S.M.) (R)

Table 10: Cheese-Milk Derivatives (S.M.: reference can be found in the supplementary materials)

Abbreviations

ANN= Artificial Neural Network
AOS= Artificial Olfactive System (“Electronic Nose”)
APCI = Atmospheric Pressure Chemical Ionization
ATR-FTIR= Attenuated Total Reflectance-Fourier Transform Infrared
ATR-MIR= Attenuated Total Reflectance Mid-Infrared
BPNN = Back Propagation Neural Network
BVM= Balsamic Vinegar of Modena
CB= Cocoa Butter
CBE= Cocoa Butter Equivalents
CZE= Capillary Zone Electrophoresis
CZE-UV= Capillary Zone Electrophoresis with indirect UV detection
DART–HRMS = Direct Analysis Real Time–High-Resolution Mass Spectrometry
DHPLC= Denaturing High Performance Liquid Chromatography
DRIFT= Diffuse Reflectance Infrared Fourier Transform Spectroscopy
ELISA= Enzyme-Linked Immunosorbent Assay
EN= Electronic Nose
ET= Electronic Tongue
FT-RAMAN= Fourier Transform Raman Spectroscopy
GC-O= Gas Chromatography-Olfactometry
GC-MS= Gas Chromatography-Mass Spectrometry
HFCS= High Fructose Corn Syrup
HILIC= Hydrophilic Interaction Liquid Chromatography
HPAEC-PAD= High Performance Anion-Exchange Chromatography with Pulsed Amperometric
Detection
HPLC-FLD= High Performance Liquid Chromatography with Fluorescence Detector
HPLC-RID= High Performance Liquid Chromatography with Refractive Index Detector
HPLC-PDA= High Performance Liquid Chromatography with Photodiode Array Detection
HRM = High Resolution Melting
HRMAS NMR= High Resolution Magic Angle Spinning Nuclear Magnetic Resonance
HS-SPME= Headspace Solid Phase Microextraction
IC= Ion Chromatography
ICA= Independent Components Analysis
ICP-AES= Inductively Coupled Plasma Atomic Emission Spectroscopy
ICP-MS= Inductively Coupled Plasma Mass Spectrometry
ICP-oa-TOF-MS= Inductively Coupled Plasma - Orthogonal Acceleration-Time of Flight Mass spectrometry
IMS= Ion Mobility Spectrometry
IRMS= Isotopic Ratio Mass Spectrometry
LDA= Linear Discriminant Analysis
LESA-MS= Liquid Extraction Surface Analysis Mass Spectrometry
LS-SVM = Least Squares-Support Vector Machine
MALDI= Matrix-Assisted Laser Desorption Ionization
MEKC-LIF= Micellar Electrokinetic Chromatography- Laser Induced Fluorescence
NIRS= Near Infrared Spectroscopy
OSC= Orthogonal Signal Correction
OCSIMCA = One-Class Soft Independent Modelling of Class Analogy
PCA= Principal Component Analysis

PCA-MLR = Principal component analysis-Multiple Linear Regression
PCR-RFLP= Polymerase Chain Reaction- Restriction Fragment Length
PDO = Protected Designation of Origin
PLS= Partial Least Squares
PRM = Parallel Reaction Monitoring
PTR-MS= Proton Transfer Reaction-Mass Spectrometry
RFLP-LOC-CE= Restriction Fragment Length-Lab on Chip-Capillary Gel Electrophoresis
RP-HPLC= Reverse Phase High Performance Liquid Chromatography
SCIRA-MS= Stable Carbon Isotopic Ratio Analysis by Mass Spectrometry
SIFT-MS= Selected Ion Flow Tube Mass Spectrometry
SIMCA= Soft Independent Modelling of Class Analogy
Si-PLS= Synergy interval Partial Least Square
SNIF-NMR= Site specific Natural Isotopic Fractionation- Nuclear Magnetic Resonance
SNIF-NMR= Site-specific Natural Isotopic Fractionation- Nuclear Magnetic Resonance
SNP= Single Nucleotide Polymorphism
SP-LDI-MS= Silica Plate Laser Desorption- Ionization Mass Spectrometry
SPME= Solid Phase Microextraction
SSR= Simple Sequence Repeats
SVM= Support Vector Machine
TAG= Triacylglycerol
TBVM= Traditional Balsamic vinegar of Modena
TIMS= Thermal Ionization Mass Spectrometry
TOF= Time of Flight
UPLC-ELSD= Ultra Performance Liquid Chromatography-Evaporative Light Scattering Detector
UPLC-HRMS= Ultra high Performance Liquid Chromatography – High Resolution Mass Spectrometry

Bibliographic references

Abad-García, B., Garmón-Lobato, S., Sánchez-Ilárduya, M., Berrueta, L., Gallo, B., Vicente, F., & Alonso-Salces, R. (2014). Polyphenolic contents in Citrus fruit juices: authenticity assessment. European Food Research and Technology, 238(5), 803-818.

Abdanan, M., Minaei, S., Hancock, N., & Karimi, T. (2014). An intelligent system for egg quality classification based on visible-infrared transmittance spectroscopy. *Information Processing in Agriculture, 1*(2), 105-114.

Ammari, F., Redjdal, L., & Rutledge, D. (2015). Detection of orange juice frauds using front-face fluorescence spectroscopy and Independent Components Analysis. *Food Chemistry, 168*, 211-217.

Arana, V., Medina, J., Alacorn, R., Moreno, E., Heintz, L., Schafer H., & Wist, J. (2015). Coffee's country of origin determined by NMR: the Colombian case. *Food Chemistry, 175*, 500-506.

Avula, B., Smillie, T., Wang, Y.-H., Zweigenbaum, J., & Khan, I. (2015). Authentication of true cinnamon (Cinnamon verum) utilising direct analysis in real time (DART)-QToF-MS. *Food Additives & Contaminants: Part A, 32*(1), 1-8.

Ballin, N. (2010). Authentication of meat and meat products. *Meat Science, 86*(3), 577-587.

Banach, J., Żywica, R., Szpendowski, J., & Kiełczewska, K. (2012). Possibilities of Using Electrical Parameters of Milk for Assessing its Adulteration with Water. *International Journal of Food Properties, 15*(1-2), 274-280.

Banach, U., Tiebe, C., & Hubert, T. (2012). Multigas sensors for the quality control of spice mixtures. *Food control, 26*(1), 23-27.

Black, C., Haughey, S., Chevalier, O., Galvin-King, P., & Elliott, C. (2016). A comprehensive strategy to detect the fraudulent adulteration of herbs: The oregano approach. *Food Chemistry , 210*, 551-557.

Bonatto, C., & Silva, L. (2015). Cocoa content influences chocolate molecular profile investigated by MALDI-TOF mass spectrometry. *Journal of the Science of Food and Agriculture, 95*(8), 1753-1756.

Bontempo, L., Camin, F., Paolini, M., Micheloni, C., & Laursen, K. (2016). Multi-isotopic signatures of organic and conventional italian pasta along the production chain. *Journal of Mass Spectrometry , 51*, 675-683.

Bontempo, L., Ceppa, F., Perini, M., Tonon, A., Gagliano, G., Marianella, R., . . . Camin, F. (2014). Use of δ18O authenticity thresholds to differentiate tomato passata from diluted tomato paste. *Food Control, 35*(1), 413-418.

Borges, E., Volmer, D., Brandelero, E., Gelinski, J., Gallimberti, M., & Barbosa Jr, F. (2016). Monitoring the Authenticity of Organic Grape Juice via Chemometric Analysis of Elemental Data. *Food Analytical Methods, 9*(2), 362-369.

Borras, E., Ferrè, J., Boqué, R., Mestres, M., Acena, L., & Busto, O. (2015). Data fusion methodologies for food and beverages authentication and quality assessment- A review. *Analytica Chimica Acta , 891*, 1-14.

Brondi, A., Torres, C., Garcia, J., Jerusa, S., & Trevisan, M. (2017). Differential Scanning Calorimetry and Infrared Spectroscopy Combined with Chemometric Analysis to the Determination of Coffee Adulteration by Corn. *J. Braz. Chem. Soc., 28*(7), 1308-1314.

Cagliani, L., Culeddu, N., Chessa, M., & Consonni, R. (2015). NMR investigations for a quality assessment of Italian PDO saffron (Crocus sativus L.). *Food Control , 50*, 342-348.

Cai, T., Ting, H., & Jin-Ian, Z. (2016). Novel identification strategy for ground coffee adulteration based on UPLC-HRMS oligosaccharide profiling. *Food Chemistry , 190*, 1046-1049.

Cajka, T., Danhelova, H., Zachariasova, M., Riddellova, K., & Hajslova, J. (2013). Application of direct analysis in real time ionization–mass spectrometry (DART–MS) in chicken meat metabolomics aiming at the retrospective control of feed fraud. *Metabolomics, 9*(3), 545-557.

Cajka, T., Hajslova, J., Pudil, F., & Riddellova, K. (2009). Traceability of honey origin based on volatiles pattern processing by artificial neural networks. *Journal of Chromatography A, 1216*(9), 1458-1462.

Camin, F., Bontempo, L., Perini, M., Tonon, A., Breas, O., Guillou, C., . . . Gagliano, G. (2013). Control of wine vinegar authenticity through δ18O analysis. *Food Control, 29*(1), 107-111.

Camin, F., Wehrens, R., Bertoldi , D., Bontempo, L., Ziller , L., Perini , M., . . . Larcher, R. (2012). H, C, N and S stable isotopes and mineral profiles to objectively guarantee the authenticity of grated hard cheeses. *Analytica Chimica Acta, 711*, 54-59.

Capici, C., Mimmo, T., Kerschbaumer, L., Cesco, S., & Scampicchio, M. (2015). Determination of Cheese Authenticity by Carbon and Nitrogen Isotope Analysis: Stelvio Cheese as a Case Study. *Food Analytical Methods, 8*(8), 2157-2162.

Casale, M., Bagnasco, L., Zotti, M., Di Piazza, S., Sitta, N., & Oliveri, P. (2016). A NIR spectroscopy-based efficient approach to detect fraudulent additions within mixtures of dried porcini mushrooms. *Talanta, 160*, 729-734.

Casazza, A., Morcia, C., Ponzoni, E., Gavazzi, F., Benedettelli, S., & Breviario, D. (2012). A reliable assay for the detection of soft wheat adulteration in Italian pasta is based on the use of new DNA molecular markers

capable of discriminating between Triticum aestivum and Triticum durum. *Journal of Cereal Science, 56*(3), 733-740.
Cavanna, D., Zanardi, S., Dall'Asta, C., & Suman, M. (2019). Ion mobility spectrometry coupled to gas chromatography: A rapid tool to assess eggs freshness. *Food Chemistry, 271*, 691-696.
Chen, L., Xue, X., Ye, Z., Zhou, J., Chen, F., & Zhao, J. (2011). Determination of Chinese honey adulterated with high fructose corn syrup by near infrared spectroscopy. *Food Chemistry, 128*(4), 1110-1114.
Cheng, Y., & Dong, Y. (2011). Screening melamine contaminant in eggs with portable surface-enhanced Raman Spectroscopy on gold nanosubstrate. *Food Control, 22*(5), 685-689.
Çinar, S., Ekşi, A., & Coşkun, I. (2014). Carbon isotope ratio (13C/12C) of pine honey and detection of HFCS adulteration. *Food Chemistry, 157*, 10-13.
Consonni, R., & Cagliani, L. (2008b). Ripening and geographical characterization of Parmigiano Reggiano cheese by 1H NMR spectroscopy. *Talanta, 76*(1), 200-205.
Dankowska, A. (2017). Data fusion of fluorescence and UV spectroscopies improves the detection of cocoa butter adulteration. *European Journal of Lipid Science and Technology, 119*(8).
Dankowska, A., Domagała, A., & Kowalewski , W. (2017). Quantification of Coffea arabica and Coffea canephora var. robusta concentration in blends by means of synchronous fluorescence and UV-Vis spectroscopies. *Talanta, 172*, 215-220.
Das, C., Chakraborty, S., Acharya, K., Bera, N. K., Chattopadhyay, D., Karmakar, A., & Chattopadhyay, S. (2017). FT-MIR supported Electrical Impedance Spectroscopy based study of sugar adulterated honeys from different floral origin. *Talanta, 171*, 327-334.
de Castro, L., Maihara, V., Silva, P., & Figueira, R. (2012). Artificial and natural radioactivity in edible mushrooms from Sao Paulo, Brazil. *Journal of Environmental Radioactivity, 113*, 150-154.
Di Anibal, C., Odena, M., Ruisánchez, I., & Callao, M. (2009). Determining the adulteration of spices with Sudan I-II-III-IV dyes by UV-Visible spectroscopy and multivariate classification techniques. *Talanta, 79*(3), 887-892.
Diomande, D., Antheaume, I., Leroux, M., Lalande, J., Balayssac, S., Remaud, G., & Tea, I. (2015). Multi-element, multi-compound isotope profiling as a means to distinguish the geographical and varietal origin of fermented cocoa (Theobroma cacao L.) beans. *Food Chemistry, 188*, 576-582.
Domingues, D., Pauli, E., de Abreu, J., Massura, F., Cristiano, V., Santos, M., & Nixdorf, S. (2014). Detection of roasted and ground coffee adulteration by HPLC by amperometric and post-column derivatization UV-Vis detection. *Food Chemistry , 146*, 353-362.
Draisci, R., Palleschi, L., Ferretti, E., Lucentini, L., & Cammarata, P. (2000). Quantitation of anabolic hormones and their metabolites in bovine serum and urine by liquid chromatography-tandem mass spectrometry. *Journal of Chromatography A, 870*(1-2), 511-522.
Ebrahimi-Najafabadi, H., Leardi, R., Oliveri, P., Caolino, C., Jalali-Heravi, M., & Lanteri, S. (2012). Detection of addition of barley to coffee using near infrared spectroscopy and chemometric techniques. *Talanta , 99*, 175-179.
Ernest, T., Xing-yi, H., Wu, L., & Huang, D. (2014). Feasibility study on the use of Fourier transform near-infrared spectroscopy together with chemometrics to discriminate and quantify adulteration in cocoa beans. *Food Research International, 55*, 288-293.
Figueira , J., Câmara, H., Pereira, J., & Câmara, J. (2014). Evaluation of volatile metabolites as markers in Lycopersicon esculentum L. cultivars discrimination by multivariate analysis of headspace solid phase microextraction and mass spectrometry data. *Food Chemistry, 145*, 653-663.
Frank, N., Bessaire, T., Tasser, A., Goyion, A., & Delatour, T. (2017). Development of a quantitative multi-compound method for the detection of 14 nitrogen-rich adulterants by LC-MS/MS in food materials. *Food Additives & Contaminants: Part A, 34*(11), 1842-1852.
Garber, E., & Brewer, V. (2010). Enzyme-linked immunosorbent assay detection of melamine in infant formula and wheat food products. *Journal of Food Protection, 73*(4), 701-707.
Giancaspro, A., Colasuonno, P., Zito, D., Blanco, A., Pasqualone, A., & Gadaleta, A. (2016). Varietal traceability of bread 'Pane Nero di Castelvetrano' by denaturing high pressure liquid chromatography analysis of single nucleotide polymorphisms. *Food Control, 59*, 809-817.
Gori, A., Maggio, R., Cerretani, L., Nocetti, M., & Caboni, M. (2012). Discrimination of grated cheeses by Fourier transform infrared spectroscopy coupled with chemometric techniques. *International Dairy Journal, 23*(2), 115-120.
Grundy, H., Reece, P., Sykes, M., Clough, J., Audsley, N., & Stones, R. (2007). Screening method for the addition of bovine blood-based binding agents to food using liquid chromatography triple quadrupole mass spectrometry. *Rapid Communication in Mass Spectrometry, 21*(18), 2919-2925.
Guerreiro, T., de Oliveira , D., Ferreira, M., & Catharino, R. (2014). High-throughput analysis by SP-LDI-MS for fast identification of adulterations in commercial balsamic vinegars. *Analytica Chimica Acta, 838*, 86-92.

Ha, Y., Thienes, C., Agapov, A., Laznicka, A., Han, S., Nadala, C., & Samadpour, M. (2018). Comparison of ELISA and DNA Lateral Flow Assays for Detection of Pork, Horse, Beef, Chicken, Turkey, and Goat Contamination in Meat Products. *Journal of AOAC International*.

Haughey, S., Galvin-King, P., Ho, Y.-C., Bell, S., & Elliott, C. (2015). The feasibility of using near infrared and Raman spectroscopic techniques to detect fraudulent adulteration of chili powders with Sudan dye. *Food Control, 48*, 18-23.

Hohmann, M., Christoph, N., Wachter, H., & Holzgrabe, U. (2014). 1H NMR Profiling as an Approach To Differentiate Conventionally and Organically Grown Tomatoes. *Journal of Agricultural and Food Chemistry, 62*(33), 8530-8540.

Hong, X., Wang, J., & Qiu, S. (2014a). Authenticating cherry tomato juices—Discussion of different data standardization and fusion approaches based on electronic nose and tongue. *Food Research International, 60*, 173-179.

Hori, K., Kiriyama, T., & Tsumura, K. (2016). A Liquid Chromatography Time-of-Flight Mass Spectrometry-Based Metabolomics Approach for the Discrimination of Cocoa Beans from Different Growing Regions. *Food Analytical Methods, 9*(3), 738-743.

Horn, B., Esslinger, S., Pfister, M., Fauhl-Hassek, C., & Riedl, J. (2018). Non-targeted detection of paprika adulteration using mid-infrared spectroscopy and one-class classification – Is it data preprocessing that makes the performance? *Food Chemistry, 257*, 112-119.

Hrbek, V., Vaclavik, L., Elich, O., & Hajslova, J. (2014). Authentication of milk and milk-based foods by direct analysis in real time ionization–high resolution mass spectrometry (DART–HRMS) technique: A critical assessment. *Food Control, 36*(1), 138-145.

Jandrić, Z., Roberts, D., Rathor, M., Abrahim, A., Islam, M., & Cannavan, A. (2014). Assessment of fruit juice authenticity using UPLC-QToF MS: a metabolomics approach. *Food Chemistry, 148*, 7-17.

Langford, V., Reed, C., Milligan, D., McEwan, M., Barringer, S., & Harper, W. (2012). Headspace analysis of Italian and New Zealand parmesan cheeses. *Journal of Food Science, 77*(6), C719-C726.

Limm, W., Karunathilaka, S. R., Yakes, B. J., & Mossoba, M. M. (2018). A portable mid-infrared spectrometer and a non-targeted chemometric approach for the rapid screening of economically motivated adulteration of milk powder. *International Dairy Journal, 85*, 177-183.

Lin, H., Zhao, J., Sun, L., Chen, Q., & Zhou, F. (2011). Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis. *Innovative Food Science & Emerging Technologies , 12*(2), 182-186.

Liu, C., Hao, G., Su, M., Chen, Y., & Zheng, L. (2017). Potential of multispectral imaging combined with chemometric methods for rapid detection of sucrose adulteration in tomato paste. *Journal of Food Engineering, 215*, 78-83.

Liu, C., Zhao, L., Sun, Z., Cheng, N., Xue, X., Wu, L., & Cao, W. (2018). Determination of three flavor enhancers using HPLC-ECD and its application in detecting adulteration of honey. *Analytical Methods, 10*(7), 743-748.

Liu, R., Hei, W., He, P., & Li, Z. (2011). Simultaneous determination of fifteen illegal dyes in animal feeds and puoltry products by ultra-high performance liquid chromatography tandem mass spectrometry. *Journal of Chromatography B, 879*(24), 2416-2422.

Loaiza, P., Brudanin, V., Piquemal, F., Reyss, J., Stekl, I., Warot, G., & Zampaolo, M. (2012). Air radioactivity levels following the Fukushima reactor accident measured at the Laboratoire Souterrain de Modane, France. *Journal of Environmental Radioactivity, 114*, 66-70.

Lopez-Calleja, I., Cruz, S., Pegels, N., Gonzalez, I., Garcia, T., & Martin, R. (2013). High resolution TaqMan real-time PCR approach to detect hazelnut DNA encoding for ITS rDNA in foods. *Food Chemistry, 141*(3), 1872-1880.

Mayer, F., Haase, I., Graubner, A., Heising, F., Paschke-Kratzin, A., & Fischer, M. (2012). Use of polymorphisms in the γ-gliadin gene of spelt and wheat as a tool for authenticity control. *Journal of Agricultural and Food Chemistry, 60*(6), 1350-1357.

Mazzei, P., & Piccolo, A. (2012). 1 H HRMAS-NMR metabolomic to assess quality and traceability of mozzarella cheese from Campania buffalo milk. *Food Chemistry, 132*(3), 1620-1627.

Medina, J., Bernal, A., Wist, J., Carlo Rodriguez, D., Esseiva, P., & Arona, V. (2017). Comparison of Attenuated Total Reflectance Mid-Infrared, Near Infrared, and 1H-Nuclear Magnetic Resonance Spectroscopies for the Determination of Coffee's Geographical Origin. *International journal of analytical chemistry*.

Miano, B., Righetti, L., Piro, R., Dall'Asta, C., Folloni, S., Galaverna, G., & Suman, M. (2018). Direct analysis real-time–high-resolution mass spectrometry for Triticum species authentication. *Food Additives & Contaminants: Part A*.

Nakashita, R., Suzuki, Y., Akamatsu, F., Iizumi, Y., Korenaga, T., & Chikaraishi, Y. (2008). Stable carbon, nitrogen, and oxygen isotope analysis as a potential tool for verifying geographical origin of beef. *Analytica Chimica Acta, 617*(1-2), 148-152.

Pardo, M. (2015). Evaluation of a dual-probe real time PCR system for detection of mandarin in commercial orange juice. *Food Chemistry, 172*, 377-384.

Parvathya, V., Swethaa, V., Sheejaa, T., Leelab, N., Chempakamb, B., & Sasikumara, B. (2014). DNA Barcoding to Detect Chilli Adulteration in Traded Black Pepper Powder. *Food Biotechnology, 28*(1), 25-40.

Petrakis, E., & Polissiou, M. (2017). Assessing saffron (Crocus sativus L.) adulteration with plant-derived adulterants by diffuse reflectance infrared Fourier transform spectroscopy coupled with chemometrics. *Talanta, 162*, 558-566.

Prandi, B., Varani, M., Faccini, A., Lambertini, F., Suman, M., Leporati, A., Tedeschi, T., & Sforza, S. (2018). Species specific marker peptides for meat authenticity assessment: a multispecies quantitative approach applied to Bolognese sauce. *Food Control*.

Quelal-Vasconez, M., Perez-Esteve, E., Arnauld-Bonachera, A., Barat, J., & Taulens, P. (2018). Rapid fraud detection of cocoa powder with carob flour using near infrared spectroscopy. *Food Control, 92*, 183-189.

Reg. CE 853/2004. (s.d.).

Ribeiro, M., Boralle, N., Redigolo Pezza, H., Pezza, L., & Toci, A. (2017). Authenticity of roasted coffee using1H NMR spectroscopy. *Journal of Food Composition and Analysis, 57*, 24-30.

Ribeiro, R., Mársico, E., Carneiro, C. d., Monteiro, M., Júnior, C., & de Jesus, E. (2014). Detection of honey adulteration of high fructose corn syrup by Low Field Nuclear Magnetic Resonance (LF 1H NMR). *Journal of Food Engineering, 135*, 39-43.

Ropodi, A., Pavlidis, D., Mohareb, F., Panagou, E., & Nychas, G.-J. (2015). Multispectral image analysis approach to detect adulteration of beef and pork in raw meats. *Food Research International, 67*, 12-18.

Sak-Bosnar, M., & Sakac, N. (2012). Direct potentiometric determination of diastase activity in honey. *Food Chem, 135*(2), 827-831.

Sak-Bosnar, M., & Sakač, N. (2012). Direct potentiometric determination of diastase activity in honey. *Food Chemistry, 135*(2), 827-831.

Scarano, D., Rao, R., Masi, P., & Corrado, G. (2015). SSR fingerprint reveals mislabeling in commercial processed tomato products. Food Control, 51, 397-401.

Shrestha, S., Deleuran, L., Olesen, M., & Gislum, R. (2015). Use of Multispectral Imaging in Varietal Identification of Tomato. *Sensors (Switzerland), 15*(2), 4496-4512.

Silletti, S., Morello, L., Gavazzi, F., Giani, S., Braglia, L., & Breviario, D. (2018). Untargeted DNA-based methods for the authentication of wheat species and related cereals in food products. *Food Chemistry, 271*, 410-418.

Škorpilova, T., Šimoniova, A., Rohlik, B.-A., & Pipek, P. (2014). Differentiation between Fresh and Thawed Chicken Meat by the Measurement of Aconitase Activity. *Czech Journal of Food Sciences, 32*(5), 509-513.

Snyder, A., Sweeney, C., Rodriguez-Saona, L., & Giusti, M. (2014). Rapid authentication of concord juice concentration in a grape juice blend using Fourier-Transform infrared spectroscopy and chemometric analysis. *Food Chemistry, 147*, 295-301.

Sobrino-Gregorio, L., Bataller, R., Soto, J., & Escriche, I. (2018). Monitoring honey adulteration with sugar syrups using an automatic pulse voltammetric electronic tongue. *Food Control, 91*, 254-260.

Souto, U., Barbosa, M., Dantas, H., de Pontes, A., Lyra, W., Diniz, P., . . . da Silva, E. (2015). Identification of adulteration in ground roasted coffees using UV-Vis spectroscopy and SPA-LDA. *LWT Food Scienc and Technology , 63*, 1037-1041.

Spiteri, M., Jamin, E., Thomas, F., Rebours, A., Lees, M., Rogers, K., & Rutledge, D. (2015). Fast and global authenticity screening of honey using 1H-NMR profiling. *Food Chemistry, 189*, 60-66.

Suman, M., Riani, G., & Dalcanale, E. (2007). MOS-based artifical olfactory system for the assessment of egg products freshness. *Sensors and Actuators B: Chemical, 125*(1), 40-47.

Trincherini, P., Baffi, C., Barbero, P., Pizzoglio, E., & Spalla, S. (2014). Precise determination of strontium isotope ratios by TIMS to authenticate tomato geographical origin. *Food Chemistry, 145*, 349-355.

van Ruth, S., Alewijn, M., Rogers, K., Newton-Smith, E., Tena, N., Bollen, M., & Koot, A. (2011). Authentication of organic and conventional eggs by carotenoid profiling. *Food Chemistry, 126*(3), 1299-1305.

Vermeulen, P., Suman, M., Fernández Pierna, J. A., & Baeten, V. (2018). Discrimination between durum and common wheat kernels using near infrared hyperspectral imaging. *84*, 74-82.

Voldřich, M., Rajchl, A., Čížková, H., & Cuhra, P. (2009). Detection of Foreign Enzyme Addition into the Adulterated Honey. *Czech Journal of Food Sciences, 27*, S280-S282.

Wang, W., Zhu, Y., Chen, Y., Xu, X., & Zhou, G. (2015). Rapid visual detection of eight meat species using optical thin-film biosensor chips. *Journal of AOAC International, 98*(2), 410-414.

Werner, R., & Roßmann, A. (2015). Multi element (C, H, O) stable isotope analysis for the authentication of balsamic vinegars. *Isotopes in Environmental and Health Studies, 51*(1), 58-67.

Willems, J., & Low, N. (2014). Authenticity analysis of pear juice employing chromatographic fingerprinting. *Journal of Agricultuiral and Food Chemistry, 62*(48), 11737-11747.

Yang, W.-L., Hu, M.-H., Chen, S.-W., Wang, Q., Zhu, S., Dai, J., & Li, X.-Z. (2015). Identification of Adulterated Cocoa Powder Using Chromatographic Fingerprints of Polysaccharides Coupled with Principal Component Analysis. *Food Analytical Methods, 8*(9), 2360-2367.

Yazgan Karacaglar, N. N., Bulat, T., Boyaci, I. H., & Topcu, A. (2018). Raman spectroscopy coupled with chemometric methods for the discrimination of foreign fats and oils in cream and yogurt. *Journal of Food and Drug Analysis*.

Yimenu, S., Kim, J., & Kim, B. (2017). Prediction of egg freshness during storage using electronic nose. *Poultry Science, 96*(10), 3733-3746. doi:10.3382/ps/pex193

Zanardi, E., Caligiani, A., Palla, L., Mariani, M., Ghidini, S., Di Ciccio, P., . . . Ianieri, A. (2015). Metabolic profiling by (1)H NMR of ground beef irradiated at different irradiation doses. *Meat Science, 103*, 83-89.

Zeng, J., Cao , X., Liu, Y., Chen, J., & Ren, K. (2015). A single cataluminescence sensor based on spectral array and its use in the identification of vinegars. *Analytica Chimica Acta, 864*, 64-73.

Zhang, W., Pan, L., Tu, S., Zhan, G., & Tu, K. (2015). Non-destructive internal quality assessment of eggs using a synthesis of hyperspectral imaging and multivariate analysis. *Journal of Food Engineering, 157*, 41-48.

Supplementary Material

Additional references

Alamprese, C., Casale, M., Sinelli, N., Lanteri, S., & Casiraghi, E. (2013). Detection of minced beef adulteration with turkey meat by UV–vis, NIR and MIR spectroscopy. *LWT - Food Science and Technology, 53*(1), 225-232.

Aldeguer, M., López-Andreo, M., Gabaldón, J., & Puyet, A. (2014). Detection of mandarin in orange juice by single-nucleotide polymorphism qPCR assay. *Food Chemistry, 145*, 1086-1091.

Ali, M., Hashim, U., Mustafa, S., Che Man, Y., Dhahi, T., Kashif, M., . . . Abd Hamid, S. (2012). Analysis of pork adulteration in commercial meatballs targeting porcine-specific mitochondrial cytochrome b gene by TaqMan probe real-time polymerase chain reaction. *Meat Science, 91*(4), 454-459.

Al-Jowder, O., Kemsley, E., & Wilson, R. (1997). Mid-infrared spectroscopy and authenticity problems in selected meats: a feasibility study. *Food Chemistry, 59*(2), 195-201.

Aquino, F., Augusti, R., Alves, J., Diniz, M., Morais, S., Alves, B., . . . Sabino, A. (2014). Direct infusion electrospray ionization mass spectrometry applied to the detection of forgeries: Roasted coffees adulterated with their husks. *Microchemical Journal, 117*, 127-132.

Ayaz, Y., Ayaz, N., & Erol, I. (2006). Detection of Species in Meat and Meat Products Using Enzyme-linked Immunosorbent Assay. *Journal of Muscle Foods, 17*(2), 214-220.

Barbin, D., Sun, D.-W., & Su, C. (2013). NIR hyperspectral imaging as non-destructive evaluation tool for the recognition of fresh and frozen–thawed porcine longissimus dorsi muscles. *Innovative Food Science & Emerging Technologies, 18*, 226-236.

Barnwell, P., McCarthy, P., Lumley, I., & Griffin, M. (1994). The Use of Reversed-phase High-performance Liquid Chromatography to Detect Common Wheat (Triticum aestivum) Adulteration of Durum Wheat (Triticum durum) Pasta Products Dried at Low and High Temperatures. *Journal of Cereal Science, 20*(3), 245-252.

Boggia, R., Casolino, M., Hysenaj, V., Oliveri, P., & Zunin, P. (2013). A screening method based on UV-Visible spectroscopy and multivariate analysis to assess addition of filler juices and water to pomegranate juices. *Food Chemistry, 140*(4), 735-741.

Bontempo, L., Camin, F., Manzocco, L., Nicolini , G., Wehrens, R., Ziller, L., & Larcher, R. (2011). Traceability along the production chain of Italian tomato products on the basis of stable isotopes and mineral composition. *Rapid Comunications in Mass Spectrometry, 25*(7), 899-909.

Boscaini, E., van Ruth, S., Biasioli, F., Gasperi, F., & Märk, T. (2003). Gas chromatography-olfactometry (GC-O) and proton transfer reaction-mass spectrometry (PTR-MS) analysis of the flavor profile of grana padano, parmigiano reggiano, and grana trentino cheeses. *Journal of Agricultural and Food Chemistry, 51*(7), 1782-1790.

Brescia, M., Monfreda, M., Buccolieri, A., & Carrino, C. (2005). Characterisation of the geographical origin of buffalo milk and mozzarella cheese by means of analytical and spectroscopic determinations. *Food Chemistry, 89*(1), 139-147.

Bryan, G., Dixon, A., Gale, M., & Wiseman, G. (1998). A PCR-based Method for the Detection of Hexaploid Bread Wheat Adulteration of Durum Wheat and Pasta. *Journal of Cereal Science, 28*(2), 135-145.

Buchgraber, M., Androni, S., & Anklam, E. (2007). Determination of cocoa butter equivalents in milk chocolate by triacylglycerol profiling. *Journal of Agricultural and Food Chemistry, 55*(9), 3284-3291.

Buchgraber, M., Senaldi, C., Ulberth, F., & Anklam, E. (2004). Detection and quantification of cocoa butter equivalents in cocoa butter and plain chocolate by gas liquid chromatography of triacylglycerols. *Journal of AOAC International, 87*(5), 1153-1163.

Buzi, K., Pinto, J., Ramos, P., & Biondi, G. (2009). Microbiological analysis and electrophoretic characterization of mozzarella cheese made from buffalo milk. *Ciecia E Tecnologia De Alimentos, 29*(1), 7-11.

Caligiani, A., Cirlini, M., Palla, G., Ravaglia, R., & Arlorio, M. (2007). GC-MS detection of chiral markers in cocoa beans of different quality and geographic origin. *Chirality, 19*(4), 329-324.

Caligiani, A., Palla, L., Acquotti, D., Marseglia, A., & Palla, G. (2014). Application of 1H NMR for the characterisation of cocoa beans of different geographical origins and fermentation levels. *Food Chemistry, 157*, 94-99.

Calvano, C., De Ceglie , C., Aresta, A., Facchini, L., & Zambonin, C. (2013a). MALDI-TOF mass spectrometric determination of intact phospholipids as markers of illegal bovine milk adulteration of high-quality milk. *Analytical and Bioanalytical Chemistry, 405*(5), 1641-1649.

Calvano, C., Monopoli, A., Loizzo, P., Faccia, M., & Zambonin, C. (2013b). Proteomic approach based on MALDI-TOF MS to detect powdered milk in fresh cow's milk. *Journal of Agricultural and Food Chemistry, 61*(8), 1609-1617.

Camin, F., Wietzerbin, K., Blanch Cortes, A., Haberhauer, G., Lees, M., & Versini, G. (2004). Application of Multielement Stable Isotope Ratio Analysis to the Characterization of French, Italian, and Spanish Cheeses. *Journal of Agricultural and Food Chemistry, 52*(21), 6592-6601.

Cevoli, C., Gori, A., Nocetti, M., Cuibus, L., Caboni, M., & Fabbri, A. (2013). FT-NIR and FT-MIR spectroscopy to discriminate competitors, non compliance and compliance grated Parmigiano Reggiano cheese. *Food Research International, 52*(1), 214-220.

Che Man, Y., Syahariza, Z., Mirghani, M., Jinap, S., & Bakar, J. (2005). Analysis of potential lard adulteration in chocolate and chocolate products using Fourier transform infrared spectroscopy. *Food Chemistry, 90*(4), 815-819.

Chen , H., Tan, C., & Wu, T. (2017). Detection of melamine adulteration in milk by near-infrared spectroscopy and one-class partial least squares. *Spectrochim Acta A Mol. Biomol. Spectrosc., 173*, 832-836.

Chen, S., Fan, L., Song, J., Liu, C., & Zhang, H. (2013). Identification of irradiated meats by determining o- and m-tyrosine as markers. *Meat Science, 93*(2), 226-232.

Chua, L., Abdul-Rahaman, N.-L., Sarmidi, M., & Aziz, R. (2012). Multi-elemental composition and physical properties of honey samples from Malaysia. *Food Chemistry, 135*(3), 880-887.

Cocchi, M., Durante, C., Foca, G., Marchetti, A., Tassi, L., & Ulrici, A. (2006). Durum wheat adulteration detection by NIR spectroscopy multivariate calibration. *Talanta, 68*(5), 1505-1511.

Consonni, R., Cagliani, L., & Cogliati, C. (2012). NMR based geographical characterization of roasted coffee. *Talanta , 88*, 420-426.

Consonni, R., Cagliani, L., & Cogliati, C. (2013). Geographical discrimination of honeys by saccharides analysis. *Food Control, 32*(2), 543-548.

Consonni, R., Cagliani, L., Rinaldini, S., & Incerti, A. (2008a). Analytical method for authentication of Traditional Balsamic Vinegar of Modena. *Talanta, 75*(3), 765-769.

Cordella, C., Militão, J., Clément, M.-C., Drajnudel, P., & Cabrol-Bass, D. (2005). Detection and quantification of honey adulteration via direct incorporation of sugar syrups or bee-feeding: preliminary study using high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) and chemometrics. *Analytica Chimica Acta, 531*(2), 239-248.

Cotte, J., Casabianca, H., Chardon, S., Lheritier, J., & Grenier-Loustalot, M. (2003). Application of carbohydrate analysis to verify honey authenticity. *Journal of Chromatography A, 1021*(1-2), 145-155.

Cotte, J., Casabianca, H., Lhéritier, J., Perrucchietti, C., Sanglar, C., Waton, H., & Grenier-Loustalot, M. (2007). Study and validity of 13C stable carbon isotopic ratio analysis by mass spectrometry and 2H site-specific natural isotopic fractionation by nuclear magnetic resonance isotopic measurements to characterize and control the authenticity of honey. *Analytica Chimica Acta, 582*(1), 125-136.

Croft, L. (1987). Stable isotope mass spectrometry in honey analysis. *TrAC Trends in Analytical Chemistry, 6*(8), 206-209.

Cubero-Leon, E., De Rudder, O., & Maquet, A. (2018). Metabolomics for organic food authentication: Results from a long-term field study in carrots. *Food Chemistry, 239*, 760-770.

D'Archivio, A., Giannitto, A., Incani, A., & Nisi, S. (2014). Analysis of the mineral composition of Italian saffron by ICP-MS and classification of geographical origin. *Food Chemistry, 157*, 485-489.

de Alda-Garcilope, C., Gallego-Picó , A., Bravo-Yagüe , J., Garcinuño-Martínez , R., & Fernández-Hernando , P. (2012). Characterization of Spanish honeys with protected designation of origin "Miel de Granada" according to their mineral content. *Food Chemistry, 135*(3), 1785-1788.

Deng, X.-J., Guo, D.-H., Zhao, S.-Z., Han, L., Sheng, Y.-G., Yi, X.-H., Zhou,Y., Peng, T. (2010). A novel mixed-mode solid phase extraction for simultaneous determination of melamine and cyanuric acid in food by hydrophilic interaction chromatography coupled to tandem mass chromatography. *Journal of Chromatography B, 878*(28), 2839-2844.

Di Anibal, C., Rodriguez, M., & Albertengo, L. (2015). Synchronous fluorescence and multivariate classification analysis as a screening tool for determining Sudan I dye in culinary spices. *Food Control, 56*, 18-23.

Dionisi, F., Golay, P., Hug, B., Baumgartner, M., Callier, P., & Destaillats, F. (2004). Triacylglycerol analysis for the quantification of cocoa butter equivalents (CBE) in chocolate: feasibility study and validation. *Journal of Agricultural and Food Chemistry, 52*(7), 1835-1841.

D.P.R. n 187, 9 February 2001-Regolamento per la revisione della normativa sulla produzione e commercializzazione di sfarinati e paste alimentari, a norma dell'articolo 50 della legge 22 febbraio 1994, n. 146. (2001). *Official Italian Journal, 117*, p. 6-12.

EC. (2001). Council directive 2001/110/EC of 20 December 2001 relating honey. Official Journal of the European Communities 12.1.2002 L10/47-52; Codex. (2001). Codex Alimentarius standard for honey 12-1981. Revised Codex standard for honey. Standards and standard methods (Vol. 11). Retrieved December, 2014, from http:// www.codexalimentarius.net]

Ehling, S., Tefera, S., & Ho, I. (2007). High-performance liquid chromatographic method for the simultaneous detection of the adulteration of cereal flours with melamine and related triazine by-products ammeline, ammelide, and cyanuric acid. *Food Additives and Contaminants, 24*(12), 1319-1325.

Elliott, C. (2019). *Elliott Review into the Integrity and Assurance of Food Supply Networks – Final Report*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/350726/elliot-review-final-report-july2014.pdf/ Accessed 5th April 2019

Fadzillah, N., Man, Y., Rohman, A., Rosman, A., Ismail, A., Mustafa, S., & Khatib, A. (2015). Detection of Butter Adulteration with Lard by Employing (1)H-NMR Spectroscopy and Multivariate Data Analysis. *Journal of Oleo Science, 64*(7), 697-703.

Fadzlillah, N., Rohman, A., Ismail, A., Mustafa, S., & Khatib, A. (2013). Application of FTIR-ATR spectroscopy coupled with multivariate analysis for rapid estimation of butter adulteration. *Journal of Oleo Science, 62*(8), 555-562.

Fang, X., & Zhang, C. (2016). Detection of adulterated murine components in meat products by TaqMan© real-time PCR. *Food Chemistry, 192*, 485-490.

Fermo, P., Beretta , G., Maffei Facino , R., Gelmini , F., & Piazzalunga , A. (2013). Ionic profile of honey as a potential indicator of botanical origin and global environmental pollution. *Environmental pollution, 178*, 173-181.

Ferreira, I., & Caçote, H. (2003). Detection and quantification of bovine, ovine and caprine milk percentages in protected denomination of origin cheeses by reversed-phase high-performance liquid chromatography of beta-lactoglobulins. *Journal of Chromatography A, 1015*(1-2), 111-118.

Fontecha, J., Mayo, I., Toledano, G., & Juárez, M. (2006). Triacylglycerol composition of protected designation of origin cheeses during ripening. Authenticity of milk fat. *Journal of Dairy Science, 89*(3), 882-887.

Fragni, R., Trifirò, A., & Nucci, A. (2015). Towards the development of a multi-element analysis by ICP-oa-TOF-MS for tracing the geographical origin of processed tomato products. *Food Control, 48*, 96-101.

Fuselli, F., Deluca, A., Montepeloso, E., Ibba, G., Tidona, F., Longo, L., & Marianella, R. (2015). Detection of fraudulent addition of bovine whey in water buffalo ricotta cheese by isoelectric focusing. *Journal of the Science of Food and Agriculture, 95*(13), 2757-2762.

Gallardo-Velázquez, T., Osorio-Revilla, G., Zuñiga-de Loa, M., & Rivera-Espinoza, Y. (2009). Application of FTIR-HATR spectroscopy and multivariate analysis to the quantification of adulterants in Mexican honeys. *Food Research International, 42*, 313-318.

Giaretta, N., Di Giuseppe, A., Lippert, M., Parente, A., & Di Maro, A. (2013). Myoglobin as marker in meat adulteration: a UPLC method for determining the presence of pork meat in raw beef burger. *Food Chemistry, 141*(3), 1814-1820.

Giunchi, A., Berardinelli, A., Ragni, L., Fabbri, A., & Silaghi, F. (2008). Non-destructive freshness assessment of shell eggs using FT-NIR spectroscopy. *Journal of Food Engineering, 89*(2), 142-148.

Goodacre, R., & Anklam, E. (2001). Fourier transform infrared spectroscopy and chemometrics as a tool for the rapid detection of other vegetable fats mixed in cocoa butter. *Journal of the American Oil Chemists' Society, 78*(10), 993-1000.

Guerreiro, J., Barros, M., Fernandes, P., Pires, P., & Bardsley, R. (2013). Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses. *Food Chemistry, 136*(3-4), 1526-1532.

Guo, W., Liu, Y., Zhu, X., & Wang, S. (2011). Dielectric properties of honey adulterated with sucrose syrup. *Journal of Food Engineering, 107*(1), 1-7.

Haase, I., & Fischer, M. (2007). Differenzierung von Theobroma cacao und Theobroma grandiflorum mittels PCR. *Journal für Verbraucherschutz und Lebensmittelsicherheit, 2*(4), 422-428.

Heidarbeigi, K., Mohtasebi , S., Foroughirad, A., Ghasemi-Varnamkhasti, M., Rafiee, S., & Rezaei, K. (2015). Detection of adulteration in Saffron samples using electronic nose. *International Journal of Food Properties, 18*(7), 1391-1401.

Hong, X., & Wang, J. (2014b). Detection of adulteration in cherry tomato juices based on electronic nose and tongue: Comparison of different data fusion approaches. *Journal of Food Engineering, 126*, 89-97.

Hong, X., & Wang, J. (2014c). Use of Electronic Nose and Tongue to Track Freshness of Cherry Tomatoes Squeezed for Juice Consumption: Comparison of Different Sensor Fusion Approaches. *Food and Bioprocess Technology, 8*(1), 158-170.

Hong, X., Wang, J., & Qi, G. (2015). E-nose combined with chemometrics to trace tomato-juice quality. *Journal of Food Engineering, 149*, 38-43.

Hsieh, C.-W., Li, P.-H., Cheng, J.-Y., & Ma, J.-T. (2013). Using SNIF-NMR method to identify the adulteration of molasses spirit vinegar by synthetic acetic acid in rice vinegar. *Industrial Crops and Products, 50*, 904-908.

Huang, Q., Qiu, N., Ma, M., Jin, Y., Yang, H., Geng, F., & Sun, S. (2012). Estimation of egg freshness using S-Ovalbumin as an indicator. *Poultry Science, 91*(3), 739-743.

Ihunegbo, F., Tesfalidet, S., & Jiang, W. (2010). Determination of melamine in milk powder using zwitterionic HILIC stationary phase with UV detection. *Journal of Separation Science, 33*(6-7), 988-995.

Jaiswal, P., Jha, S. N., Kaur, J., & Borah, A. (2017). Detection and quantification of anionic detergent (lissapol) in milk using attenuated total reflectance-Fourier Transform Infrared spectroscopy. *Food. Chem., 221*, 815-821.

Jakes, W., Gerdova, A., Defernez, M., Watson, A., McCallum, C., Limer, E., . . . Kemsley, E. (2015). Authentication of beef versus horse meat using 60 MHz 1H NMR spectroscopy. *Food Chemistry, 175*, 1-9.

Kamruzzaman, M., Makino, Y., Oshita, S., & Liu, S. (2015). Assessment of Visible Near-Infrared Hyperspectral Imaging as a Tool for Detection of Horsemeat Adulteration in Minced Beef. *Food and Bioprocess Technology, 8*(5), 1054-1062.

Karoui, R., Dufour, E., Pillonel, L., Picque, D., Cattenoz, T., & Bosset, J.-O. (2004). Determining the geographic origin of Emmental cheeses produced during winter and summer using a technique based on the concatenation of MIR and fluorescence spectroscopic data. *European Food Research and Technology, 219*(2), 184-189.

Karoui, R., Schoonheydt, R., Decuypere, E., Nicolaï, B., & Baedermaeker, J. (2007). Front face fluorescence spectroscopy as a tool for the assessment of egg freshness during storage at a temperature of 12.2 °C and 87% relative humidity. *Analytica Chimica Acta, 582*(1), 83-91.

Kemps, B., De Ketelaere, B., Bamelis, F., Mertens, K., Decuypere, E., De Baerdemaeker, J., & Schwagele, F. (2007). Albumen freshness assessment by combining visible near-infrared transmission and low-resolution proton nuclear magnetic resonance spectroscopy. *Poultry Science, 86*(4), 752-759.

Knödler, M., Most, M., Schieber, A., & Carle, R. (2010). A novel approach to authenticity control of whole grain durum wheat (Triticum durum Desf.) flour and pasta, based on analysis of alkylresorcinol composition. *Food Chemistry, 118*(1), 177-181.

Kuswandi, B., Cendekiawan, K., Kristiningrum, N., & Ahmad, M. (2015). Pork adulteration in commercial meatballs determined by chemometric analysis of NIR Spectra. *Journal of Food Measurement and Characterization, 9*(3), 313-323.

Leygonie, C., Britz, T., & Hoffman, L. (2012). Impact of freezing and thawing on the quality of meat: Review. *Meat Science, 91*(2), 93-98.

Li, A., Ha, Y., Wang, F., & Li, Y. (2010). Detection of Hydrocarbons in Irradiated Chilled Beef by HS-SPME–GC–MS and Optimization of the Method. *JAOCS, Journal of the American Oil Chemists' Society, 87*(7), 731-736.

Li, J., Ding, X.-M., Liu, D.-D., Guo, F., Chen, Y., Zhang, Y.-B., & Liu, H.-M. (2013). Simultaneous determination of eight illegal dyes in chili products by liquid chromatography-tandem mass spectrometry. *Journal of Chromatography B, 942-943*, 46-52.

Li, S., Shan, Y., Zhu, X., Zhang, X., & Ling, G. (2012). Detection of honey adulteration by high fructose corn syrup and maltose syrup using Raman spectroscopy. *Journal of Food Composition and Analysis, 28*(1), 69-74.

Liu, F., Yusuf, B., Zhong, J., Feng, L., He, Y., & Wang, L. (2011). Variety Identification of Rice Vinegars Using Visible and Near Infrared Spectroscopy and Multivariate Calibrations. *International Journal of Food Properties, 14*(6), 1264-1276.

Liu, J., Zhao, Q.-S., Li, M.-Q., Chen, J., & Zhao, G.-H. (2015). Rapid Detection of Adulterate Starch in Chocolate by Near-infrared Spectroscopy. *Modern Food Science and Technology, 31*(3), 260-265.

Liu, Y., Ying, Y., Ouyang, A., & Li, Y. (2007). Measurement of internal quality in chicken eggs using visible transmittance spectroscopy technology. *Food Control, 18*(1), 18-22.

Ma, P., Liang, F., Sun, Y., Jin, Y., Chen, Y., Wang, X., . . . Song, D. (2013). Rapid determination of melamine in milk and milk powder by surface-enhanced Raman spectroscopy and using cyclodextrin-decorated silver nanoparticles. *Microchimica Acta, 180*(11-12), 1173-1180.

Macarthur, R., Crews, C., & Brereton, P. (2000). An improved method for the measurement of added vegetable fats in chocolate. *Food Additives and Contaminants, 17*(8), 653-664.

Masiri, J., Benoit, L., Thienes, C., Kainrath, C., Barrios-Lopez, B., Agapov, A., . . . Samadpour, M. (2017). A rapid, semi-quantitative test for detection of raw and cooked horse meat residues. *Food Control, 76*, 102-107.

Maurer, N., & Rodriguez-Saona, L. (2013). Rapid Assessment of Quality Parameters in Cocoa Butter Using ATR-MIR Spectroscopy and Multivariate Analysis. *Journal of the American Oil Chemists' Society, 90*(4), 475-481.

Monahan, F., Schmidt, O., & Moloney, A. (2018). Meat provenance: Authentication of geographical origin and dietary background of meat. *Meat Science, 144*, 2-14.

Mondello, L., Casilli, A., Tranchida , P., Dugo, P., Costa, R., Festa, S., & Dugo, G. (2004). Comprehensive multidimensional GC for the characterization of roasted coffee beans. *Journal of separation science , 27*(5-6), 442-450.

Montowska, M., Alexander, M., Tucker, G., & Barrett, D. (2015). Authentication of processed meat products by peptidomic analysis using rapid ambient mass spectrometry. *Food Chemistry, 187*, 297-304.

Morales, V., Corzo, N., & Sanz, M. (2008). HPAEC-PAD oligosaccharide analysis to detect adulterations of honey with sugar syrups. *Food Chemistry, 107*(2), 922-928.

Morin , J., Jamin, E., Guyader, S., & Thomas, F. (2018). Coffee. In J. Morin, & M. Lees, *FoodIntegrity Handbook: a guide to food authenticity issues and analytical solutions* (p. 295-314). Nantes: Eurofins Analytical France.

Morsy, N., & Sun, D.-W. (2013). Robust linear and non-linear models of NIR spectroscopy for detection and quantification of adulterants in fresh and frozen-thawed minced beef. *Meat Science, 93*(2), 292-302.

Mounicouy, S., Szpunary, J., Andreyz, D., Blakez, C., & Lobinsky, R. (2003). Concentrations and bioavailability of cadmium and lead in cocoa powder and related products. *Food Additives and Contaminants, 20*, 343-352.

Natonek-Wiśniewska, M., Krzyścin, P., & Piestrzyńska-Kajtoch, A. (2013). The species identification of bovine, porcine, ovine and chicken components in animal meals, feeds and their ingredients, based on COX I analysis and ribosomal DNA sequences. *Food Control, 34*(1), 69-78.

Navarro-Pascual-Ahuir, M., Lerma-García, M., Simó-Alfonso, E., & Herrero-Martínez, J. (2015). Rapid differentiation of commercial juices and blends by using sugar profiles obtained by capillary zone electrophoresis with indirect UV detection. *Journal of Agricultural and Food Chemistry, 63*(10), 2639-2646.

Niu , W., & Liu, Y. (2014). Discrimination of vinegars using a catalytic nanomaterials-based chemiluminescence sensor array. *Luminescence, 29*(2), 138-142.

Novotná, H., Kmiecik, O., Gałązka , M., Krtková, V., Hurajová, A., Schulzová, V., . . . Hajšlová, J. (2012). Metabolomic fingerprinting employing DART-TOFMS for authentication of tomatoes and peppers from organic and conventional farming. *Food Additives and Contaminants-Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 29*(9), 1335-1346.

Oliveira , R., Oliveira, L., Franca, A., & Augusti, R. (2009). Evaluation of the potential of SPME-GC-MS and chemometrics to detect adulteration of ground roasted coffee with roasted balrey. *Journal of food composition and analysis , 22*(3), 257-261.

Padovan, G., De Jong, D., Rodrigues, L., & Marchini, J. (2003). Detection of adulteration of commercial honey samples by the 13C/12C isotopic ratio. *Food Chemistry, 82*(4), 633-636.

Pan, X., Chen, J., Chen, Q., Huang, B., & Han, J. (2018). Authentication of pork in meat mixtures using PRM mass spectrometry of myosin peptides. *RSC Advances, 8*(20), 11157–11162.

Papotti, G., Bertelli, D., Graziosi, R., Maietti, A., Tedeschi, P., Marchetti, A., & Plessi, M. (2015). Traditional balsamic vinegar and balsamic vinegar of Modena analyzed by nuclear magnetic resonance spectroscopy coupled with multivariate data analysis. *LWT- Food Science and Technology, 60*(2), 1017-1024.

Paradkar, M., & Irudayaraj, J. (2002). Discrimination and classification of beet and cane inverts in honey by FT-Raman spectroscopy. *Food Chemistry, 76*(2), 231-239.

Pauli, E., Barbieri, F., Garcia, P., Madeira, T., Acquaro, J., Scarmino leda, S., . . . Nixdorf, S. (2014). Detection of ground roasted coffee adulteration with roasted soybean and wheat (with pulsed amperometric detection). *Food Research international , 61*, 112-119.

Pellegrino,, L., & Tirelli, A. (2000). A sensitive HPLC method to detect hen's egg white lysozyme in milk and dairy products. *International Dairy Journal, 10*(7), 435-442.

Petrakis, E., Cagliani, L., Polissiou, M., & Consonni, R. (2015). Evaluation of saffron (Crocus sativus L.) adulteratrion with plant adulterants by 1H NMR metabolite fingerprinting. *Food Chemistry, 173*, 890-896.

Picariello, G., Sacchi, R., Fierro, O., Melck, D., Romano, R., Paduano, A., . . . Addeo, F. (2013). High resolution 13C NMR detection of short- and medium-chain synthetic triacylglycerols used in butterfat adulteration. *European Journal of Lipid Science and Technology, 115*(8), 858-864.

Piergiovanni, A. (2007). Extraction and separation of water-soluble proteins from different wheat species by acidic capillary electrophoresis. *Journal of Agricultural and Food Chemistry, 55*(10), 3850-3856.

Pillonel, L., Albrecht, B., Badertscher, R., Buetikofer, U., & Bosset, J. (2003a). Analytical methods for the determination of the geographic origin of Emmental cheese. Parameters of proteolysis and rheology. *Italian Journal of Food Science, 15*(1), 49-62.

Pillonel, L., Ampuero, S., Tabacchi, R., & Bosset, J. (2003b). Analytical methods for the determination of the geographic origin of Emmental cheese: volatile compounds by GC/MS-FID and electronic nose. *European Food Research and Technology, 216*(2), 179-183.

Pillonel, L., Badertscher, R., Froidevaux, P., Haberhauer, G., Hölzl, S., Horn, P., . . . Bosset, J. (2003c). Stable isotope ratios, major, trace and radioactive elements in emmental cheeses of different origins. *LWT-Food Science and Technology, 36*(6), 615-623.

Pizarro, C., Esteban-Diez, I., & Gonzalez-Saiz, J. (2007). Mixture resolution according to the percentage of robusta variety in order to detect adulteration in roasted coffee by near infrared spectroscopy. *Analytica Chimica Acta , 585*(2), 266-276.

Pontillon, J. (1995). Determination of milk fat in chocolates by gas-liquid chromatography of triglycerides and fatty acids. *Journal of the American oil Chemists' Society, 72*(8), 861-866.

Ponzi, L. (2017, May). Cibo criminale (Criminal Food). *Proceedings of the 4th Food Integrity Conference-Turning Science into Solutions* (p. 87-88), ParmaPrandi, B., Lambertini, F., Faccini, A., Suman, M., Leporati, A., Tedeschi, T., & Sforza, S. (2017). Mass spectrometry quantification of beef and pork meat in highly processed food: Application on Bolognese sauce. *Food Control, 74*, 61-69.

Radovic, B., Careri, M., Mangia, A., Musci, M., Gerboles, M., & Anklam, E. (2001a). Contribution of dynamic headspace GC–MS analysis of aroma compounds to authenticity testing of honey. *Food Chemistry, 72*(4), 511-520.

Radovic, B., Goodacre, R., & Anklam, E. (2001b). Contribution of pyrolysis-mass spectrometry (Py-MS) to authenticity testing of honey. *Journal of Analytical and Applied Pyrolysis, 60*(1), 79-87.

Rahman, M., Ali, M., Hamid, S., Bhassu, S., Mustafa, S., Al Amin, M., & Razzak, M. (2015). Lab-on-a-Chip PCR-RFLP Assay for the Detection of Canine DNA in Burger Formulations. *Food Analytical methods, 8*(6), 1598-1606.

Reis, N., Botelho, B., Franca, A., & Oliveira, L. (2017). Simultaneous Detection of Multiple Adulterants in Ground Roasted Coffee by ATR-FTIR Spectroscopy and Data Fusion. *Food Analytical Methods, 10*, 2700-2709.

Reis, N., Franca, A., & Oliveira, L. (2013). Performance of diffuse reflectance infrared Fourier transform spectroscopy and chemometrics for detection of multiple adulterants in roasted ground coffee. *LWT Food Science and Technology , 53*(2), 395-401.

LWT Food Science and Technology , 53(2), 395-401.

Reis, N., Franca, A., & Oliveira, L. (2016). Concomitant Use of Fourier Transform Infrared Attenuated Total Reflectance Spectroscopy and Chemometrics for Quantification of Multiple Adulterants in Roasted and Ground Coffee. *Journal of Spectroscopy, 2016*, 1-7.

Rektorisova, M., & Tomaniova, M. (2018). Cocoa, cocoa preparation, chocolate and chocolate-based confectionery. In J. Morin, & M. Lees, *FoodIntegrity Handbook: a guide to food authenticity issues and analytical solutions* (p. 137-153). Nantes: Eurofins Analytics France.

Rentsch , J., Weibel , S., Ruf , J., Eugster, A., Beck, K., & Köppel, R. (2013). Interlaboratory validation of two multiplex quantitative real-time PCR methods to determine species DNA of cow, sheep and goat as a measure of milk proportions in cheese. *European Food Research and Technology, 236*(1), 217-227.

Restani, P., Velona, T., Carpen, A., Duranti, M., & Galli, C. (1996). gamma-Casein as a marker of ripening and/or quality of Grana Padano cheese. *Journal of Agricultural and Food Chemistry, 44*(8), 2026-2029.

Rodriguez Mondal, A., Desmarcheller, A., Konings, E., Acheson-Shalom, R., & Delatour, T. (2010). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) method extension to quantify simultaneously melamine and cyanuric acid in egg powder and soy protein in addition to milk products. *Journal of Agricultural and Food Chemistry, 58*(22), 11574-11579.

Rohmana, A., Sismindari, Erwanto, Y., & Che Man, Y. (2011). Analysis of pork adulteration in beef meatball using Fourier transform infrared (FTIR) spectroscopy. *Meat Science, 88*(1), 91-95.

Rossi, M., Hidalgo, A., & Pompei, C. (2001). Reaction between albumen and 3,3',5,5'-tetramethylbenzidine as a method to evaluate egg freshness. *Journal of Agricultural Food Chemistry, 49*(8), 3522-3526.

Ruiz Orduna, A., Husby, E., Yang, C., Ghosh, D., & Beaudry, F. (2015). Assessment of meat authenticity using bioinformatics, targeted peptide biomarkers and high-resolution mass spectrometry. *Food additives and contaminants. Part A, Chemistry, Analysis, Control, Exposure and Risk Assesment, 32*(10), 1709-1717.

Ruiz-Matute, A., Rodríguez-Sánchez, S., Sanz, M., & Martínez-Castro, I. (2010). Detection of adulterations of honey with high fructose syrups from inulin by GC analysis. *Journal of Food Composition and Analysis, 23*(3), 273-276.

Russo, R., Cusano, E., Perissi, A., Ferron, E., Severino, V., Parente, A., & Chambery, A. (2014). Ultra-high performance liquid chromatography tandem mass spectrometry for the detection of durum wheat contamination or adulteration. *Journal of Mass Spectrometry, 49*(12), 1239-1246.

Russo, R., Severino, V., Mendez, A., Lliberia, J., Parente, A., & Chambery, A. (2012). Detection of buffalo mozzarella adulteration by an ultra-high performance liquid chromatography tandem mass spectrometry methodology. *Journal fo Mass Spectrometry, 47*(11), 1407-1414.

Sabatino, L., Scordino, M., Gargano, M., Belligno, A., Traulo, P., & Gagliano , G. (2011). HPLC/PDA/ESI-MS evaluation of saffron (Crocus sativus L.) adulteration. *Natural Product Communications, 6*(12), 1873-1876.

Sáiz-Abajo, M., González-Sáiz , J., & Pizarro, C. (2005). Orthogonal signal correction applied to the classification of wine and molasses vinegar samples by near-infrared spectroscopy. Feasibility study for the detection and quantification of adulterated vinegar samples. *Analytical and Bioanalytical Chemistry, 382*(2), 412-420.

Sakaridis, I., Ganopoulos, I., Argiriou, A., & Tsaftaris, A. (2013). A fast and accurate method for controlling the correct labeling of products containing buffalo meat using High Resolution Melting (HRM) analysis. *Meat Science, 94*(1), 84-88.

Santos, P., Pereira-Filho, E., & Rodriguez-Saona, L. (2013). Rapid detection and quantification of milk adulteration using infrared microspectroscopy and chemometrics analysis. *Food Chemistry, 138*(1), 19-24.

Santos, P., Wentzell, P., & Pereira-Filho, E. (2012). Scanner Digital Images Combined with Color Parameters: A Case Study to Detect Adulterations in Liquid Cow's Milk. *Food Analytical Methods, 5*(1), 89-95.

Schneider, N., Werkmeister, K., Becker, C., & Pischetsrieder, M. (2011). Prevalence and stability of lysozyme in cheese. *Food Chemistry, 128*(1), 145-151.

Scholl, P., Farris, S., & Mossoba, M. (2014). Rapid turbidimetric detection of milk powder adulteration with plant proteins. *Journal of Agricultural and Food Chemistry, 62*(7), 1498-1505.

Sciuto, S., Dell'Atti, L., Esposito, G., Martucci, F., Guglielmetti, C., & Acutis, P. L. (2017). Rapid Screening Technique To Identify Sudan Dyes (I to IV) in Adulterated Tomato Sauce, Chilli Powder, and Palm Oil by Innovative High-Resolution Mass Spectrometry. *Journal of food protection, 80*(4), 640-644.

Shao, Y., Xie, C., Jiang, L., Shi, J., Zhu, J., & He, Y. (2015). Discrimination of tomatoes bred by spaceflight mutagenesis using visible/near infrared spectroscopy and chemometrics. *Spectrochimica Acta-Part A: Molecular and Biomolecular Spectroscopy, 140*, 431-436.

Silva, B., Andrade, P., Mendes, G., Valentão, P., Seabra, R., & Ferreira, M. (2000). Analysis of phenolic compounds in the evaluation of commercial quince jam authenticity. *Journal of Agricultural and Food Chemistry, 48*(7), 2853-2857.

Soares, S., Amaral, J., Oliveira, M., & Mafra, I. (2013). A SYBR Green real-time PCR assay to detect and quantify pork meat in processed poultry meat products. *Meat Science, 94*(1), 115-120.

Soares, S., Amaral, J., Oliveira, M., & Mafra, I. (2015). Improving DNA isolation from honey for the botanical origin identification. *Food Control, 48*, 130-136.

Sonnante, G., Montemurro, C., Morgese, A., Sabetta, W., Blanco, A., & Pasqualone, A. (2009). DNA microsatellite region for a reliable quantification of soft wheat adulteration in durum wheat-based foodstuffs by real-time PCR. *Journal of Agricultural and Food Chemistry, 57*(21), 10199-10204.

Sousa, M., Dias, L., Veloso, A., Estevinho, L., Peres, A., & Machardo, A. (2014). Practical procedure for discriminating monofloral honey with a broad pollen profile variability using an electronic tongue. *Talanta, 128*, 284-292.

Spangenberg, J., & Dionisi, F. (2001). Characterization of Cocoa Butter and Cocoa Butter Equivalents by Bulk and Molecular Carbon Isotope Analyses: Implications for Vegetable Fat Quantification in Chocolate. *Journal of Agricultural and Food Chemistry, 49*(9), 4271-4277.

Spaniolas, S., Tsachaki, M., Bennett, M., & Tucker, G. (2008). Evaluation of DNA extraction methods from green and roasted coffee beans. *Food Control , 19*(3), 257-262.

Suman, M., Cavanna, D., Zerbini, M., Ricchetti, D., Sanfelici, D., Cavandoli, E., & Mirone, L. (2018). Eggs & egg products. In J. Morin, & M. Lees, *Foodintegrity Handbook* (p. 27 - 42). Nantes: Eurofins Analytics France.

Sun, Y.-L., Kang, H.-M., Kim, Y.-S., Baek, J.-P., Zheng, S.-L., Xiang, J.-J., & Hong, S.-K. (2014). Tomato (Solanum lycopersicum) variety discrimination and hybridization analysis based on the 5S rRNA region. *Biotechnology and Biotechnical Equipment, 28*(3), 431-437.

Terzi, V., Malnati, M., Barbanera, M., Stanca, A., & Faccioli, P. (2003). Development of analytical systems based on real-time PCR for Triticum species-specific detection and quantitation of bread wheat contamination in semolina and pasta. *Journal of Cereal Science, 38*(1), 87-94.

Teye, E., Huang, X., Dai, H., & Chen, Q. (2013). Rapid differentiation of Ghana cocoa beans by FT-NIR spectroscopy coupled with multivariate classification. *Spectrochimica Acta-Part A: Molecular and Biomolecular Spectroscopy, 114*, 183-189.

Teye, E., Huang, X., Han, F., & Botchway, F. (2014b). Discrimination of Cocoa Beans According to Geographical Origin by Electronic Tongue and Multivariate Algorithms. *Food Analytical Methods, 7*(2), 360-365.

Teye, E., Huang, X., Takrama, J., & Haiyang, G. (2014a). Integrating NIR Spectroscopy and Electronic Tongue Together with Chemometric Analysis for Accurate Classification of Cocoa Bean Varieties. *Journal of Food Process Engineering, 37*(6), 560-566.

Tezcan, F., Uzaşçı, S., Uyar, G., Oztekin, N., & Erim, F. (2013). Determination of amino acids in pomegranate juices and fingerprint for adulteration with apple juices. *Food Chemistry, 141*(2), 1187-1191.

Thomas, F., & Jamin, E. (2009). 2H NMR and 13C-IRMS analyses of acetic acid from vinegar, 18O-IRMS analysis of water in vinegar: international collaborative study report. *Analytica Chimica Acta, 649*(1), 98-105.

Toledo, B., Hantao , L., Ho, T., Augusto, F., & Anderson, J. (2014). A chemometric approach toward the detection and quantification of coffee adulteration by solid-phase microextraction using polymeric ionic liquid sorbent coatings. *Journal of Chromatography A , 1346*, 1-7.

Tomaszewska-Gras, J. (2012). Detection of butter adulteration with water using differential scanning calorimetry. *Journal of Thermal Analysis and Calorimetry, 108*(2), 433-438.

Vaclavik, L., Schreiber, A., Lacina, O., Cajka, T., & Hajslova, J. (2012). Liquid chromatography–mass spectrometry-based metabolomics for authenticity assessment of fruit juices. *Metabolomics, 8*(5), 793-803.

Vardin, H., Tay , A., Ozen, B., & Mauer, L. (2008). Authentication of pomegranate juice concentrate using FTIR spectroscopy and chemometrics. *Food Chemistry, 108*(2), 742-748.

Voldrich, M., Rajchl, A., Cizkova, H., & Cuhra, P. (2009). Detection of Foreign Enzime Addition into the Adulterated Honey. *Czech J. Food Sci., 27*.

von Bargen, C., Brockmeyer, J., & Humpf , H. (2014). Meat authentication: a new HPLC-MS/MS based method for the fast and sensitive detection of horse and pork in highly processed food. *Journal of Agricultural and Food Chemistry, 62*(39), 9428-9435.

Wang, S., Guo, Q., Wang, L., Lin, L., Shi, H., Cao, H., & Cao, B. (2015). Detection of honey adulteration with starch syrup by high performance liquid chromatography. *Food Chemistry, 172*, 669-674.

Woodcock, T., Downey, G., & O'Donnell, C. (2009). Near infrared spectral fingerprinting for confirmation of claimed PDO provenance of honey. *Food Chemistry, 114*(2), 742-746.

Woodcock, T., Fagan, C., O'Donnell, C., & Downey, G. (2008). Application of Near and Mid-Infrared Spectroscopy to Determine Cheese Quality and Authenticity. *Food and Bioprocess Technology, 1*(2), 117-129.

Xia, X., Ding, S., Li, X., Gong, X., Zhang, S., Jiang, H., . . . Shen, J. (2009). Validation of a confirmatory method for the determination of melamine in egg by gas chromatography-mass spectrometry and ultra-performance liquid chromatography-tandem mass spectrometry. *Analytica Chimica Acta, 651*(2), 196-200.

Xie, H., Bu, L., & Peng, X. (2012). Ultraviolet spectroscopy method for classifying vinegars. *Advanced Material Research, 346*, 865-874.

Yilmaz, M., Tatlisu, N., Toker, O., Karaman, S., Dertli, E., Sagdic, O., & Arici, M. (2014). Steady, dynamic and creep rheological analysis as a novel approach to detect honey adulteration by fructose and saccharose syrups: Correlations with HPLC-RID results. *Food Research International, 64*, 634-646.

Yin, R., Sun, Y., Yu, S., Wang, Y., Zhang, M., Xu, Y., Xue, J., & Xu, N. (2016). A validated strip-based lateral flow assay for the confirmation of sheep-specific PCR products for the authentication of meat. *Food Control, 60*, 146-150.

Zanardi, E., Caligiani, A., Padovani, E., Mariani, M., Ghidini, S., Palla, G., & Ianieri, A. (2013). Detection of irradiated beef by nuclear magnetic resonance lipid profiling combined with chemometric techniques. *Meat Science, 93*(2), 171-177.

Zhao, J., Li, H., Chen, Q., Huang, X., Sun, Z., & Zhou, F. (2010). Identification of egg's freshness using NIR and support vector data description. *Journal of Food Engineering, 98*(4), 408-414.

Zhao, M., Downey, G., & O'Donnell, C. (2014). Detection of adulteration in fresh and frozen beefburger products by beef offal using mid-infrared ATR spectroscopy and multivariate data analysis. *Meat Science, 96*(1), 1003-1011.

Zhou, J., Qi, Y., Ritho, J., Duan, L., Wu, L., Diao, Q., . . . Zhao, J. (2014). Analysis of maltooligosaccharides in honey samples by ultra-performance liquid chromatography coupled with evaporative light scattering detection. *Food Research International, 56*, 260-265.

Zhu, X., Li, S., Shan, Y., Zhang, Z., Li, G., Su, D., & Liu, F. (2010). Detection of adulterants such as sweeteners materials in honey using near-infrared spectroscopy and chemometrics. *Journal of Food Engineering, 101*(1), 92-97.

The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: a proposed validation workflow to bring about a harmonized approach

Daniele Cavanna[1,2], Laura Righetti[2], Chris Elliott[3], Michele Suman[1]

[1]Barilla Advanced Laboratory Research, Via Mantova 166 - 43122 Parma, Italy
[2]Department of Food and Drug, University of Parma, Parco Area delle Scienze 95/A- 43124 Parma, Italy
[3]Faculty of Medicine, Health & Life Sciences, Queen's University Belfast, BT9 6NG Belfast, Northern Ireland

Introduction

Food fraud should be considered as a highly dynamic activity in which fraudsters aim to escape the regulatory and industry controls, for instance by hiding or changing the type of adulterants employed.

So far, most of the methods applied for the control of food fraud are targeted methods, which are focused on the detection of one or a few classes of compounds. In many cases the extraction procedures are complex and expensive, but enable to lower the analytes detection limits (up to sub ppt-levels) also in complex matrices (Kaufmann, Butcher, Maden, Walker, & Widmer, 2015). However, recent advances in mass spectrometry, mainly high resolution mass spectrometry (HRMS) together with improvements in user friendly software, have allowed non-targeted approaches to be developed (Kaufmann, 2014). In this type of methods, target inclusion lists are not used, since the molecules to be detected are not known *a priori.*

Indeed, the analysis aims to study the global sample fingerprint itself.

The increasing popularity of HRMS is mainly due to the introduction of benchtop instruments, such as Time-of-Flight (ToF) and Orbitrap, and the advantages of using full-scan acquisition mode with high sensitivity, combined with high resolving power (up to 100,000 FWHM) and accurate mass measurement (<5 ppm). In addition, the acquisition of high resolution full scanned data permits the combination of target analysis with screening of non-target compounds, novel compound identification, and retrospective data analysis. Moreover, a broad range of *m/z* values can be recorded simultaneously, without any target compounds list and

individual optimization (Kaufmann, 2012). In most cases a generic sample preparation is performed which allows, in principle, a very broad view of any potential compounds of interest. Therefore, the advantages of applying analytical methods not focused on a narrow group of targeted analytes are very evident, underling the need to widen the range of monitored compounds to combat the complexities of food adulteration.

Besides the previously mentioned advantages of using non-targeted approaches, some critical aspects have also to be taken into account. Non-targeted data-handling is much more demanding compared to that required in classical targeted approaches. In targeted analysis, results are usually evaluated compound-by-compound using univariate statistics. By contrast, the data collected for non-targeted approaches typically needs to be evaluated using multivariate statistical models (Riedl, Esslinger, & Fauhl-Hassek, 2015). In addition, the huge diversity in data processing workflows applied through the available scientific literature makes the evaluation of method performance extremely challenging. An agreed, harmonized and 'official' workflow for development and validation of non-targeted methods is very much required. A science-based approach was presented by Alewijn and co-workers (Alewijn, van der Voet, & van Ruth, 2016) in which a validation roadmap is described starting from the criteria for the selection of suitable samples for the training set, passing through the identification of the most appropriate analytical methods and continuing with a description of some initial validation steps (i.e. repeatability, permutation tests). Subsequently, a cross-validation of the training set and the prediction of an external set of samples are suggested as crucial points for a robust validation.

This publication was followed by a preliminary attempt of untargeted analysis harmonization, which has been recently suggested by the US Pharmacopoeia (USP Pharmacopeial Convention., 2016). In this document, great attention is placed on the criteria that must be used to build-up a "reference" and a "test" samples set able to provide a reliable predictive model. The concepts of sensitivity and specifity rate are introduced, together with the idea that the evaluation of the receiver operating characteristic (ROC) curves could represent an important tool to assess the goodness of the method.

According to this guideline, after the selection of the analytical approach and of the appropriate modelling technique, the method should be developed and optimized with the "reference set" and the "test set". Subsequently, sensitivity and specificity rates should be evaluated together with the ROC curves.

If the criteria are fulfill, the method should be validated with new samples and, after its relese for use, a monitoring process should be put in place with the aim to check the reliability of the method over a period of time (USP Pharmacopeial Convention., 2016).

This guideline is very welcome but the process should be considered to only be at the beginning. Furthermore, USP guideline is generic and so is not able to provide suggestions specific for each analytical technique.

Therefore, this article has been aimed to assess the current state of the art on non- targeted mass spectrometry in food fraud detection and to propose a harmonized workflow.

Literature Overview

Few keywords ("non-targeted", "mass spectrometry", "authenticity", "fraud" and the target matrix) were mainly set and used to obtain a representative set of studies from the Scopus database. A wide range of food matrices were considered, while mass spectrometry was the only analytical technique considered. Searches included articles published from 2011 to 2017. Altogether, 49 articles were evaluated, from 16 different scientific journals.

As detailed in Tables 1 & 2, information available in literature on food fraud deals with a large number of commodities, including meat, spices, wine and cereals. The most commonly addressed issues that pertain to geographical origin protection, proof of the authenticity, followed by the detection of different types of adulterations (i.e mixtures, dilutions, substitutions). .

As detailed in Table 3 (that resumes the validation parameters verified by the authors of each of the articles selected for this review) non-targeted mass spectrometry workflows appear to be non-standardized to date. Experimental design as well as crucial parameters (i.e. number of samples, number of replicates, sample sources) quite often are not clearly detailed in the articles or completely absent in some cases.

Another challenging issue that should be addressed in the future is the representativeness of the samples used to build the models. Authentic samples must be provided by certified producers and/or guaranteed by official center (i.e. PDO). In fact, the use of non-authentic samples could well result to a misleading classification model.

From an analytical point of view, sample preparation is usually simple and rapid: every cleanup step could potentially decrease the number of detected compounds, with a depletion of the chromatographic fingerprint (Vuckovic, 2012). Liquid chromatography coupled to mass spectrometry (especially high resolution mass spectrometry) is the most widely used approach,

followed by Direct Analysis in Real Time (DART)-MS, GC-MS, Proton Transfer Reaction (PTR)-MS and other techniques.

Generally, accepted chemometric models are used for data processing, with the exception of proteomic studies. Prediction clusters (multivariate models aiming to predict class membership with no marker selection) and discriminative model with markers identification (that ends with the identification of significant compounds responsible for class membership) are the chemometric approaches mainly presented in the literature so far.

The most applied unsupervised technique is the Principal Component Analysis (PCA), while Orthogonal Partial Least Square – Discriminant Analysis ((O)PLS-DA) and Linear Discriminant Analysis (LDA) are the most commonly employed supervised models. The identification of markers represents the most challenging and time-consuming step. Indeed, quite often they remain unknown (only the *m/z* values are provided) or only a tentative identification is presented (Kind & Fiehn, 2007). Only a few studies followed the criteria proposed by the Standard Initiative in metabolomics (Sumner, *et al*., 2007) as subsequently amended and supplemented (Schymanski, *et al*., 2014), performing a first level identification, thus by unambiguously confirming the identity with Reference Standards injection.

Most of the investigated papers reached a "Level II" that corresponds to compound identified by HRMS/MS spectra matching with literature or libraries ("Level IIa") or by diagnostic evidence when only one structure fits the experimental data but no standard or literature information is available for confirmation ("Level IIb").

On the other hands, metabolites are considered putatively characterized when evidence exists for possible structures but there is not enough information for one exact structure only ("Level III").

Finally, unknown compounds can be classified as "Level IV" when an unequivocal molecular formula can be assigned but a structure cannot be hypothesized and as "Level V" when a specific exact mass is important for the study but no information are useful to identify the compound (Sumner, *et al*., 2007) (Schymanski, *et al*., 2014).Another critical step is represented by the validation of chemometric models, essential to assess their reliability, but quite often this step is not undertaken in published studies. When applied, different and sometimes incomplete validation approaches are presented, suggesting, once again, the urgent need for a harmonization approach to method validation. These considerations will be further detailed in the following paragraphs.

Design of experiments (DOEs)

One important point is to define the study question being asked (i.e. geographical origin, organic or conventional regimen, percentage of adulteration). After the definition of a question, the study design has to be planned, involving the choice of which samples to collect and a reliable procedure for handling and measuring the samples minimizing the effect of nuisance variation (Bevilacqua , et al., 2017). However, quite often in non-targeted methods the number of experimental variables usually greatly exceeds the number of objects, especially with the development of new mass spectrometry-based technologies. Variables, i.e. metabolites or proteins, are represented by mass/retention time combinations and it is typical to have a huge number of features varying from several hundred to many thousands, depending on the experimental and analytical conditions. This increase in experimental possibilities, however, does not correspond to a proportional increase in the number of samples, leading to a serious complication for the statistical analysis, including the risk of type I and type II errors and thus lowering the prediction power (Franceschi, Vrhovsek, Mattivi, & Wehrens, 2012). Power analyses, that are an important aspect of experimental design, are often avoided and sample size determination seems to be driven by sample availability, laboratory practice or extrapolated from the existing literature. Approaches developed in other fields (Blaise, 2013) allow an efficient *a priori* evaluation of the number of samples to be included in a study in order to identify statistically significant variations throughout the data set. Indeed, power laws, analysis of variance (ANOVA) and ANOVA Simultaneous Component Analysis (ASCA) (Khakimov, Gürdeniz, & Engelsen, 2015) (Smilde, et al., 2005) can be used to decide on the minimum number of samples required (Blaise, *et al*., 2016). On-line software tools are also available for simple power/sample size calculations such as G*Power (http://www.gpower.hhu.de/en.html) or Glimmpse (https://glimmpse.samplesizeshop.org/#/) (Kreidler, et al., 2013). Despite recent efforts (Blaise, 2013) (Blaise, et al., 2016), there is still no widely accepted criteria for sample size determination even though it is necessary to gain confidence in the results generated.

Apart from the size of the data set, samples have to be representative and biological variation should be accurately defined. An appropriate number of biological replicates have to be included in each group to confidently answer the question in a statistically robust manner. Samples should be selected minimizing the non-controlled variations among them as much as possible (such as storage conditions) checking also that groups are balanced (Xia, Broadhurst, Wilson, & Wishart, 2013) (that is, when the number of samples in different classes do not vary

greatly). Ideally, we would measure one source of variation while controlling all other sources of variation. For example, when identify the metabolic changes resulting from the agricultural regimen of tomatoes, all source of variation must be controlled, such as genotypes, environmental factor, years of harvesting. By controlling other sources of variation, we are confident that the observations are due to the parameter we are testing, that means the organic/conventional condition. However, when performing an open field study, it is difficult to control these sources of variation and so instead we need to ensure the variability is equivalent between different groups of tomatoes. For example, we have to ensure the range of tomatoes varieties, growing locations and harvesting years is equivalent in the samples group because we know that this will influence the measured metabolome. So, when investigating the differences between a group of tomatoes cultivated in organic and in conventional conditions, and the organic samples are from Spain while the conventional are cultivated in Netherlands, there are two sources of variations, agricultural conditions and geographical origin. Where these sources of variation are not matched then the metabolic changes observed are a result of a combination of the two or more of them. If the geographical origin of tomatoes is similar within the groups, then the influence is removed and the study now has only one sources of variation, the conventional/organic regimen.

In any authenticity research, "authentic" samples are essential in order to understand the natural variation within a population. These samples can be provided by producers or prepared by researchers under controlled conditions. Also, the adulterated samples could be formulated 'fit for purpose', e.g. by dilution (relevant in wine or juice authentication) (Rubert, Zachariasova, & Hajslova, 2015).

One of the critical steps in non-targeted approach is the sample preparation which has hugely important consequences on the accuracy of the analytical results produced. Ideally, no sample preparation would be required for the analysis of samples as every manipulation might influence the reproducibility. This is feasible only for some matrices (i.e. wine) or by using direct analysis techniques, such as DART (Rubert, Lacina, Fauhl-Hassek, & Hajslova, 2014), Rapid Evaporative Ionization Mass Spectrometry (REIMS) (Balog, *et al*., 2016) and PTR-MS (Granato, Koot, & van Ruth, 2015) (Black, Chevallier, & Elliott, 2016).

In order to further increase the number of detected molecules (thousands), both positive and negative ionization modes should be performed. Therefore, sample handling should be minimal, non-selective, aiming to get coverage of a wider range of analytes by simple preparation steps. As no single analytical technique is suitable for the detection and

identification of the “true fingerprints”, multiplatform approaches represent the best solution. In this context, sample pre-treatment should be the minimum possible to make it compatible with the instrumental techniques.

Data processing to extract meaningful markers

For processing massive information based on separation techniques and mass spectrometry, effective software tools capable of rapid data mining procedures must be employed. Note that data matrices contain thousands of variables (*m/z*, retention time, intensity), and they have to be converted into more manageable information (Bevilacqua et al 2017). According to the literature review, metabolomic approaches are the most often applied. In these type of studies, data processing and data pretreatment must be carried out in order to permit the identification of significant compounds, which capture the bulk of variation between different datasets and may therefore potentially serve as biomarkers (Riedl, Esslinger, & Fauhl-Hassek, 2015). Data processing usually involves four basic steps: deconvolution, alignment, filtering and gap filling (Riccadonna & Franceschi, 2018) (Mastrangelo, Ferrarini, Rey-Stolle, García, & Barbas, 2015). The features, defined by their *m/z* and retention time, and their intensities in different samples are used for the statistical analysis. Samples can be grouped and can be observed using score plots, heatmaps or hierarchical clustering. After data pretreatment, a statistical comparison can be performed using the multivariate data analysis (MVDA). Usually this step involves unsupervised models (PCA) and supervised classification tools, such as PLS-DA and OPLS-DA (Franceschi et al. 2012). These supervised methods are performed to maximize differences between groups and to highlight potential biomarkers.

Validation

In food authenticity studies, non-targeted approaches results end up with a prediction of class belonging or with the validation of discriminative models and thus identification of few markers. To verify the relevance of modelling results, appropriate model validation is essential for non-targeted approaches, but it is often found to be used insufficiently or inconsistently.

As example, only 35% of the articles presented in this opinion mentioned the use of an external set of samples for the validation of the model while 25% of papers do not perform any validation study (not even a cross-validation study). Furthermore, when target compounds are identified, only 12% of the articles clearly indicated that the relative reference standard were injected. In addition, only 10% of works performed an ROC curves evaluation and only 30% of papers monitored the analytical variability by using quality control (QC) samples.

Analytical validation

The first type of validation that should be considered is the analytical one, which usually include the randomization of the injection samples order and the quality control samples (QCs). QCs samples are a pool of equal volumes of each sample of the set. These mixtures must be injected at the beginning of the run to equilibrate and stabilize the system as well as at regular intervals throughout the sequence run (i.e. every 10 injections). For non-targeted analysis, the quality of the instrumental performance is checked by the tight clustering in the center of the plot of QC in the preliminary PCA (Godzien, Alonso-Herranz, Barbas, & Armitage, 2015). The US FDA has proposed other useful criteria for analytical method validation (Food and Drug Administration, 2001) to calculate the relative standard deviation (RSD) of each analyte in the QCs: features with RSD% higher than 40% should be filtered out, since they would not be good candidates as markers.

So far, only a few studies presented in literature mentioned the preparation and evaluation of QCs, either showing the related scores plots, that is a fundamental evidence of a good analytical procedure. In addition, the use of an internal standard (Want, et al., 2013) is strongly recommended, even though not largely applied in the food fraud field as yet. By adding this compound(s), the performance of the extraction procedure and the chromatographic run can be checked, and it could also be used as an additional tool for post processing data evaluation (i.e. normalizing the peaks area with the Internal Standard ones) (Khamis, Adamko, & El-Aneed, 2017). In the research presented by Khamis *et al*., for example, a deuterated internal standards is used to compensate the matrix effect and the relative ion suppression. Moroever, in the work of Ulaszewska *et al*, $^{13}C_1$ labeled creatinine and trans cinnamic acid-d_5 were used in urine metabolomics as additional control of the extraction efficiency (Ulaszewska, et al., 2016). Recently, Boysen *et al* presented a workflow, called best-matched internal standard (B-MIS) normalization, that should lead to normalize peaks signal according to the isotope-labeled internal standards that have a similar behavior during the analysis (Boysen, Heal, Carlson, & Ingalls, 2018).

Internal validation

Multivariate approaches suffer from overfitting, thus validation is an obligatory component of any analysis. Typically, cross-validation approaches are used, in which a proportion of the data (e.g., 10-40%, the “validation set”) are randomly removed, and the model is built with the remaining “training set”. This procedure is repeated many times until each sample has been in

the test set exactly once (leave-*n*-out procedure). The accuracy of the model on these left-out samples gives an estimate of the predictive power for unseen samples and also the robustness of the model to perturbations of the data.

Model performance is usually described by the goodness-of-fit parameter (R^2X), the proportion of the variance of the response variable that is explained by the model (R^2Y) and the predictive ability parameter (Q^2). Most of the papers investigated through this review reported quite nice prediction abilities. Indeed, Rubert *et al* while investigating the wine authentication presented values always higher than 0.7 in all the supervised models created (Rubert, Lacina, Fauhl-Hassek, & Hajslova, 2014). Another interesting example is the Oregano study performed by Black *et al* in which the three parameters are always higher than 0.9 in all the models created (Black, Haughey, Chevalier, Galvin-King, & Elliott, 2016) indicating excellent classification performance as well as prediction ability.

In a few studies, also permutation testing and Monte Carlo simulation were used as a tool to avoid overfitting (Riedl, Esslinger, & Fauhl-Hassek, 2015). In addition, sensitivity (percentage of samples correctly classified) and specificity (percentage of samples correctly rejected) are used to evaluate the classification performance (USP Pharmacopeial Convention, 2016). The visual tool is represented by the receiver operator characteristic (ROC) curves, that has not been extensively applied in food studies (Springer et al. 2014; Righetti et al. 2018), but widely elsewhere (Xia, Broadhurst, Wilson, & Wishart, 2013). Recently (Righetti, *et al.*, 2018), this tool was employed to verify the relaibility of markers to identify the durum what aduteration in a confirmatory study. The authors reported area under the curve (AUC) for the most significant markers values ranging from 65% to 100% (Righetti, *et al.*, 2018), and thus being classified as excellent markers.

External validation

Evaluating the repeatability and the performance of a model is an invaluable and crucial step before the introduction of this new model in routine practice. Independent external validation should be performed by the assessment of an external set of samples that were not used for model building. Samples should be critically selected in order to demonstrate the validity of the model and expand the application of the method, modeling all the possible sources of variability of the considered matrix. Therefore, for example different geographical origins, growing seasons, cultivars, and producers have to be considered and included in the building of the model.

Interesting examples of this approach can be found in the articles studied for this opinion.
In the honey floral origin discrimination performed by Jandric´ *et al.*, 33 samples from 4 different botanical origin were used for the model validation (Jandrić, et al., 2015). Moreover, the geographic origin discrimination between different Spanish Extra Virgin Olive Oils (EVOO) presented by Gil-Solsona *et al*. was validated with 15 samples from cultivars representative of all the Spanish EVOOs and collected in a different season with respect to the samples used for the model creation.
Finally, a complete confirmatory metabolomic study with an higher amount of samples and with the introduction of more sources of variability was executed by Righetti *et al*. in order to confirm the markers selected in the preliminary study. (Righetti, *et al.*, 2018).

Marker validation

If the ultimate goal of the non-targeted approach is to move markers from the research laboratory to the food authorities control routine practice, the significance of the markers must be confirmed:

- *During the external validation study, where more sources of variability are considered;*
- *Evaluating the marker performance with the area under the curve (AUC) of the ROC;*
- *By a survey of blind real samples;*
- *By the analysis of admixture samples (that are samples in which authentic sample is mixed up with different percentage of adulterant), especially if a legislation level is established for the target fraud.*

Another step of validation can be considered the integration of multiplatform data as well as data fusion. Indeed, multiplatform characterization of food samples with subsequent data fusion has been shown to improve prediction ability of multivariate models in authenticity testing (Biancolillo, Bucci, Magrì, Magrì, & Marini, 2014). The concatenation of analytical information from complementary instrumental techniques can be established on different levels. Data fusion is becoming more and more important in food authentication but appropriate preprocessing and model validation are required (Riedl, Esslinger, & Fauhl-Hassek, 2015) (Biancolillo, Bucci, Magrì, Magrì, & Marini, 2014).
In this context, ring trails are highly recommended to assess the reliability of non-targeted approaches across different laboratories. To the best of our knowledge, they have not been applied yet in the food fraud analysis, but some attempts of "metabo-ring tests" were reported recently in literature (Martin, *et al*., 2015) (Cajka, Smilowitz, & Fiehn, 2017). Bringing

together different mass spectrometers across Europe, the authors obtained consistent results and interestingly, no effect of the LC-MS instrumentation (TOF, QTOF,LTQ-Orbitrap) was reported. It should be noted that the Standard Operating Procedure (SOP) for sample preparation and specific statistical design were the same and undoubtedly played an important role in the quality of the study. In the work presented by Martin *et al.*, eleven different mass spectrometers were used for two tests: in the first one two groups of urine samples were analyzed, one of them spiked with 32 standard metabolites. Interstingly,all the mass spectrometer instruments were able to discriminate between the two groups of samples and most of the spiked compounds were identified as features responsible of the clusterization. In the second test, blood samples collected from rats fed with low or high Vitamin D diets were analyzed. The separation of the two groups was not satisfactory but it was due to the low biological contrast of the two groups and not to differences through the platform. Moreover, the trends detected with different instruments were comparable. (Martin, *et al.*, 2015)

In the work presented by *Cajka et al.*, nine mass spectrometers were used to perform a non targeted lipidomic study on human plasma samples. The classification results obtained were in good agreement and the most discriminative lipids found by each instrument overlapped in the 92% of the cases.

Coming back to the food field, however, large amounts of samples are very difficult to find and specific ring tests for each target fraud (and for each commodity) should be developed.

Another opportunity is to make the collection of massive amounts of non-targeted data available to all investigators who are interested in undertaking analysis. As an example, "MetaboLights" is an on-line repository where, for each study, data are shared by following a data-protocol deposition procedure to fully detail the experiment (Kale, *et al*., 2016). The database is cross-species and cross-technique and covers metabolite structures and their reference spectra, as well as their biological roles, locations, concentrations, and experimental data from different experiments. If data for the food fraud of interest can be shared, the raw files could be potentially used as independent data set for an external validation: the model should be able to correctly cluster these samples, even if probably the number of extracted features will not be exactly the same. Additionally, if the non-targeted study aims to identify relevant markers, the shared results could be helpful to identify robust compounds potentially already detected in different laboratories or, if not, to merge complementary results. A possible step in this direction would be to make data-protocol-algorithm deposition a prerequisite for publication in all peer-reviewed scientific journals.

Applicability in official and legal trials

Official procedures require a high degree of result reproducibility across different laboratories, instrumentation and analysts. So far non-targeted method have been reported to be mainly "in-house" developed and validated, with little focus on inter-laboratory reproducibility. For this reason, standardization of analyses performed at different laboratories will be a challenge, but it should be recognized as essential to allow data to be more widely comparable. To take advantage of some recent analytical breakthroughs, the combination of non-targeted and semi-targeted strategies is strongly recommended to provide a robust approach for a short but well defined list of compounds. Due to their differing objectives, the application of non-targeted methods for quantitative or confirmatory purpose remains challenging. Yet these approaches should be considered essential to select and identify significant markers, mainly when the adulterant is unknown. However, once markers are identified, optimized methods should be applied for the accurate quantification required by a confirmatory purpose, as recently presented by different research groups (Wielogorska, *et al*., 2018) (Jandric, Islam, Singh, & Cannavan, 2017). Subsequently, these 'biomarker target methods' can be shared and applied through different laboratories.

To summarize, in order to be able to present untargeted analysis in legal trials and have them accepted, non-targeted models must be fully validated as detailed in this article ensuring that identified compounds are highly specific of the food fraud detected.

If the non-targeted approach ends with a prediction cluster, only unknown samples having an adulteration level in line with those used to build the model (e.g. that are classified in a specific "not-compliant cluster") can be declared as fraudulent and thus illegal.

Conclusions

This Opinion has set out to summarize the scientific activities published up to now on the non-targeted mass spectrometry approaches to food fraud detection. The authors have outlined a possible approach for the development and validation of these types of methods, taking into account that at the moment there is no harmonized, agreed or 'official' workflows. Additionally, global considerations on the applicability of these methods for legal purposes are provided.

Processes harmonization does appear to only be at the beginning and both public and private institutions will have to increase their efforts in order to finalize a shared approach, able to

guide the development of robust non-targeted methods for food fraud detection using spectrometric techniques.

Acknowledgement

This work was supported by the European Union, Seventh Framework Program (FP7/2007-2013), under Grant Agreement 613688 (Ensuring the Integrity of the European food chain − FoodIntegrity).

The authors would like also to thank Prof. Saskia van Ruth from Wageningen University and Research - The Netherlands, for her scientific suggestions devoted to further improve the technical quality of the present paper.

Commodiy	Sub-commodity	Number of papers	Method											Chemometric Mehtod								
			LC-MS	LC-HRMS	HRMS	DART-MS	REIMS-HRMS	MALDI-HRMS	MALDI-LRMS	LESA-HRMS	SPME-MS	PTR-MS	GC-MS	PCA	(O)PLS-DA	SIMCA	(L)DA	k-NN	(H)CA	k-Means	PCO	CP-ANN
EVOO		4	2					1				1		2	3			1	1			1
FISH		1	1																			
Balsamic Vinegar		1							1					1								
Coffee		2			1							1		2	2							
Cheese		2		1								1		1	1	1						
Tomatoes		2				1							1	2			1					
Fruit Juices		3	1	3										3	2	1	1					
Honey		5		3								1	1	5	3		1	1				
Carrot		1		1										1	1							
Meat		8	2	2		2	1			1		1		4	1	1	2					
Herbs and spices	Saffron	3		2		1							1	3	2					1		1
	Oregano	1		1										1	1							
Wine		4		2		1					1		1	4	3		1					
Spirit drinks	Whisky	3		1	2									2	1		1					
Sea Buckthorn Oil		1		1		1																
Soybean		1		1		1								1	1							
Cereals	Wheat	6	1	4									1	4	3	1			1		1	
Tiger nut		1												1	1							

PCA: Principal component analysis, (O)PLS-DA: (orthognol) partial least squares discriminant analysis, (L)DA: (Linear) discriminant analysis , SIMCA: Soft independent modelling by class analogy, k-NN: k-Nearest-Neighbors, (H)CA: (hierarchical) cluster analysis, PCO: principal coordinates analysis, CP-ANN: Counter propagation artificial neural network, LC-MS: liquid chromatography-mass spectrometry, LC-HRMS. liquid chromatography-high resolution mass spectrometry, HRMS: high resolution mass spectrometry, DART-MS. direct analysis in real time-mass spectrometry, REIMS-HRMS: rapid evaporative ionization mass spectrometry, MALDI-HRMS: matrix-assisted laser desorption/ionization- high resolution mass spectrometry, MALDI-LRMS: matrix-assisted laser desorption/ionization- low resolution mass spectrometry,

LESA-HRMS: liquid extraction surface analysis-high resolution mass spectrometry, SPME-MS. Solid phase micro extraction- mass spectrometry, PTR-MS: proton transfer reaction–mass spectrometry, GC-MS: gas chromatography-mass spectrometry

Table1: Commodities, number of papers, instrumental method and chemometric method used in the articles that meet the selection criteria used for the literature search

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Kalogiouri, Alygizakis, Aalizadeh, & Thomaidis, 2016)	Authenticity	Extra virgin olive oil	Defective olive oil	Liquid-liquid extraction, centrifugation and filtration	LC-HRMS	N.A.	XCMS	PLS-DA; CP-ANN	2 - classes	Predictive clusters	From the Kolovi and Adramitiani varieties	22
(Di Girolamo, et al., 2015)	Adulteration	Extra virgin olive oil	Corn oil addition	Liquid-liquid extraction and centrifugation	MALDI-HRMS	Normalization	Biotyper 3.1, ClinProTools	HCA; PCA	2 - classes	Predictive clusters	Pure oil samples and home maid mixtures	8
(Gil-Solsona, et al., 2016)	Geographical origin discrimination	Extra virgin olive oil	False geographical origin declaration	Liquid-liquid extraction (polar fraction) / dilution (non polar fraction)	LC-HRMS	Mean centering normalization, log2 transformation	EZInfo	PCA; PLS-DA; OPLS-DA	2-classes	Disciminative	From local cooperatives	105
(Ruiz-Samblás, et al., 2012)	Monovarietal extra virgin olive oil identification	Extra virgin olive oil	False monovarietal extra virgin olive oil declaration	No sample prep	PTR-MS	Autoscaling	MATLAB	PLS-DA	5-classes	Predictive clusters	Samples purchased from common markets in Spain	30
(Wulff, Nielsen, Deelder, Jessen, & Palmblad, 2013)	Authenticity	Fish	N.A.	Proteins extraction and digestion	LC-MS	N.A.	DataAnalysis	N.A.	N.A.	N.A.	Samples from different fishes species	69

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Guerreiro , de Oliveira , Ferreira , & Catharino , 2014)	Adulteration	Balsamic Vinegar of Modena	Addition of other vinegars	Dilution	MALDI-LRMS	N.A.	Unscrambler	PCA	6-classes	Disciminative	All vinegars purchased from commercial establishments	N.A.
(Garrett, et al., 2013)	Arabica Coffee Geographical origin discrimination	Coffee	False geographical origin declaration	Solid-Liquid extraction and centrifugation	HRMS	Pareto scaling	MetaboAnalyst	PCA; PLS-DA	2-classes	Disciminative	Coffe samples from a field study	8 for each region
(Özdestan, et al., 2013)	Specialty coffees discrimination	Coffee	False organic declaration	No sample prep	PTR-MS	Autoscaling	Pirouette	PCA; PLS-DA	2-classes	Predictive clusters	Samples collected from retail outlests in the Netherlands	110
(Popping, De Dominicis, Dante, & Nocetti, 2017)	Geographical origin of Paramigiano Reggiano	Cheese	False geographical origin declaration	Solid-Liquid extraction	LC-HRMS	Pareto scaling	SIMCA	SIMCA	2-classes	Predictive clusters	Samples commercially available	84
(Galle, et al., 2011)	Geographical origin of PDO cheese of Leiden	Cheese	False geographical origin declaration	No sample prep	PTR-MS	Autoscaling	Pirouette; STATISTICA	PCA; PLS-DA	2-classes	Predictive clusters	Samples purchased from supermarkets and from registered cheese farms	59
(Figueira, Câmara, Pereira, & Câmara, 2014)	Cultivars discrimination	Tomatoes	N.A.	Dilution and SPME extraction	GC-MS	Normalization	SPSS	PCA	5-classes	Predictive clusters	Samples obtained from local producers	5

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Novotná, et al., 2012)	Organic and conventional farming discrimination	Tomatoes	False organic declaration	Solid-Liquid extraction and centrifugation	DART-MS	Normalization	statistiXL	PCA; LDA	2-classes	Predictive clusters	Samples obtained from a field study	40
(Jandric, et al., 2014)	Grapefruit and Oranje juices authenticity	Fruit Juices	Low price fruit juices addition	Centrifugation and filtration	LC-HRMS	N.A.	EZinfo	PCA; OPLS-DA	6-classes for each juice	Disciminative	Fruits purchased from Austrian markets	28
(Jandric, Islam, Singh, & Cannavan, 2017) 2017	Indian citrus fruit/ fruit juices adulteration	Fruit Juices	N.A.	Centrifugation and filtration	LC-HRMS	Pareto Scaling	SIMCA	PCA; SIMCA;OPLS-DA	4-classes	Disciminative	Fruit juices produced in the laboratory from authentic citrus juices	N.A.
(Vaclavik, Schreiber, Lacina, Cajka, & Hajslova, 2012)	Authenticity	Fruit Juices	Addition of low price juices to expensive juices	Centrifugation. filtration and dilution	LC-MS/LC-HRMS	Pareto scaling	MarkerView, statistiXL	PCA; LDA	3-classes	Disciminative	Purchased from Czech and Canadian markets	84
(Li, et al., 2017)	Single/multi floral and geographical origin authenticity	Honey	N.A.	Liquid-liquid extraction, Centrifugation and filtration	LC-HRMS	Range Scaling	SIMCA	PCA; PLS-DA	2-classes	Disciminative	Honey samples collected from apiaries	210
(Silva, et al., 2017)	Sugarcane honey authenticity	Honey	N.A.	No sample prep, direct extraction with SPME	GC-MS	N.A.	SPSS; STATISTICA	PCA; LDA	2-classes	Discriminative	Samples were from Madeira Island, Portugal	16

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Jandrić, et al., 2015)	Floral origin authentication	Honey	N.A.	Dilution, centrifugation and filtration	LC-HRMS	Pareto scaling	SIMCA	PCA; OPLS-DA	4-classes	Disciminative	Samples obtained from honey producers	83
(Spiteri, et al., 2016)	Botanical origin	Honey	N.A.	Liquid-liquid extraction (Orbitrap); QuEChERS (TOF-MS)	LC-HRMS	Log transformation and Pareto Scaling	R; Matlab	PCA; PLS-DA	5-classes	Discriminative	Samples collected from different supermarkets	56
(Kus & van Ruth, 2015)	Floral origin authentication	Honey	N.A.	No sample prep	PTR-MS	Autoscaling	Pirouette	PCA; k-NN	6-classes	Predictive clusters	Samples collected in Poland	62
(Cubero-Leon, De Rudder, & Maquet, 2018)	Organic authentication	Carrot	Conventional carrots in declared organic carrots	Solid-Liquid extraction, centrifugation and filtration	LC-HRMS	Pareto scaling	SIMCA	PCA; OPLS-DA	2-classes	Predictive clusters	Samples from field study	210

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Balog, et al., 2016)	Identification of the species of origin for meat products	Meat	Addition of not declared meat species	No sample prep, direct analysis with REIMS	REIMS-HRMS	N.A.	MATLAB	PCA; LDA	Different classifications are explored	Predictive clusters	Samples from regional abattoir	15 bovine and 5 equine samples . All samples were divided into four pieces, and a total of 30 sampling points were taken in four separate experiments. Total: 600 sampling points
(Vaclavik, et al., 2011)	Animal fats (lard and beef tallow) authentication	Meat	Pork lard and beef tallow admixtures	Solid-Liquid extraction and centrifugation	DART-HRMS	Constant row sum transformation	StatistiXL, Statistica	PCA; LDA	2-classes	Predictive clusters	Samples purchased from the retail market	29
(von Bargen, Brockmeyer, & Humpf, 2014)	Horse and pork detection in higly processed food	Meat	Not declared meat mixtures	Solid-Liquid extraction, digestion, desalting	LC-MS	N.A.	N.A.	N.A.	N.A.	N.A.	Samples purchased from the market and in-house processed samples	23

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Prandi, et al., 2017)	Highly processed meat authentication	Meat	Fraudulent beef and pork meat mixtures in Bolognese sauce	Proteins extraction and digestion	LC-MS	N.A.	N.A.	N.A.	N.A.	N.A.	Samples prepared in a pilot plant	N.A.
(Cajka, Danhelova, Zachariasova, Riddellova, & Hajslova, 2013)	Chicken feeding discrimination	Meat	Bone meal addition to chicken feed	Solid-Liquid extraction and centrifugation	DART-MS	Constant row sum transformation; Pareto scaling	SIMCA	PCA; OPLS-DA	2-classes	Predictive clusters	Chickens grown under controlled conditions	116 (chicken) 24 (feed)
(Sarah, et al., 2016)	Discrimination of pork from beef, chevon from chicken in thermally processed meat	Meat	Fraudulent meats mixtures	Proteins extraction and digestion	LC-HRMS	N.A.	N.A.	N.A.	N.A.	N.A.	Authentic meat purchased directly from commercial abattoirs	4
(Montowska, Alexander, Tucker, & Barrett, 2015)	Beef, pork, horse, chicken and turkey meat authentication in processed samples	Meat	Fraudulent meats mixtures	Protein digestion and centrifugation	LESA-HRMS	N.A.	N.A.	N.A.	N.A.	N.A.	Meat products purchased at supermarket	18

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Oliveira, Alewijn, Boerrigter-Eenling, & van Ruth, 2015)	Discrimination of Conventional, Free Range, and Organic pork meat	Meat	Fraudulent organic declaration	No sample prep (PTR-MS); fat melted and dissolved in DCM (LC-HRMS)	LC-HRMS , PTR-MS	Autosaling (LC-HRMS); Pareto scaling (PTR-MS)	Pirouette	PCA, SIMCA	3-classes	Predictive clusters	Samples purchased by a meat producing company in the Netherlands	41
(Rubert, Lacina, Zachariasova, & Hajslova, 2016)	Authentic PDO saffron vs unknown samples	Saffron	False origin declaration	Solid-Liquid extraction and centrifugation	DART-MS; LC-HRMS	Pareto scaling	MarkerView, SIMCA	PCA; OPLS-DA	2-classes	Discriminative	Samples purchased from markets.	44
(Aliakbarzadeh, Sereshti, & Parastar, 2016)	Geographical regions origin of Iranian saffron authentication	Saffron	False origin declaration	Ultrasound-assisted solvent extraction (UASE) and dispersive liquid–liquid microextraction (DLLME)	GC-MS	MCR–ALS technique for chromatographic peaks resolving	MATLAB	PCA; k-Means; CP-ANN	5-classes	Discriminative	Sample obtained from producers and agricultural research centers	17
(Guijarro-Díez, Nozal, Marina, & Crego, 2015)	Authenticity certification for Spanish and Iranian Saffron	Saffron	Fraudulent mixtures	Solid-Liquid extraction, centrifugation, dilution and filtration	LC-HRMS	Log transformation and Pareto Scaling	SIMCA	PCA; PLS-DA; OPLS-DA	2-classes	Discriminative	Samples provided by "Carmencita" company	20

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Black, Haughey, Chevalier, Galvin-King, & Elliott, 2016)	Adulteration	Oregano	Different adulterants (i.e.addition of olive or myrtle leaves etc)	Solid-Liquid extraction, centrifugation and filtration	LC-HRMS	Mean centering and Pareto scaling	SIMCA	PCA; OPLS-DA	6-classes	Discriminative	Samples purchased from markets.	N.A.
(Vaclavik, Lacina, Hajslova, & Zweigenbaum, 2011)	Discrimination of red wine varieties	Wine	N.A.	No sample prep, direct injection in HPLC	LC-HRMS	Mean centering and log transformation	Mass Profiler Professional	PCA; PLS-DA	3-classes	Discrminative	Samples purchased from retail markets.	45
(Springer, et al., 2014)	Wine botanical origin verification	Wine	False botanical origin declaration	No sample prep, direct extraction with SPME	GC-MS	Square root transformation, row and column scaling	SIMCA	PCA; PLS-DA; OPLS-DA	4-classes	Predictive clusters	Commercial wine samples	198
(Rubert, Lacina, Fauhl-Hassek, & Hajslova, 2014)	Wine varietal authentication	Wine	N.A.	Direct injection for fingerpirnt; Liquid-liquid extraction for Polyphenol profiling	LC-HRMS; DART-MS	Total area sum normalization	MarkerView, SIMCA	PCA; OPLS-DA	3-classes (white varieties), 3-class (red varieties)	Discriminative	Commercial wines	343
(Ziółkowska, ìWąsowicz, & Jeleń, 2016)	Grape variety and geographical origin discrimination	Wine	False grape variety and false geographical origin declaration	vial SPME extraction	SPME-MS	Autoscaling and log transformation	Statistica	PCA; LDA	Varieties: 3-classes (white),2-classes (red). Origin: 9-classes (white) 5-classes (red)	Predictive clusters	Wines purchased in wine shops	79

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Kew, Goodall, Clarke, & Uhrín, 2017)	Characterization of Whiskie types	Spirit Drinks-whisky	Whiskie adulteration	Dilution	HRMS	Mean centering and scaling to unit variance	SIMCA	PCA;OPLS-DA	2-classes	Discriminative	Commercial samples	85
(Collins, Zweigenbaum, & Ebeler, 2014)	Characterization of Whiskies type and aging	Spirit Drinks-whisky	Whiskey producers and ages	Direct injection	LC-HRMS	N.A.	XLSTAT	LDA	N.A.	Discriminative	Commercial samples	63
(Garcia, et al., 2013)	Whiskies adulteration detection	Spirit Drinks-whisky	Addition of not declared components	Dilution	HRMS	N.A.	Pirouette	PCA	5-class (different brands); 2-class (adulterated and not adulterated)	Predictive clusters	Commercial samples	80
(Hurkova, Rubert, Stranska-Zachariasova, & Hajslova, 2017)	Adulteration	Sea Buckthorn Oil	Addition of sunflower oil	Liquid-Liquid extraction, centrifugation.	DART-MS; LC-HRMS	N.A.	N.A.	N.A.	N.A.	Discriminative	Samples purchased at supermarket	N.A.
(Hrbek, et al., 2017)	GMO and Non-GMO Soybean discrimination	Soybean	Presence of GMO soybean	Solid-Liquid extraction, centrifugation and filtration	DART-MS; LC-HRMS	Pareto scaling	SIMCA	PCA; OPLS-DA	2-classes	Discriminative	Samples from field study	49
(Pastor, et al., 2016)	Small grains and corn flour authentication	Cereals - Wheat	Fraudulent mixed flour bakery products	Solid-Liquid extraction, centrifugation	GC-MS	N.A.	PAST	PCO; CA	2-classes	Discriminative	Samples from field study	40

Reference	Objective	Commodity	Fraud investigated	Sample prep	Method	Data pre treatment	Software	Chemometric methods	Classifications	Output	Sample source	Total No Samples
(Prandi , et al., 2012)	Durum wheat authenticity	Cereals - Wheat	Addition of not declared common wheat	Enzimatic cleavage	LC-MS	N.A.	N.A.	N.A.	N.A.	Discriminative	Analysis of wheat varieties and commercial samples	N.A.
(Geng, Harnly, & Chen, 2016)	Whole grain and refined wheat flours discrimination in bread	Cereals- Wheat	Refined wheat addition to whole wheat for bread preparation	Solid-liquid extraction, centrifugation and filtration	LC-HRMS	N.A.	SIMCA	PCA; SIMCA	2-classes	Discriminative	Samples purchased from local grocery stores or online	72
(Righetti, et al., 2016)	Ancient Triticum species characterization	Cereals- Wheat	N.A.	Solid-liquid extraction and centrifugation	LC-HRMS	Total area sum normalization, Pareto scaling	MarkerView, SIMCA	PCA; PLS-DA; OPLS-DA	3-classes	Discriminative	Wheat from field study	77
(Matthews, et al., 2012)	Common and durum wheat characterization	Cereals- Wheat	N.A.	Solid-liquid extraction and centrifugation	LC-HRMS	Pareto scaling	SIMCA	PCA; OPLS-DA; HCA	3-classes	Discriminative	Commercial samples	45
(Righetti, et al., 2018)	Common and durum wheat discrimination	Cereals- Wheat	Fraudulent common wheat addition	Solid-liquid extraction and centrifugation	LC-HRMS	Pareto scaling	SIMCA, SPSS	PCA; OPLS-DA	2-classes	Discriminative	Wheat samples from a field study	225
(Rubert, Hurkova, Stranska, & Hajslova, 2017)	Discrimination between PDO Valencian tiger nut and African tiger nut	Tiger nut	False PDO Valencian tiger nut declaration	Solid-Liquid extraction, centrifugation	LC-HRMS	Pareto scaling	MarkerView, SIMCA	PCA; OPLS-DA	2-classes	Discriminative	Xufa de València	45

PCA: Principal component analysis, (O)PLS-DA: (orthogonal) partial least squares discriminant analysis (L)DA: (Linear) discriminant analysis, , SIMCA: Soft independent modelling by class analogy, k-NN: k-Nearest-Neighbors, (H)CA: (hierarchical) cluster analysis, PCO: principal coordinates analysis, CP-ANN: Counter propagation artificial neural network LC-MS: liquid chromatography-mass spectrometry, LC-HRMS. liquid chromatography-high resolution mass spectrometry, HRMS: high resolution mass spectrometry, DART-MS. direct analysis in real time-mass spectrometry, REIMS-HRMS: rapid evaporative ionization mass spectrometry, MALDI-HRMS: matrix-assisted laser desorption/ionization- high resolution mass spectrometry, MALDI-LRMS: matrix-assisted laser desorption/ionization- low resolution mass spectrometry,

LESA-HRMS: liquid extraction surface analysis-high resolution mass spectrometry, SPME-MS. Solid phase micro extraction- mass spectrometry, PTR-MS: proton transfer reaction–mass spectrometry, GC-MS: gas chromatography-mass spectrometry, N.A.: information not available in the paper

Table 2: Summary of analytical approaches of the articles that meet the selection criteria, related to different raw materials

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Kalogiouri, Alygizakis, Aalizadeh, & Thomaidis, 2016)	N.A.	yes	N.A.	yes	Model validation with 4 samples used to evaluate the external accuracy. Markers confirmed with the injection of reference standards	yes	Samples produced in the 2014-2015 period and cultivated in different regions on Lesvos	Standard defective olive oils from the International Olive Council, other samples from the Kolovi and Adramitiani varieties	From the Kolovi and Adramitiani varieties	15	N.A.	N.A.
(Di Girolamo, et al., 2015)	4 replicates for all the samples	N.A.	N.A.	N.A.	Mixtures analyses. Two of them treated as blind samples and analyzed by an other investigator	N.A.	4 different italian cultivars; different commercial brands	Pure oil samples	In-house prepared mixtures	N.A.	N.A.	0.5%, 1%, 5%, 10%, 20% w/w addtion of Corn oil to EVOO

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Gil-Solsona, et al., 2016)	N.A.	yes	N.A.	yes	Cross- validation; analysis of 15 external samples	N.A.	Samples from 6 different spanish regions	Samples certified by the Spanish Agriculture ministry	Samples collected in different seasons	15	N.A.	N.A.
(Ruiz-Samblás, et al., 2012)	Samples analyzed in triplicate	N.A.	N.A.	yes	30% of samples used for internal cross validation and external validation	N.A.	Five different varieties of olive fruit (Arbequina, Cornicabra, Frantoio, Hojiblanca, and Picual) cultivated in Spain and Italy	N.A.	N.A.	N.A.	N.A.	N.A.
(Wulff, Nielsen, Deelder, Jessen, & Palmblad, 2013)	4 technical replicates	N.A.	N.A.	N.A.	Other samples prepared at different times,using different LC systems with different columns	N.A.	All included species were assigned by experienced zoologists	N.A.	N.A.	47	N.A.	N.A.
(Guerreiro , de Oliveira , Ferreira , & Catharino , 2014)	N.A.	N.A.	N.A.	yes	N.A.	N.A.	Other vinegars: red wine, white wine, apple and alcohol	PGI certification from the region of Modena (Italy)	N.A.	N.A.	N.A.	10%, 20% and 50% addition of each vinegar to the balsamic vinegar

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Garrett, et al., 2013)	All the samples in triplicate	N.A.	PCA: 2PCs 66% variance	yes	Leave one-out cross-validation	N.A.	Arabica coffee cultivars grown in two brasilian regions under the same edaphoclimatic conditions	N.A.	/	/	N.A.	N.A.
(Özdestan, et al., 2013)	All the samples in triplicate	N.A.	PCA: 2PCs 88.5% variance	yes	10-fold cross-validation	N.A.	Samples obtained in the last two months of 2011/early 2012. Organic and conventional samples originated from Southand Central America, Africa and Asia. A group of samples had a mixture of origins.	All groups of samples included fair trade certified coffees (54 samples in total)	N.A.	N.A.	N.A.	N.A.
(Popping, De Dominicis, Dante, & Nocetti, 2017)	All the samples in triplicate	yes	PCA-Class correctly classified 87.5% of samples	no	Leave one-out cross-validation; classification of an external set of samples	N.A.	Different geographical origin and ageing	A group of Parmigiano Reggiano samples were certified as PDO	Blind samples	32	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Galle, et al., 2011)	Analysis in triplicate	N.A.	PCA: 2PCs 45.6% variance	yes	Leave one-out cross-validation	N.A.	Artisanal farmers' cheese PDO samples and commercial Dutch cumin cheese samples of varying brands without PDO protection	All the artisanal farmers' cheese of Leiden samples were certified as PDO	N.A.	N.A.	N.A.	N.A.
(Figueira, Câmara, Pereira, & Câmara, 2014)	All the samples in triplicate	N.A.	PCA: 2PCs 87.4% variance	yes	Reference standards injection for a group of marker compounds; Leave one-out cross-validation	N.A.	All cultivars grew at southeast of Madeira Island (Portugal) and cultivated under similar conditions	N.A.	N.A.	N.A.	N.A.	N.A.
(Novotná, et al., 2012)	N.A.	N.A.	N.A.	yes	Leave one-out cross-validation	N.A.	Four locations, two farming methods, five varieties, two harvest years	N.A.	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Jandric, et al., 2014)	N.A.	yes	Different results according to the mixtures analyzed	yes	Some markers confirmed with reference standards. Markers detected in the admixtures samples	N.A.	N.A.	N.A.	In-house prepared mixtures	N.A.	N.A.	Pineapple juice adulterated with orange, apple, grapefruit and clementine were prepared at 1%, 5%, 10% and 15% adulteration. Orange juice was adulterated with apple, grapefruit and clementine at the same adulteration levels.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Jandric, Islam, Singh, & Cannavan, 2017)	3 replicates	N.A.	PCA: 2PCs 80,1% variance	yes	Leave one-out cross-validation; some markers confirmed with the injection of the reference standards. Subsequent Target method to detect markers	N.A.	Different citrus varieties grown in separate rows	Authentic citrus fruits obtained from the Indian Agriculture Research Institute.	Mixtures of citrus fruits juice adulterated with other juices	N.A.	N.A.	Mixtures of citrus fruits juice adulterated with each other were prepared at 1, 2, 5 and 10%
(Vaclavik, Schreiber, Lacina, Cajka, & Hajslova, 2012)	Mixtures in triplicate; single preparation for the pure samples	N.A.	PCA: 2PCs 45% variance (pos) 30,6% variance (neg)	yes	Leave one-out cross-validation and Sevenfold cross-validation LDA model created also with the mixtures	N.A.	Collected samples represented both fresh-pressed juices and juices prepared from concentrate and were produced in various countries	N.A.	In-house prepared mixtures	N.A.	N.A.	10%, 15%, 25%, 50%, 75% of apple juice addition to oragnge juice. 10%, 15%, 25%, 50%, 75%, 90% of grapefruit addition to orange juice
(Li, et al., 2017)	N.A.	yes	N.A.	yes	N.A.	N.A.	Honey samples collected from different provinces of China between March and September 2014	N.A.	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Silva, et al., 2017)	Triplicate analysis	N.A.	PCA: 2PCs 86,1% variance	yes	Leave one-out cross-validation	N.A.	Different producers and different sampling years.	Authentic samples obtained by certified producers	N.A.	N.A.	N.A.	N.A.
(Jandrić, et al., 2015)	n = 4	N.A.	PCA: 2PCs 49,6% variance.	yes	seven-fold cross-validation; classification of an external set of samples	N.A.	Samples from different species, collected from different locations in the North and South Islands of New Zealand	Authentic honey samples obtained from honey producers	Samples with manuka, clover, khamahi and rata floral origin	33	N.A.	N.A.
(Spiteri, et al., 2016)	Extraction in triplicate for Orbitrap-MS and in duplicate for TOF-MS	yes	PCA: 2PCs 36,4% variance (Orbitrap) 30,1% (TOF-MS)	yes	N.A.	N.A.	Four botanical origins, collected from different retail outlets	The samples were tested for their authenticity using pollen analysis, isotope ratio mass spectrometry and an NMR-profiling approach	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Kus & van Ruth, 2015)	Analysis in triplicate	N.A.	PCA: 2PCs 41,6% variance	yes	N.A.	N.A.	Six floral origin collected in 2010 and 2011	Floral analysis confirmed the origin	N.A.	N.A.	HPLC-DAD	N.A.
(Cubero-Leon, De Rudder, & Maquet, 2018)	15 from each field	yes	PCA: 2PCs 24,8% variance (ESI+) 24% (ESI-)	yes	Seven-fold cross-validation; external samples set validation; metbaolites identity confirmed with reference standards injection	yes	Carrot from different varieties, cultivated in two Belgium regions, in organic and conventional farming, over four consecutive years.	Certified organic and conventional samples were collected directly from the fields.	Carrot samples of Nerac and Namur varieties from organic and conventional farming	70	N.A.	N.A.
(Balog, et al., 2016)	Each meat species divided into four pieces	N.A.	N.A.	N.A.	Leave-20%-out cross-validation in the case of minced meat patties and leave-oneanimal-out cross-validation in the case of authentic meat samples	N.A.	Five equine and five bovine samples supplied by an Irish abattoir including two Hereford Cross, two Limousin Cross, and one Blonde Cross breed. 10 Scottish bovine) samples supplied by a Scottish abattoir; all samples were from different animals	N.A:	N.A.	N.A.	N.A.	1.25%:98.75%; 2.5%:97.5%; 5%:95%;10%:90 %; 33%:67% mixtures of horse, wagyu, venison and grain beef meats

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Vaclavik, et al., 2011)	Mixtures prepared in duplicate	N.A.	PCA: 2PCs 74,51% variance	yes	Leave one-out cross-validation; mixtures analyses	N.A.	Samples from different Czech producers	Pork lard and beef tallow pure samples	Mixtures prepared in the laboratory	12 admixtures samples	N.A.	Mixes of neat lard and beef tallow in the ratios 5:95, 10:90, 25:75, 50:50, 75:25, and 90:10 (w/w). Fat isolated from pork and beef mixed in the ratios 10:90, 25:75, 50:50, 75:25, and 90:10 (w/w)
(von Bargen, Brockmeyer, & Humpf, 2014)	N.A.	N.A.	N.A.	yes	N.A.	N.A.	Different supermarkets, date of purchase, packaging, storage conditions, end-products	N.A.	Processed samples prepared in-house and purchased from the market	N.A.	N.A.	yes

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Prandi, et al., 2017)	Samples extracted in duplicate	N.A.	N.A.	yes	Analysis of blind samples	N.A.	N.A.	N.A.	N.A.	3	N.A.	80:20; 60:40; 40:60; 20:80 (% of beef:pork)
(Cajka, Danhelova, Zachariasova, Riddellova, & Hajslova, 2013)	N.A.	N.A.	N.A.	yes	Four-fold cross-validation	N.A.	Group 1: chicken fed with the feed with the addition of chicken bone meal. Group 2: chicken fed with the feed without the addition of chicken bonemeal	N.A.	Chicken fed without the "contaminated" feed six months after the model creation	16	N.A.	N.A.
(Sarah, et al., 2016)	10 for each meat type	N.A.	N.A.	yes	Markers confirmation with a target LC-MS method	N.A.	Different abattoirs	N.A.	N.A.	N.A.	LC-MS target method	N.A.
(Montowska, Alexander, Tucker, & Barrett, 2015)	3 technical replicates	N.A.	N.A.	yes	N.A.	N.A.	Meat purchased at English and Polish supermarkets or manufactured in pilot plant	N.A.	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Oliveira, Alewijn, Boerrigter-Eenling, & van Ruth, 2015)	2 replicates (PTR-MS)	N.A.	N.A.	N.A.	20% of the samples used for the external validation of SIMCA models	N.A.	Commercial samples of conventional, "free range" and organic productions	N.A	N.A.	8	N.A.	N.A.
(Rubert, Lacina, Zachariasova, & Hajslova, 2016)	Duplicate of six samples	N.A.	PCA: 2PCs 58,7% variance (ESI+) 64,1% (ESI-)	yes	7-round internal cross-validation	N.A.	Samples collected from Spanish, Czech, Turkish markets.	10 samples labeled as PDO (Protected Designation of Origin).	N.A.	N.A.	N.A.	N.A.
(Aliakbarzadeh, Sereshti, & Parastar, 2016)	N.A.	N.A.	PCA: 2PCs 48,2% variance	yes	N.A.	N.A.	Samples from 8 different regions of Iran.	N.A.	N.A.	N.A.	N.A.	N.A.
(Guijarro-Díez, Nozal, Marina, & Crego, 2015)	Extraction in triplicate	yes	N.A.	yss	1/3 out cross-validation	N.A.	Samples from Spain and Iran.	Authentic samples certified according to ISO 3632 and HPLC analysis of dyes	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Black, Haughey, Chevalier, Galvin-King, & Elliott, 2016)	N.A.	yes	PCA: 2PCs 33,5% variance (ESI+)	yes	1/7 out cross-validation; analysis of commercial samples	N.A.	Samples sourced from different parts of the world	N.A.	Samples purchased at various retailers and web sites	78	yes (FTIR)	N.A.
(Vaclavik, Lacina, Hajslova, & Zweigenbaum, 2011)	N.A.	N.A.	PCA: 2PCs 45,6% variance (ESI+)	yes	1/3 out cross validation	N.A.	3 wine varieties from different geographical origins.	N.A.	N.A.	N.A.	N.A.	N.A.
(Springer, et al., 2014)	N.A.	N.A.	PCA: 2PCs explained from 25% to 39% variance (ESI+)	N.A.	2 set of samples measured on different devices; 1/7 out cross validation	yes	Wine varieties from different botanical origins.	Declaration of varietal purity (100%) according to the producers.	Additional wine samples, including Riesling , Pinot Gris and Pinot Blanc origin	74	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Rubert, Lacina, Fauhl-Hassek, & Hajslova, 2014)	Duplicate of six samples	N.A.	PCA for varieties: 2PCs 19,4% variance ESI+	yes	Seven round internal cross-validation	N.A.	Samples from 12 different countries, produced by different vineyards, from different grape varieties	Samples provided by the Federal Institute for Risk Assessment (Berlin, Germany).	N.A.	N.A.	N.A.	N.A.
(Ziółkowska, ìWąsowicz, & Jeleń, 2016)	3-5 replicates each sample	N.A.	PCA for varieties: 2PCs 67,6% variance (white wines);92,0% (red wines)	yes	20% of samples used as evaluation set; leave-one out cross-validation	N.A.	Commercially available white wines produced in Chile , USA , France, Bulgaria, Moldova, Spain, Argentina, Australia and South Africa, and red wines in Chile, Bulgaria, California, France and Moldova	N.A.	N.A.	N.A.	GC-MS	N.A.
(Kew, Goodall, Clarke, & Uhrín, 2017)	Three samples analyzed in replicate	N.A.	N.A.	yes	N.A.	N.A.	Samples from the 2010, 2012 and 2014 years. Scotch Whisky samples consisted of a mixture of malts (n=55) and blends (n=30)	Authentic reference samples provided by the Scotch Whisky Research Institute (SWRI).	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Collins, Zweigenbaum, & Ebeler, 2014)	Samples analyzed in triplicate	N.A.	N.A.	yes	N.A.	N.A.	37 bourbon whiskeys, 13 rye whiskeys, 6 Tennessee whiskeys and 7 other American whiskeys	N.A.	N.A.	N.A.	N.A.	N.A.
(Garcia, et al., 2013)	N.A.	N.A.	Different brands: PCA: 2PCs 50,9% variance (ESI+); 23,0% (ESI-) Adulteration: PCA: 2PCs 85,7% variance (ESI+); 89,6% (ESI-)	yes	N.A.	N.A.	Scotch whiskies from different brands.	Counterfeit samples provided by the Brazilian Federal Police.	N.A.	N.A.	N.A.	25, 50, 70, 80 and 90% addition of adulterant (brazilian distilled whisky)
(Hurkova, Rubert, Stranska-Zachariasova, & Hajslova, 2017)	N.A.	N.A.	N.A.	yes	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Hrbek, et al., 2017)	N.A.	yes	PCA: 2PCs 52,8% variance (ESI+) 68,3% (ESI-)	yes	*k*-fold internal cross-validation, analysis of certified reference materials of non GMO samples	N.A.	different varieties grown in different areas and agricoltural conditions.	Samples provided by the Crop Research Institute (CRI) of Prague.	N.A.	N.A.	N.A.	1%,5% and 10% of GMO addition to the CRM material
(Pastor, et al., 2016)	Three samples from every cereal flour sample	N.A.	N.A.	yes	N.A.	N.A.	23 small grain samples and 17 corn hybrid samples	Samples obtained from the Institute of Field and Vegetable Crops BNS Seme,^ Novi Sad, Serbia	N.A.	N.A.	N.A.	N.A.
(Prandi , et al., 2012)	Calibration curve prepared in triplicate	N.A.	N.A.	yes	Analysis of home made mixtures and commercial samples	N.A.	Wheat varieties were provided by Società Produttori Sementi SpA	N.A.	Mixtures and commercial samples analyses	N.A.	N.A.	5%, 20%, 56%,70%, 75%, 90% of common wheat addition
(Geng, Harnly, & Chen, 2016)	N.A.	N.A.	PCA: 2PCs 90,2% variance	yes	1/7 out cross-validation	N.A.	For every sample flour, bread crumb and bread crust were analyzed	N.A.	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Righetti, et al., 2016)	Double extraction of 8 samples (10% of the samples set)	yes	PCA: 2PCs 50% ESI+ and 47.2% ESI- explained variance	yes	Leave one-third out cross-validation	N.A.	Three varieties, two growing locations, two farming conditions, two hatvest years.	N.A.	N.A.	N.A.	N.A.	N.A.
(Matthews, et al., 2012)	Triplicate extractions	N.A.	N.A.	yes	Seven fold cross-validation	N.A.	45 varieties from common and durum wheat	Wheat seed provided from Colorado State University	N.A.	N.A.	N.A.	N.A.

Reference	Extraction replicates	QC	Performance	Marker identification	Validation	ROC curve	Representitive (provenance/verification of state)	Reference samples	Test samples	Validation set	Confirmed by ref/secondary method	Mixed adulterants
(Righetti, et al., 2018)	N.A.	yes	PCA: 2PCs > 52% variance for both ESI+ and ESI- models	yes	Markers confirmed with a new metabolomic experiment. Analysis of mixtures in duplicate	yes	2 locations, 2 agricultural conditions, 2 harvest years	Samples from Odisseo (durum, n=26) and Blasco (common, n=26) wheat lines	Samples from five varieties of durum (Triticum durum Desf.) and common wheat (Triticum aestivum L.), grown with conventional and organic farming	173	N.A.	1%, 2%, 3%, 5%, 10%, 15% addition of common wheat to durum wheat
(Rubert, Hurkova, Stranska, & Hajslova, 2017)	N.A.	yes	PCA: 2PCs > 58% variance for both ESI+ and ESI- models	yes	1/7 out cross-validation	N.A.	Valencian Tiger nuts and African tiger nuts from Niger, Burkina Faso and Mali	All the Valencian tiger nuts were PDO samples	N.A.	N.A.	N.A.	N.A.

N.A.: information not available in the paper

Table 3: Summary of the validation parameters verified in each of the articles selected for this review

Bibliographic references

Alewijn, M., van der Voet, H., & van Ruth, S. (2016). Validation of multivariate classification methods using analytical fingerprints – concept and case study on organic feed for laying hens. *Journal of Food Composition and Analysis, 51*, 15-23.

Aliakbarzadeh, G., Sereshti, H., & Parastar, H. (2016). Pattern recognition analysis of chromatographic fingerprints of Crocus sativus L. secondary metabolites towards source identification and quality control. *Analytical and Bioanalytical Chemistry, 408*(12), 3295-3307.

Balog, J., Perenyi, D., Guallar-Hoyas, C., Egri, A., Pringle, S., Stead, S., Chevallier, O.P., Elliott C.T., &Takats, Z. (2016). Identification of the Species of Origin for Meat Products by Rapid Evaporative Ionization Mass Spectrometry. *Journal of Agricultural and Food Chemistry, 64*(23), 4793-4800.

Bevilacqua , M., Bro , R., Marini, F., Rinnan , A., Rasmussen, M., & Skov, T. (2017). Recent chemometrics advances for foodomics. *Trends in Analytical Chemistry, 96*, 42-51.

Biancolillo, A., Bucci, R., Magrì, A., Magrì, A., & Marini, F. (2014). Data-fusion for multiplatform characterization of an Italian craft beer aimed at its authentication. *Analytica Chimica Acta, 820*, 23-31.

Black, C., Chevallier, O., & Elliott, C. (2016). The current and potential applications of Ambient Mass Spectrometry in detecting food fraud. *Trends in Analytical Chemistry, 82*, 268-278.

Black, C., Haughey, S., Chevalier, O., Galvin-King, P., & Elliott, C. (2016). A comprehensive strategy to detect the fraudulent adulteration of herbs: The oregano approach. *Food Chemistry , 210*, 551-557.

Blaise, B. (2013). Data-Driven Sample Size Determination for Metabolic Phenotyping Studies. *Analytical Chemistry, 85*(19), 8943-8950.

Blaise, B., Correia, G., Tin, A., Young, J., Vergnaud, A., Lewis, M., Pearce, J.T.M., Elliott, P., Nicholson, J.K., Holmes, E., & Ebbels, T. (2016). Power Analysis and Sample Size Determination in Metabolic Phenotyping. *Analytical Chemistry, 88*(10), 5179-5188.

Boysen, A., Heal, K., Carlson, L., & Ingalls, A. (2018). Best-Matched Internal Standard Normalization in Liquid Chromatography–Mass Spectrometry Metabolomics Applied to Environmental Samples. *Analytical Chemistry, 90*(2), 1363-1369.

Cajka, T., Danhelova, H., Zachariasova, M., Riddellova, K., & Hajslova, J. (2013). Application of direct analysis in real time–mass spectrometry (DART–MS) in chicken meat metabolomics aiming at retrospective control of feed fraud. *Metabolomics, 9*, 545-557.

Cajka, T., Smilowitz, J., & Fiehn, O. (2017). Validating quantitative untargeted lipidomics across nine liquid chromatography–high-resolution mass spectrometry platforms. *Analytical Chemistry, 89(22)*, 12360-12368.

Collins, T., Zweigenbaum, J., & Ebeler, S. (2014). Profiling of nonvolatiles in whiskeys using ultra high pressure liquid chromatography quadrupole time-of-flight mass spectrometry (UHPLC-QTOF MS). *Food Chemistry, 163*, 186-196.

Cubero-Leon, E., De Rudder, O., & Maquet, A. (2018). Metabolomics for organic food authentication: Results from a long-term field study in carrots. *Food Chemistry, 239*, 760-770.

Di Girolamo, F., Masotti, A., Lante, I., Scapaticci, M., Calvano, C., Zambonin, C., Muraca, M., & Putignani, L. (2015). A Simple and Effective Mass Spectrometric Approach to Identify the Adulteration of the Mediterranean Diet Component Extra-Virgin Olive Oil with Corn Oil. *International Journal of Molecular Sciences, 16*, 20896-20912.

Figueira, J., Câmara, H., Pereira, J., & Câmara, J. (2014). Evaluation of volatile metabolites as markers in Lycopersicon esculentum L. cultivars discrimination by multivariate analysis of headspace solid phase microextraction and mass spectrometry data. *Food Chemistry, 145*, 653-663.

Food and Drug Administration. (2001). Guidance for Industry: Bioanalytical Method Validation. *US Department of Health and Human Services, FDA, Center For Drug Evaluation and Research*.

Franceschi, P., Vrhovsek, U., Mattivi, F., & Wehrens, R. (2012). Metabolic Biomarker Identification with Few Samples. In K. Varmuza, *Chemometrics in Practical Applications* (p. 141-156). London: IntechOpen

Galle, S., Koot, A., Soukoulis, C., Cappellin, L., Biasioli, F., Alewijn, M., & van Ruth, S. (2011). Typicality and Geographical Origin Markers of Protected Origin Cheese from The Netherlands Revealed by PTR-MS. *Journal of Agricultural and Food Chemistry, 59*(6), 2554-2563.

Garcia, J., Vaz, B., Corilo, Y., Ramires, C., Saraiva, S., Sanvido, G., Schmidt, E.M., Maia, D.R.J., Cosso, R.G., Zacca, J.J., & Eberlin, M. (2013). Whisky analysis by electrospray ionization-Fourier transform mass spectrometry. *Food Research International, 51*(1), 98-106.

Garrett, R., Schmidt, E., Pereira, L., Kitzberger, C., Scholz, M., Eberlin, M., & Rezende, C. (2013). Discrimination of arabica coffee cultivars by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and chemometrics. *LWT - Food Science and Technology, 50*, 496-502.

Geng, P., Harnly, J., & Chen, P. (2016). Differentiation of bread made with whole grain and refined wheat (T. aestivum) flour using LC/MS-based chromatographic fingerprinting and chemometric approaches . *Journal of Food Composition and Analysis, 47*, 92-100.

Gil-Solsona, R., Raro, M., Sales, C., Lacalle, L., Díaz, R., Ibáñez, M., Beltran, J., Sancho, J.V., &Hernández, F. (2016). Metabolomic approach for Extra virgin olive oil origin discrimination making use of ultra-high performance liquid chromatography –Quadrupole time-of-flight mass spectrometry. *Food Control, 70*, 350-359.

Godzien, J., Alonso-Herranz, V., Barbas, C., & Armitage, E. (2015). Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample. *Metabolomics, 11*(3), 518-528.

Granato, D., Koot, A., & van Ruth, S. (2015). Geographical provenancing of purple grape juices from different farming systems by proton transfer reaction mass spectrometry using supervised statistical techniques. *Journal of the science of food and agriculture, 95*(13), 2668-2677.

Guerreiro , T., de Oliveira , D., Ferreira , M., & Catharino , R. (2014). High-throughput analysis by SP-LDI-MS for fast identification of adulterations in commercial balsamic vinegars. *Analytica Chimica Acta, 838*, 86-92.

Guijarro-Díez, M., Nozal, L., Marina, M., & Crego, A. (2015). Metabolomic fingerprinting of saffron by LC/MS: novel authenticity markers. *Analytical and Bioanalytical Chemistry, 407*(23), 7197-7213.

Hrbek, V., Krtkova, V., Rubert, J., Chmelarova, H., Demnerova, K., Ovesna, J., & Hajslova, J. (2017). Metabolomic Strategies Based on High-Resolution Mass Spectrometry as a Tool for Recognition of GMO (MON 89788 Variety) and Non-GMO Soybean: a Critical Assessment of Two Complementary Methods. *Food Analytical Methods, 10*(11), 3723-3737.

Hurkova, K., Rubert, J., Stranska-Zachariasova, M., & Hajslova, J. (2017). Strategies to Document Adulteration of Food Supplement Based on Sea Buckthorn Oil: a Case Study. *Food Analytical Methods, 10*, 1317-1327.

Jandrić, J., Haughey, S., Frew, R., McComb, K., Galvin-King, P., Elliott, C., & Cannavan, A. (2015). Discrimination of honey of different floral origins by a combination of various chemical parameters. *Food Chemistry, 189*, 52-59.

Jandric, Z., Islam, M., Singh, D., & Cannavan, A. (2017). Authentication of Indian citrus fruit/fruit juices by untargeted and targeted metabolomics. *Food Control, 72*, 181-188.

Jandric, Z., Roberts, D., Rathor, M., Abrahim, A., Islam, M., & Cannavan, A. (2014). Assessment of fruit juice authenticity using UPLC–QToF MS: A metabolomics approach. *Food Chemistry, 148*, 7-17.

Kale, N., Haug, K., Conesa, P., Jayseelan, K., Moreno, P., Rocca-Serra, P., Nainala, V.C., Spicer, R.A., Williams, M., Li, X., Salek, R.M., Griffin, J.L., & Steinbeck , C. (2016). MetaboLights: An Open-Access Database Repository for Metabolomics Data. *Current Protocols in Bioinformatics, 53*, 14.13.1-14.13.18.

Kalogiouri, N., Alygizakis, N., Aalizadeh, R., & Thomaidis, N. (2016). Olive oil authenticity studies by target and nontarget LC–QTOF-MS combined with advanced chemometric techniques. *Analytical and Bioanalytical Chemistry, 408*, 7955-7970.

Kaufmann, A. (2012). The current role of high-resolution mass spectrometry in food analysis. *Analytical and Bioanalytical Chemistry, 403*(5), 1233-1249.

Kaufmann, A. (2014). Combining UHPLC and high-resolution MS: A viable approach for the analysis of complex samples? *Trends in Analytical Chemistry, 63*, 113-128.

Kaufmann, A., Butcher, P., Maden, K., Walker, S., & Widmer, M. (2015). Reliability of veterinary drug residue confirmation: High resolution mass spectrometry versus tandem mass spectrometry. *Analytica Chimica Acta, 856*, 54-67.

Kew, W., Goodall, I., Clarke, D., & Uhrín, D. (2017). Chemical Diversity and Complexity of Scotch Whisky as Revealed by High-Resolution Mass Spectrometry. *Journal of The American Society for Mass Spectrometry, 28*(1), 200-213.

Khakimov, B., Gürdeniz, G., & Engelsen, S. (2015). Trends in the application of chemometrics to foodomics studies. *Acta Alimentaria, 44*(1), 4-31.

Khamis, M., Adamko, D., & El-Aneed, A. (2017). Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrometry Reviews, 36*(2), 115-134.

Kind, T., & Fiehn, O. (2007). Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics, 8*, 105.

Kreidler, S., Muller, K., Grunwald, G., Ringham, B., Coker-Dukowitz, , Z., Sakhadeo, U., Baron, A.E., & Glueck, D. (2013). GLIMMPSE: Online Power Computation for Linear Models with and without a Baseline Covariate. *Journal of Statistical Software, 54*(10), i10.

Kus, P., & van Ruth, S. (2015). Discrimination of Polish unifloral honeys using overall PTR-MS and HPLC fingerprints combined with chemometrics. *LWT - Food Science and Technology, 62*(1), 69-75.

Li, Y., Jin, Y., Yang, S., Zhang, W., Zhang, J., Zhao, W., Chen, L., Wen, Y., Zhang, Y., Lu, K., Zhang, Y., Zhou, J., &Yang, S. (2017). Strategy for comparative untargeted metabolomics reveals honey markers of different floral and geographic origins using ultrahigh-performance liquid chromatography-hybrid quadrupole-orbitrap mass spectrometry. *Journal of Chromatography A, 1499*, 78-89.

Martin, J., Maillot, M., Mazerolles, G., Verdu, A., Lyan, B., Migne´, C., . . . Rutledge, D. (2015). Can we trust untargeted metabolomics? Results of the metabo-ring initiative, a large-scale, multi-instrument inter-laboratory study. Metabolomics, 11(4), 807-821.

Mastrangelo, A., Ferrarini, A., Rey-Stolle, F., García, A., & Barbas, C. (2015). From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. Analytica Chimica Acta, 900, 21-35

Matthews, S., Santra, M., Mensack, M., Wolfe, P., Byrne, P., & Thompson, H. (2012). Metabolite Profiling of a Diverse Collection of Wheat Lines Using Ultraperformance Liquid Chromatography Coupled with Time-of-Flight Mass Spectrometry. *PLoS ONE, 7*(8), e44179.

Montowska, M., Alexander, M., Tucker, G., & Barrett, D. (2015). Authentication of processed meat products by peptidomic analysis using rapid ambient mass spectrometry. *Food Chemistry, 187*, 297-304.

Novotná, H., Kmiecik, O., Gałązka, M., Krtková, V., Hurajová, A., Schulzová, V., Hallmann, E., Rembiałkowska, E.,& Hajšlová, J. (2012). Metabolomic fingerprinting employing DART-TOFMS peppers from organic and conventional farming. *Food Additives & Contaminants: Part A, 29*(9), 1335-1346.

Oliveira, G., Alewijn, M., Boerrigter-Eenling, R., & van Ruth, S. (2015). Compositional Signatures of Conventional, Free Range, and Organic Pork Meat Using Fingerprint Techniques. *Foods, 4*(3), 359-375.

Özdestan, Ö., van Ruth, S., Alewijn, M., Koot, A., Romano, A., Cappellin, L., & Biasioli, F. (2013). Differentiation of specialty coffees by proton transfer reaction-mass spectrometry. *Food Research International, 53*, 433-439.

Pastor, K., Ačanski, M., Vujić, D., Bekavac, G., Milovac, S., & Kravić, S. (2016). Rapid Method for Small Grain and Corn Flour Authentication Using GC/EI–MS and Multivariate Analysis. *Food Analytical Methods , 9*, 443-450.

Popping, B., De Dominicis, E., Dante, M., & Nocetti, M. (2017). Identification of the Geographic Origin of Parmigiano Reggiano (P.D.O.) Cheeses Deploying Non-Targeted Mass Spectrometry and Chemometrics. *Foods, 6*(13).

Prandi , B., Bencivenni , M., Tedeschi, T., Marchelli , R., Dossena, A., Galaverna, G., & Sforza, S. (2012). Common wheat determination in durum wheat samples through LC/MS analysis of gluten peptides. *Analytical and Bioanalytical Chemistry, 403*, 2909-2914.

Prandi, B., Lambertini, F., Faccini, A., Suman, M., Leporati, A., Tedeschi, T., & Sforza, S. (2017). Mass spectrometry quantification of beef and pork meat in highly processed food: Application on Bolognese sauce. *Food Control, 74*, 61-69.

Riccadonna, S., & Franceschi, P. (2018). Data Treatment for LC-MS Untargeted Analysis. In G. Theodoridis, H. Gika, & I. Wilson, *Metabolic Profiling - Methods in Molecular Biology* (Vol. 1738, p. 27-39). New York: Humana Press

Riedl, J., Esslinger, S., & Fauhl-Hassek, C. (2015). Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Analytica Chimica Acta, 885*, 17-32.

Righetti, L., Rubert, J., Galaverna, G., Folloni, S., Ranieri, R., Stranska-Zachariasova, M., Hajslova, J., & Dall'Asta, C. (2016). Characterization and Discrimination of Ancient Grains: A Metabolomics Approach. *International Journal of Molecular Sciences, 17*(1217), 1-14.

Righetti, L., Rubert, J., Galaverna, G., Hurkova, K., Dall'Asta, C., Hajslova, J., & Stranska-Zachariasova, M. (2018). A novel approach based on untargeted lipidomics reveals differences in the lipid pattern among durum and common wheat. *Food Chemistry, 240*, 775-783.

Rubert, J., Hurkova, K., Stranska, M., & Hajslova, J. (2017). Untargeted metabolomics reveals links between Tiger nut (Cyperus esculentus L.) and its geographical origin by metabolome changes associated with membrane lipids. *Food Additives & Contaminants: Part A, 35(4)*, 605-613.

Rubert, J., Lacina, O., Fauhl-Hassek, C., & Hajslova, J. (2014). Metabolic fingerprinting based on high-resolution tandem mass spectrometry: a reliable tool for wine authentication? *Analytical and Bioanalytical Chemistry, 406*(27), 6791-6803.

Rubert, J., Lacina, O., Zachariasova, M., & Hajslova, J. (2016). Saffron authentication based on liquid chromatography high resolution tandem mass spectrometry and multivariate data analysis. *Food Chemistry, 204*, 201-209.

Rubert, J., Zachariasova, M., & Hajslova, J. (2015). Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review. *Food Additives & Contaminants: Part A, 32*(10), 1685-1708.

Ruiz-Samblás, C., Tres, A., Koot, A., van Ruth, S., González-Casado, A., & Cuadros-Rodríguez, L. (2012). Proton transfer reaction-mass spectrometry volatile organic compound fingerprinting for monovarietal extra virgin olive oil identification. *Food Chemistry, 134*, 589-596.

Sarah, S., Faradalila, W., Salwani, M., Amin, I., Karsani, S., & Sazili, A. (2016). LC–QTOF-MS identification of porcine-specific peptide in heat treated pork identifies candidate markers for meat species determination. *Food Chemistry, 199*, 157-164.

Schymanski, E., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology, 48*, 2097-2098.

Silva, P., Freitas, J., Silva, C., Perestrelo, R., Nunes, F., & Câmara, J. (2017). Establishment of authenticity and typicality of sugarcane honey based on volatile profile and multivariate analysis. *Food Control, 73*, 1176-1188.

Smilde, A., Jansen, J., Hoefsloot, H., Lamers, R., van der Greef , J., & Timmerman, M. (2005). ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. *Bioinformatics, 21*(13), 3043-3048.

Spiteri, M., Dubin, E., Cotton, J., Poirel, M., Corman, B., Jamin, E., Lees, M.,& Rutledge, D. (2016). Data fusion between high resolution (1)H-NMR and mass spectrometry: a synergetic approach to honey botanical origin characterization. *Analytical and Bioanalytical Chemistry, 408*, 4389-4401.

Springer, A., Riedl, J., Esslinger, S., Roth, T., Glomb, M., & Fauhl-Hassek, C. (2014). Validated Modeling for German White Wine Varietal Authentication Based on Headspace Solid-Phase Microextraction Online Coupled with Gas Chromatography Mass Spectrometry Fingerprinting. *Journal of Agricultural and Food Chemistry, 82*(28), 6844-6851.

Sumner, L., Amberg, A., Barrett, D., Beale, M., Beger, R., Daykin, C., . . . Viant, M. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics, 3*, 211–221. doi:10.1007/s11306-007-0082-2

Ulaszewska, M., Trost, K., Stanstrup, J., Tuohy, K., Franceschi, P., Foong-Fong Chong, M., George, T., Minihane, A.M., Lovegrove J.A., &. Mattivi, F. (2016). Urinary metabolomic profiling to identify biomarkers of a flavonoid-rich and flavonoid-poor fruits and vegetables diet in adults: the FLAVURS trial. *Metabolomics*(12), 32.

USP Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection. (2016). *Appendix XVIII*, 2053-2067.

Vaclavik, L., Hrbek, V., Cajka, T., Rohlik, B., Pipek, P., & Hajslova, J. (2011). Authentication of animal fats using direct analysis in real time (DART) ionization-mass spectrometry and chemometric tools. *Journal of Agricultural and Food Chemistry, 59*(11), 5919-5926.

Vaclavik, L., Lacina, O., Hajslova, J., & Zweigenbaum, J. (2011). The use of high performance liquid chromatography–quadrupole time-of-flight mass spectrometry coupled to advanced data mining and chemometric tools for discrimination and classification of red wines according to their variety. *Analytica Chimica Acta, 685*(1), 45-51.

Vaclavik, L., Schreiber, A., Lacina, O., Cajka, T., & Hajslova, J. (2012). Liquid chromatography–mass spectrometry-based metabolomics for authenticity assessment of fruit juices. *Metabolomics, 8*, 793-803.

von Bargen, C., Brockmeyer, J., & Humpf, H. (2014). Meat authentication: a new HPLC-MS/MS based method for the fast and sensitive detection of horse and pork in highly processed food. *Journal of Agricultuiral and Food Chemistry, 62*(39), 9428-9435.

Vuckovic, D. (2012). Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry, 403*(6), 1523-1548.

Want, E., Masson, P., Michopoulos, F., Wilson, I., Theodoridis, G., Plumb, R., Shockcor, J., Loftus, N., Holmes, E., &Nicholson, J. (2013). Global metabolic profiling of animal and human tissues via UPLC-MS. *Nature Protocols, 8*(1), 17-32.

Wielogorska, E., Chevallier, O., Black, C., Galvin-King, P., Delêtre, M., Kelleher, C., Haughey, S., & Elliott, C. (2018). Development of a comprehensive analytical platform for the detection and quantitation of food fraud using a biomarker approach. The oregano adulteration case study. *Food Chemistry, 238*, 32-39.

Wulff, T., Nielsen, M., Deelder, A., Jessen, F., & Palmblad, M. (2013). Authentication of Fish Products by Large-Scale Comparison of Tandem Mass Spectra. *Journal of Proteome Research, 12*, 5253-5259.

Xia, J., Broadhurst, D., Wilson, M., & Wishart, D. (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics, 9*(2), 280-289.

Ziółkowska, A., Wąsowicz, E., & Jeleń, H. (2016). Differentiation of wines according to grape variety and geographical origin based on volatiles profiling using SPME-MS and SPME-GC/MS methods. *Food Chemistry, 213*, 714-720.

Author Contributions

Conceptualization: D.C., L.R., M.S.; bibliographic study and writing-original draft preparation: D.C., L.R.; writing - review and editing: all authors; supervision. C.E., M.S.

Aim of the thesis

The general objective of the thesis was the development and set up of novel non-targeted methodologies devoted to verify traceability and frauds/adulterations issues on specific raw materials for both rapid and confirmatory purposes; the priorities were identified along a risk-assessment phase in collaboration with Barilla Food Anti-Fraud Team.

Simultaneously, a similar approach was also applied in the sensomic field, trying to identify some active molecules responsible of consumer overall liking evaluation on specific food products.

The specific aims of the project were:

- The development of rapid and confirmatory methods for the detection of freshness issues in egg products (Chapter 1)
- The development of rapid and confirmatory methods for the detection of soft refined oils addition to extra virgin olive oils (Chapter 2)
- The development of a method able to assess the geographical origin of durum wheat (Chapter 3)
- Apply the same approach in the sensomic field, trying to correlate sensorial evaluations with chemical parameters (Chapter 4)

CHAPTER 1

EGG PRODUCT

Egg products

Michele Suman, Daniele Cavanna, Michele Zerbini, Diego Ricchetti, Damiano Sanfelici, Elisa Cavandoli, Leonardo Mirone

Barilla G.& R. F.lli SpA, Via Mantova 166 - 43122 Parma, Italy

General overview of the product

Eggs can be considered a milestone for human and animal diets thanks to the large content of high quality proteins and vitamins. For this reason, eggs are ingredients largely used in the industrial preparation of different types of foods (foams, emulsions, pastry and bakery products) under the mixture forms which denominate egg products.

Eggs sector is one of the most important industries in the agricultural framework all over the world since the avoidance of weather factor makes it suitable for all the different climate areas.

Main producing countries are European Union, China and North-Central America. Given the perishable nature of the main product (shell eggs and liquid egg products) there are not so big flows of products around the different countries and most of the production is fully dedicated to internal markets. This may vary in case of big issues such as Aviary Influenza in US that lead the entire country to support the import of liquid products from EU. Different discussion can be made on the powder products, especially albumen with specific properties on bakery products, that can be commercialized without concerns on expiry dates, but in any case face some trade restrictions between countries (i.e. not all European country can export to U.S.).

Main supplier are entitled for a vertical integration of the different step in the supply chain: feed-mill, shelling and heating treatment plant and farms.

Increasingly importance on animal welfare has put even more under pressure the entire sector that has to adapt itself to the law requests especially in EU (2012) and also to consumer demand that will eradicate the cage household from the supply side.

In replying to this issues market reacted to tackle the increasing demand on more animal welfare suitable farms, and in addition a more transparent supply chain in order to prevent any possible other scandal that affects market in the past.

Product Identity

Definition of the product and manufacturing process

The concept of egg products is related to all the forms of presentation of the egg: yolk, albumen or a mix of them. The term "Egg Products" refers to processed or convenience forms of eggs obtained by processing shell eggs: egg products include whole eggs, egg whites, and egg yolks in frozen, pasteurized and refrigerated liquid, and dried forms available in a number of different product formulations.

In particular, industry is interested in high quality egg products in a liquefied form, obtained from eggs shelled within 4 days and subjected to homogenization and pasteurization processes: their use is mainly related to the preparation of egg pasta and bakery products (Rossi, Proprietà funzionali degli ovoprodotti, 1998).

Transformation process

After the deposition, the eggs receive a sorting in order to provide the separation of damaged, dirty, broken ones and the classification (as the dimension) according to the characteristics defined by law (A category, B category).

A classification of the eggs also in terms of freshness, meaning the time from the deposition to the transformation in egg products or the shipping to the retail market, is done. According to this classification "extra-fresh" eggs or "fresh" eggs are clearly distinguished by "conventional" eggs.

Eggs for industry use, entering the process for transformation to liquid egg products, are destined to the management by food transformation factories where from eggs in shells, they turn into pasteurized refrigerated egg products. The flow diagram of this transformation process is presented below:

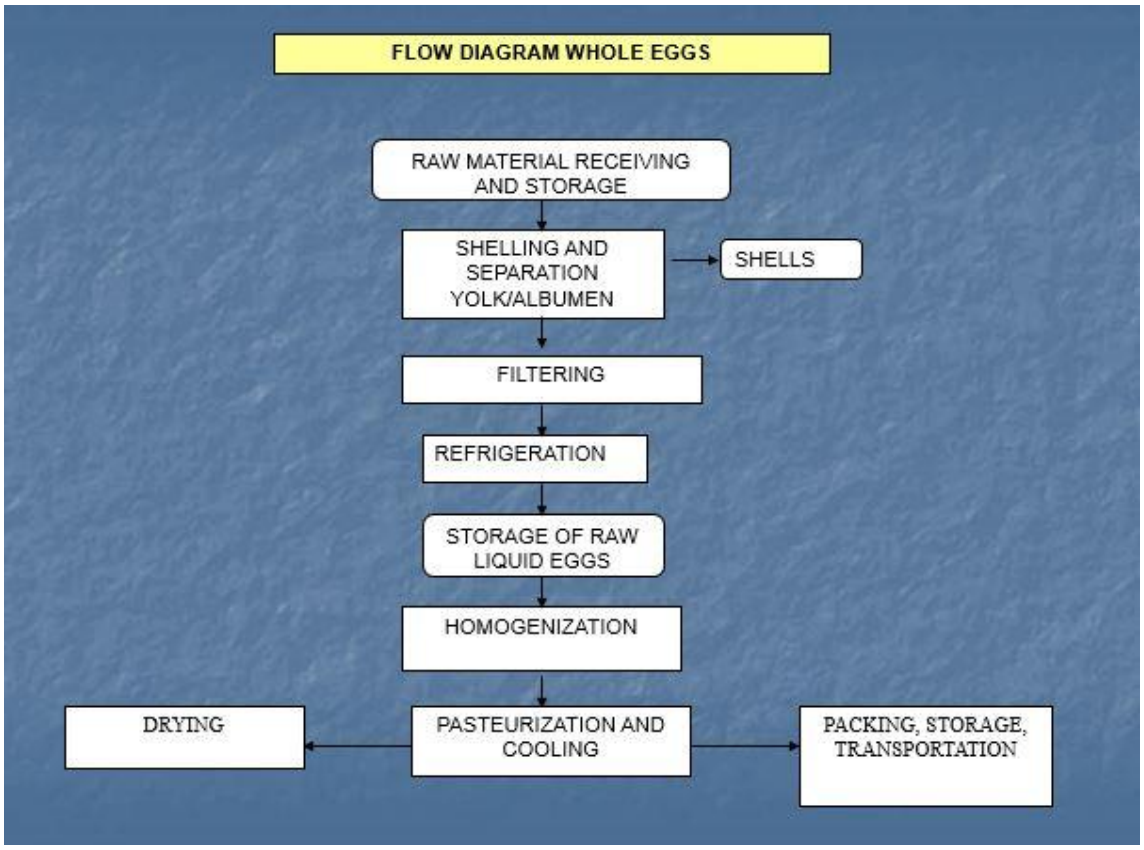


Figure 1: flow diagram of eggs transformation

Liquid egg products coming from this process, are commonly delivered to food companies in refrigerated tanks and their quality and food safety characteristics (chemical, physical parameters inserted in related technical specifications) are carefully controlled by the producers before their release and by the customers at the receiving and before the use.

Current standards of identity or related legislation (Codex, EU, US, ISO, Trade associations…)

178/2002 – This regulation lays down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety

852/2004 - This regulation and its annexes defines a set of food safety objectives that firms working with food must meet.

853/2004 - This regulation aims to ensure a high level of food safety and public health. It complements Regulation (EC) No 852/2004 on the hygiene of foodstuffs, whose rules mainly cover the approval of operators in the sector. The regulation's rules apply to unprocessed and processed products of animal origin. They generally do not apply to food that contains both products of plant origin and processed products of animal origin. European Union (EU) countries must register and, where necessary, approve establishments handling products of animal origin.

2073/2005 – This regulation concerns with microbiological criteria applied to food staff

1881/2006 – This regulation lays down the maximum limits for certain contaminants in food in particular to protect the health of the most sensitive population groups, i.e. children, the elderly and pregnant women.

589/2008 – This regulation lays down detailed rules for implementing Council Regulation (EC) No 1234/2007 as regards marketing standards for eggs:

Describe the egg characteristics for CAT A and CAT B (shell and cuticle, air space, yolk, album, germ, foreign matter and smell)

Grading Cat A by weight defining the classification in different size

Define shelf life and timing to grading, marking and packing eggs

Define how to handle the industrial eggs

Define the code to mark the eggs

Which records to be kept by producers, collectors and packing center

Checks

Non compliances and tolerances

Authenticity issues

Identification of current authenticity issues

Hereafter we will consider the pasteurized eggs used in the food preparation, starting from liquid eggs already shelled, provided by suppliers located in EU with integrated supply chain for farms, feed mill and transformation plant.

Most risky issues

The **most risky factors** that can have impact on the egg authenticity are:

- **Different hens farming with impact on animal welfare**

  At the market level there is a rising request of eggs coming from barn hens or cage free farming system, but the facilities need to be converted and it is a matter of fact a number of uncertainties about certifications that all the volumes are/will be satisfied within the animal welfare respect framework.

  Currently, there are no available analytical methods able to categorize different farming approaches (barn hens or cage free farming system) and this fact increases the fraud opportunity. In addition, eggs promiscuity is possible at farm level, during the transportation and at the transformation steps.

- **Fresh eggs**

  Eggs can be declared as fresh eggs within 28 days shelf life.

  Eggs with more than 28 days shelf life can be found on the market still declared as “fresh” in order to fraudulently exploit a higher price vs the others.

  Albumen and yolk contain enzymes, and if eggs are not stored at a sufficiently low temperature, the proteins can alter. The optimal temperature for a correct egg storage is normally about 6–8 ◦C. The enzymatic alterations of the albumen modify its viscosity, which allows to recognize the freshness: in fact when the egg is not fresh the albumen tends to liquefy and the yolk easily breaks.

  Immediately after the laying phase, the contents of the egg with its entire shell are practically sterile and can be contaminated from environmental microorganisms only if the shell is broken.

- **Eggs category**

  There is a rising request of cat A eggs in the last years, with quality parameters described by the regulation into force. The farms cannot be sufficient to satisfy the request and the Cat B eggs are cheaper than the Cat A.

  This promiscuity is possible at farm level, during the transportation and at the transformation facility.

- **Dilution with incubated eggs**

  The eggs coming from incubation process must be sent for destruction or animal feed. It is not possible to use them for human consumption. The price of these eggs is very low and can imply the illegal use in some periods (when the egg offer on the market is low or when there is much availability from the incubators). There are parameters regulated by law to avoid the usage of these eggs for human consumption.

- **Artificial colorants**

  Artificial colorants are allowed but some supply chain are free from artificial colorants.

  Eggs promiscuity is possible at farm level, feed mill and at the transformation facility.

Less risky issues

Less risky factors, but possible, are the following:

- **Usage of different animal eggs**

  The economic opportunity not always justifies this type of fraud and there are some technical and mechanical restrictions in the shelling lines.

- **Dilution with water**

  The opportunity of this type of fraud is decreased by the detectability with the current analytical methods (dry matter analyses).

Aim of the work

In the following chapters, a rapid approach and a non-targeted metabolomic method will be presented as two parallel new tools able to assess eggproducts freshness.

The first chapter was accepted for publication in "*Food Chemistry"*, while the second was accepted for publication in "*Journal of Mass Spectrometry"*. For additional details see section "Author".

Bibliographic references

Rossi, M. (1998). Proprietà funzionali degli ovoprodotti. *Rivista di Avicoltura*, 67, 28-34.

Ion Mobility Spectrometry coupled to Gas Chromatography: a rapid tool to assess eggs freshness

Daniele Cavanna[1,2], Sandro Zanardi[1], Chiara Dall'Asta[2], Michele Suman[1]

[1]Barilla Advanced Laboratory Research, Via Mantova 166 - 43122 Parma, Italy
[2]Department of Food and Drug, University of Parma, Parco Area delle Scienze 95/A- 43124 Parma, Italy

Introduction

In the last decade, consumers' demand for high quality and safer food has showed an increasing trend, in response to a number of scandals that have interested the agro-food compartment. Food industry, therefore, has enforced measures to increase its product and process transparency, also with the development of rapid but reliable methods for detecting frauds and adulterations along the production chain.

Eggs, mostly in the egg products form, are largely used for the preparation of different industrial products, and their freshness is a crucial parameter for the production of safe and high quality commodities. However, rapid and robust methods to assess eggs freshness are still lacking.

In the past, lactic acid (≤ 1000 mg/kg dry egg) and succinic acid (≤ 25 mg/kg dry egg) were monitored by law (D.L. no. 65, 1993), while the current EU legislation considers only lactic acid as reliable marker (Reg. CE 853/2004).

Both compounds are strongly related to the microbial spoilage of eggs. Since their increase is not linear overtime, they could act as "tardive" markers of ageing. Another limitation of lactic and succinic acid use as marker of eggs aging, is the amount of time required by official methods, based on enzymatic reaction, to get a reliable result.

Because of delay, potentially not fresh egg products batches could enter the production lines, leading to a subsequent interruption of the production or, in case the product batch will enter the market, to consumers complains and, possibly, to a product recall.

Therefore, development of rapid tools able to assess if a batch is as fresh as declared can potentially avoid the delivery of the degraded raw material to the plant, minimizing the economic loss for the industry and reducing the risk for the final consumer.

Nondestructive rapid methods, usually applied on shell eggs, have been already presented in literature. Spectroscopic approaches are largely diffused (Wang, Zhou, Wang, & Ma, 2015) (Zhang, Pan, Tu, Zhan, & Tu, 2015) (Lin, Zhao, Sun, Chen, & Zhou, 2011) (Zhao, et al., 2010), together with the S-ovalbumin evaluation (Huang, et al., 2012). Concerning rapid application for egg freshness, several studies based on e-nose have been reported for shell eggs (Yimenu, Kim, & Kim, 2017) (Li, Zhu, Jiang, & Wang, 2017) , while our research group has described an artificial olfactory system with demonstrated capabilities towards egg products (Suman, Riani, & Dalcanale, 2007).

Hyphenated GC-IMS (Gas Chromatography-Ion Mobility Spectrometry) is a new technique, rapidly growing in terms of interest for the analytical community. This instrumentation is based on a gas chromatograph coupled with an ion mobility spectrometer, able to detect the volatile fingerprints of liquid or solid samples without any relevant pretreatment.

Ion Mobility Spectrometry is an analytical technique able to separately detect gaseous compounds in a mixture of analytes. Ions have to pass through a fixed distance tube (named drift tube) along which an electric field is applied. The electric signal is recorded at the end of the tube with a faraday plate, while an inert gas is operating in counter flow in the tube in order to increase ions differentiation.

Molecules are separated according to the time that they take to pass through the tube (the drift time), according to their shape and charge distribution. Therefore, IMS can effectively separate isomeric compounds as well (Eiceman, Karpas, & Hill, Jr, 2016).

The selectivity of this approach can increase if it is coupled with orthogonal separation techniques (i.e. Liquid or Gas Chromatography) (Gerhardt, Birkenmeier, Sanders, Rohn, & Weller, 2017).

When coupled to a gas chromatograph, everything that elute from the column is ionized and then pass through the ion mobility system, allowing the recording of a 3D graph in which each compound is characterized by a retention time, a drift time and an intensity value. (Garrido-Delgado R. , Dobao-Prieto, Arce, & Valcárcel, 2015).

This technique was recently used for the detection of sesame oil adulteration (Zhang, et al., 2016) ,for extra virgin olive oils profiling (Gerhardt, Birkenmeier, Sanders, Rohn, & Weller, 2017), and for the identification of the best packaging conditions for this raw material (Garrido-Delgado R. , et al., 2015). In addition, GC-IMS is presented in literature as a very useful tool for a rapid discrimination through extra virgin, virgin and lampante olive oils (Garrido-Delgado R. , Dobao-Prieto, Arce, & Valcárcel, 2015) (Garrido-Delgado, Arce, & Valcárcel , 2012) .

Based on what described above, the aim of the present work is the development of a rapid, low cost and low sample demanding method able to assess egg products freshness, using GC-IMS combined with chemometric data elaboration.

Different chemometric models were applied in order to classify the samples, and an attempt of markers identification was performed through the exploitation of the SPME-GC-MS (Solid Phase Micro Extraction / Gas Chromatography-Mass Spectrometry) technique.

Materials and Methods

Chemicals and samples description

All reagents were purchased from VWR International, Ltd (Poole, United Kingdom).

Analytical standards of Butyl Acetate, 1-Butanol, Heptanal, Dimethyl disulfide and Dimethyl trisulfide were purchased from Sigma Aldrich (St. Louis, MO).

Water was purified using a Milli-Q system (Millipore, Bedford, MA).

Egg products samples were directly collected from batches used in routinely Italian egg pasta production plant flows; aiming to increase the variability of the model, the sampling period lasted six weeks and samples coming from different batches, different suppliers and with different amounts of carotenoids (in the range of tenths of mg/kg, authorized for improving appearance and correspondent consumer preference in finished products) were selected.

Experimental Design

The basic idea of this study was to create a model able to discriminate between the fresh egg products and the not fresh ones, recording the volatile fingerprints of the samples. In order to include the highest amount of degradation products, the model was created with extreme storage conditions, so the products were left at room temperature for few days.

As first step, a predictive model was built up, able to assess egg products freshness. Subsequently, as recently suggested also from US Pharmacopoeia (U.S. Pharmacopeial Convention, 2016), an external set of samples was used to validate the predictive capabilities of the model: it was then challenged with different mixtures of fresh and not fresh egg products.

For the model set up, 52 different batches were collected and analysed immediately after ($t = 0$). Batches were kept at room temperature, and randomly analysed according to the following scheme: 29 batches after 1 day ($t = 1$), 21 batches after 2 days ($t = 2$), 15 batches after 3 days ($t = 3$), 11 batches after 4 days ($t = 4$), and 9 batches after 5 days ($t = 5$). Some batches were randomly sampled in duplicate at each time point, for a total sample size of 132 samples.

For the external validation, 7 new egg product batches were collected immediately after the delivery to the plant, and then left at room temperature up to 5 days after delivery. All the batches were sampled and analysed at $t = 0$, $t = 1$ and $t = 2$. Then, a random sampling was performed at $t = 3$, $t = 4$, and $t = 5$. In total, $n = 30$ samples were used for model validation.

The full sampling plan is reported in details in the Supplementary Material (Section 1, Table S1 and S2).

Finally, the predictive capabilities of the model were challenged by analyzing 8 different admixtures of fresh egg products, obtained by adding to fresh material different percentage of aged samples stored at different conditions. The mixtures percentage and the type of additions are detailed in Table 1. These samples were prepared using also aged egg products stored at 2-8 °C for including another more "subtile" possible type of fraudulent addition.

Experimental approach for markers identification

With the aim to identify some of the spots responsible of freshness issues, the egg products used to build up the predictive model underwent SPME-GC-MS (Solid Phase Microextraction-Gas Chromatography-Mass Spectrometry) analysis as well, following the same experimental design.

Raw files were compared using the fresh samples as reference, and the mass spectra of the peaks highlighted as potential aging markers were inspected. After a first attempt of compound identification, based on NIST libraries, a further selection was performed including only those compounds probably related to freshness issues.

Reference standards were available for five of the compounds identified with this approach: Butyl Acetate, 1-Butanol, Heptanal, Dimethyl disulfide and Dimethyl trisulfide. These were separately spiked into fresh egg product samples and analyzed with the GC-IMS system, for finding a correlation between the marker spots and the target compounds.

Sample preparation

For the GC-IMS analysis, 15 drops of sample were directly transferred in a 20 mL headspace vial that was subsequently incubated at 40 °C for 20 minutes.

For the evaluation of method reproducibility, 20% of the samples were double prepared and injected.

For the GC-MS analysis, 2 grams of egg product were directly transferred in a 20 mL headspace vial that was subsequently incubated for 60 minutes at 40 °C.

GC-IMS analysis parameter

The GC-IMS instrument (Flavourspec® - G.A.S. Dortmund Company – Germany) was equipped with a syringe and an autosampler unit for headspace analysis. The injection volume was set to 0.5 mL, with an injection speed of 30 mL min^{-1} and a syringe temperature of 80°C; the injector temperature was kept at 80 °C.

The chromatographic separation was performed on a SE-54-CB-1 (5%phenyl- 1%vinyl- 94% methylpolysiloxane) capillary column (15 m x 0.53 mm , 1 μm film thickness) kept at 40°C, with Nitrogen as carrier gas.

The chromatographic separation was executed under isothermal conditions, employing a carrier gas flow ramp starting at 2 mL min^{-1} for 5 minutes , then increasing to 70 mL min^{-1} in 20 min and finally staying 5 minutes at 70 mL min^{-1}. The total GC runtime was 30 minutes.

What eluted from the chromatographic column was transferred in the ion mobility system (see Supplementary Material, Section 2, for instrumental details).

GC-MS analysis parameter

The GC-MS instrument was a Trace GC Ultra (Thermo Fisher Scientific, Inc., Waltham, MA) coupled with a DSQ II single quadrupole (Thermo Fisher Scientific, Inc., Waltham, MA); instrumental analyses were partially performed according to Costanzo et al (Costanzo, et al., 2016).

At the end of the incubation time, using a TriPlus autosampler, a preconditioned SPME fiber (at 250 °C) 50/30 μm Divinylbenzene/Carboxen/Polydimethylsiloxane (Supelco- Sigma Aldrich) was exposed to the headspace of the vial for 3 hours.

Subsequently, the injection in the GC systems was executed in splitless mode for 5 minutes.

The separation was performed with a Rxi -5ms (30m x 0.25 mm, 0.5 μm film thickness) capillary column (Restek) kept at 40 °C.

A temperature gradient was created for a better compounds separation: after 8 minutes at 40°C, temperature increased to 60 °C at 4°C min^{-1}, then to 160 °C at 6°C min^{-1} and finally to 200 °C at 20°C min^{-1}. Oven was kept at this temperature for 5 minutes before ending the cycle.

Data treatment

GC-IMS analysis resulted in a 3D graph in which each analyte is represented by a point (spot) that is characterized by the retention time (measured in seconds, on the *y* axis), the drift time (measured in milliseconds, on the *x* axis) and the intensity of the signal (Gerhardt, Birkenmeier, Sanders, Rohn, & Weller, 2017) (Figure 1).

An area set able to integrate all the markers spots was created using the LAV software (version 2.2.1 - G.A.S. Dortmund company). It was applied for the analysis of all the samples considered within this study, for model set up, model validation and admixture experiments. Data considered for the area set were reported in Table 3.

For the predictive model set up, the acquired spectra were aligned according to the "Reaction Ion Peak" (RIP) position (as reported in Figure 1A): this signal is generated by the total of all reactant ions available in the ionization chamber.

Area values of the integrated spectra were exported in an Excel spreadsheet, and processed with SIMCA (version 14.1 Umetrics, Umea, Sweden) software for chemometric elaboration. A global overview of the spots identified in egg products is reported in Figure 1B.

Firstly, data were log transformed and UV scaled, then a preliminary Principal Component Analysis (PCA) was performed with all the samples (including the extraction replicates) and subsequently, a new PCA model was created with the average area values of the replicates, as reported in Figure 2A (for details, see Supplementary Material, section 3).

On the basis of the results obtained, samples were divided into two groups: "fresh samples" (t = 0) and "not fresh samples" (from t = 1 to t = 5). A supervised orthogonal partial least square discriminant analysis (OPLS-DA) model was then built up to maximize groups separation (see Figure 2B).

The goodness of the chemometric models was firstly evaluated with an internal 1/7 out cross-validation. The supervised model was then validated with the external set of samples and its prediction capabilities were also challenged with the analysis of the mixtures previously described.

For the GC-MS compounds characterization, raw data were acquired using Xcalibur software (verison 1.4 SR1- Thermo Fisher Scientific, Inc.); the same software was used for data elaboration (qualitative analysis) in combination with SIEVE (version 2.0 - Thermo Fisher Scientific, Inc.) for peak alignment and features extraction. The fresh samples were used as reference group in order to detect all the peaks that appeared in the other time points and that were probably related to eggs ageing.

Mass spectra of the selected compounds were compared with NIST libraries.

Results

Model set up

The preliminary PCA generated with all the samples (including replicates) showed a great reproducibility of the analysis with a perfect overlay for almost all the replicates in the Score Plot.

Fresh samples are clearly clusterized and separated from the other time points. The first two Principal Components described more than 71% of the explained variance and the global goodness of fit value (R^2X) was 0.764. The excellent quality of the model was also confirmed by the value of its prediction ability parameter (Q^2 = 0.679) evaluated using the internal cross-validation.

On the basis of these results, an OPLS-DA supervised model was generated, grouping the samples as "fresh" (0 hours) and "not fresh" (all the others).

Figure 2B shows the clear clusterization of the two groups obtained with this model; moreover, the global evaluation parameters were excellent, with an R^2X value of 0.824, an R^2Y value of 0.950 and a Q^2 value of 0.920.

Model validation

For model validation, n = 30 indipendent samples were analyzed and their affinity to the "fresh" or to the "not fresh" group was predicted with the OPLS-DA model (for details, see Supplementary Materials, section 3).

The model shows a great prediction ability (97%), and all the samples were correctly classified, with the exception of the "Val2_1D" sample.

Admixture experiments

As for the external validation set of samples, the positioning of the 8 admixture samples within the two clusters, was predicted with the OPLS-DA model , as shown in Figure 2C (for details, see Supplementary Materials, section 3).

None of the mixtures showed an affinity percentage for the "fresh group" higher than the software threshold for a certain assignment of a sample to that class (70%); at least they are considered as "suspect" or directly as "not fresh". This certifies a good capability of the method to detect fraudulent mixtures of products, also at low percentage (as for example the "mix_7" sample).

Markers identification

After the SPME-GC-MS experiments, a tentative of compounds identification was performed for a group of chromatographic preaks connected to eggs ageing.

The separate reference standards addition to fresh egg products of some of these markers, and the subsequent analysis with the GC-IMS instrument, allowed to find a unique correlation with specific spots.

The obtained results are presented in Figure 3. In particular, peak 28 and peak 31 were identified as butyl acetate and heptanal, respectively. Both Dimethyl disulfide and Dimethyl trisulfide concurred to peak 1, while 1-Butanol was responsible for peak 15 and peak 17. The reason why this last compound generated two spots is still unknown but certainly both the signals are related to 1-Butanol as they appeared only when the specific reference standard was injected.

Discussion

The main target to address in this study, was the development of a rapid tool able to return a quick but reliable answer about egg product freshness, without causing delays in the industrial workflow. One of the major point to consider, was the accuracy in terms of false negative misclassification.

Th use of GC-IMS returned fully adequate results, in terms of time required from sampling to results communication (about 40 minutes), robustness and stability.

The procedure was developed considering the broadest possible variability in terms of batches, delivery date, suppliers etc; with more than 50 batches and more than 130 samples used for model build up.

The stability over time was successfully evaluated by collecting samples over six weeks. Every day a group of fresh and not fresh samples were analyzed. In spite of the long period of observation, no trend related to the day of the analysis was detected; in addition, a satisfactory overlay for almost all the replicates in the Score Plot attested the great repeatability of the instrument.

The PCA clearly shows a separation between fresh and not fresh samples. In chemometric studies, this unsupervised first approach is of utmost relevance, because it is based only on differences between samples without the need of any supervised information for driving the clusterization. This means that the variables selected for the model creation (in the present work represented by the intensity of the spots) can be regarded as discriminative. *(*Eriksson, et al., 2006*)*

This outcome actually ensures that the subsequent creation of a supervised model brings to a maximization of real differences through the samples and not to a "false" differentiation due to possible overfitting of the model.

We have therefore moved from unsupervised PCA to supervised OPLS-DA analysis.

As previously mentioned, the external validation is crucial to avoid in any case possible overfitting of the model: the correct prediction of the 97% of the validation set clearly demonstrate that the clusterization according to the time points is reliable, robust and not related only to specific analytical or statistical conditions (i.e. analytical sequence, sampling time etc).

This method was generated storing the egg products in extreme conditions (room temperature) with the aim to create an area set as large as possible. This will ensure that the majority of the degradation products could be highlighted and included. It is unlikely that fraudulent actions lead to the delivery of a completely degraded egg product batch. More likely, mixtures of fresh and not fresh batches could be used.

For this reason, from a food industry perspective the ability of our method to correctly classify as "suspect" admixtures containing low amount of aged eggs, can be regarded as fully satisfactory.

The marker compounds identified with the help of the GC-MS instrument are coherently related with thermal degradation processes, as presented in literature for many others commodities (Dragone, Mussatto, Oliveira, & Teixeira, 2009) (Friedrich & Acree, 1998) (Giri, Osako, & Ohshima, 2010). Due to their similar structure, the sulfur compounds are not separated. However, we can assume that the increase in «peak 1» intensity can be considered a marker of not freshness.

Conclusions

The present study describes the development and application of an untargeted GC-IMS fingerprinting for the assessment of egg products freshness. The method presented herein was demonstrated to be an easy to use, rapid, low cost and robust analytical tool.

A chemometric model was applied in order to classify the samples as "fresh" or "not fresh" and the results obtained are excellent. The predictive capabilities were validated with an external set of samples and 97% of them were correctly classified.

Moreover, the method was also challenged with the analysis of admixture samples, prepared with fresh egg product as predominant component and variable percentage of different aged egg products as adulterants.

The application of this analytical approach in routine laboratories of the production plants could efficiently support the rapid identification of not fresh batches in shorter time compared to those required from official control methods, avoiding therefore their entrance in the manufacturing chain and ensuring thus the production of safe and high quality egg products based food. Moreover, this volatile fingerprint approach could be applied also for the detection of other frauds related to freshness issues, as for example the illegal use of incubated eggs.

Acknowledgement

Authors would like to thank Angelo Restori and Enrica Emanuelli (Barilla Quality Control Laboratory - Parma) for the invaluable support, especially in sample batches collection.

This work was partially supported by the European Union, Seventh Framework Program (FP7/2007-2013), under Grant Agreement 613688 (Ensuring the Integrity of the European food chain FoodIntegrity).

Tables

SAMPLE	% Fresh Egg	% Addition	Type of addition
Mix_1	80	20	1 day at RT
Mix_2	80	20	5 days at RT
Mix_3	50	50	1 day at RT
Mix_4	50	50	5 days at RT
Mix_5	80	20	24 days at 2-8 °C
Mix_6	90	10	25 days at 2-8 °C
Mix_7	94	6	25 days at 2-8 °C
Mix_8	90	10	2 days at RT

Table 1: Composition of admixture samples.

Name	RT min (s)	RT max (s)	Drift time (ms)	ΔDrift time (ms)
peak1	117.04	137.17	8.471	0.683
peak2	124.59	142.83	8.944	0.499
peak3	125.22	144.09	9.357	0.466
peak4	132.14	166.74	8.327	0.310
peak5	152.90	174.29	8.933	0.340
peak6	151.64	175.55	9.610	0.533
peak7	153.53	173.67	10.197	0.733
peak8	174.92	200.72	8.622	0.533
peak9	175.55	199.46	10.040	0.967
peak10	191.28	213.94	8.828	0.382
peak11	183.73	224.63	10.701	0.433
peak12	204.50	233.44	9.392	0.767
peak13	208.90	226.52	11.076	0.355
peak14	210.79	257.35	8.422	0.199
peak15	233.44	294.48	11.579	1.645
peak16	302.03	324.68	8.615	0.638
peak17	269.94	330.34	9.491	0.505
peak18	303.91	322.79	10.487	0.892
peak19	327.20	345.44	11.185	0.333
peak20	328.45	344.81	11.703	0.500
peak21	319.64	403.33	8.806	0.849
peak22	320.90	366.21	10.593	0.850

peak23	358.03	384.45	11.316	0.765
peak24	354.88	384.45	12.444	0.978
peak25	420.95	432.90	10.472	0.566
peak26	421.58	444.86	12.550	0.595
peak27	423.47	495.20	9.487	0.510
peak28	441.08	465.00	13.060	0.935
peak29	480.10	505.89	10.125	0.425
peak30	477.58	509.67	13.288	0.927
peak31	554.34	570.70	13.520	0.548
peak32	634.89	706.62	10.693	0.379
peak33	636.77	693.40	14.427	0.928
peak34	735.56	799.11	10.778	0.463
peak35	734.93	804.78	14.595	0.927

Table 2: GC-IMS global area set integration parameters.

Sample name	Affinity % with "fresh" class	Affinity % with "not fresh" class
Mix_1	57.91	42.09
Mix_2	0.00	100.00
Mix_3	0.00	100.00
Mix_4	0.00	100.00
Mix_5	38.88	61.12
Mix_6	41.45	58.55
Mix_7	38.03	61.97
Mix_8	33.87	66.13

Table 3: Predicted group affinity of the mixtures samples. According to the software rules, values higher than 70% correspond to a certain belonging of the sample to that group.

Figures

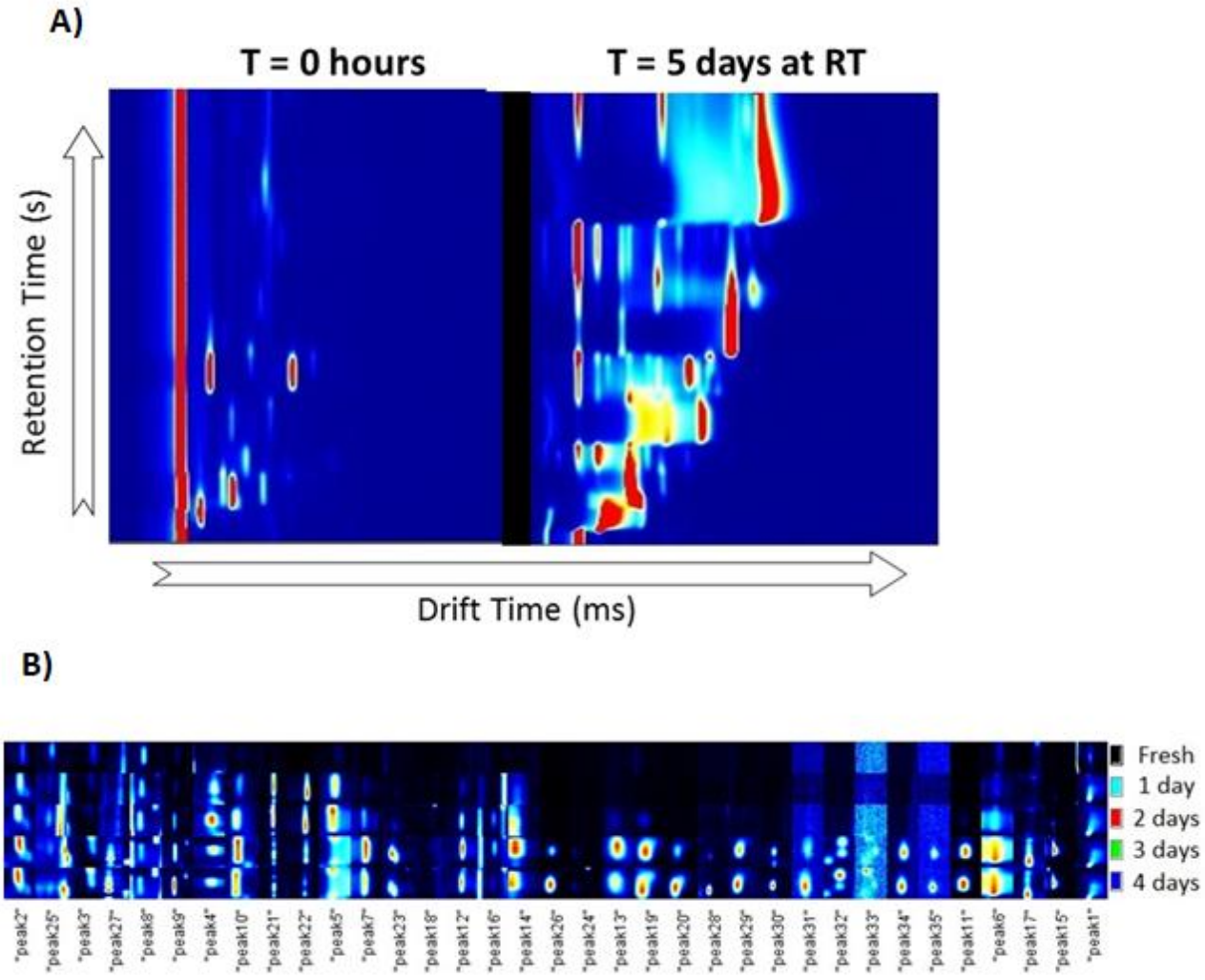


Figure 1: A) Ion Mobility spectrum of a egg product at T=0 hours (left) and after 5 days at room temperature (right). The red line identifies the Reaction Ion Peak (RIP) position. B) Global overview of the spots identified in the egg product "Sample_20" at different time points.

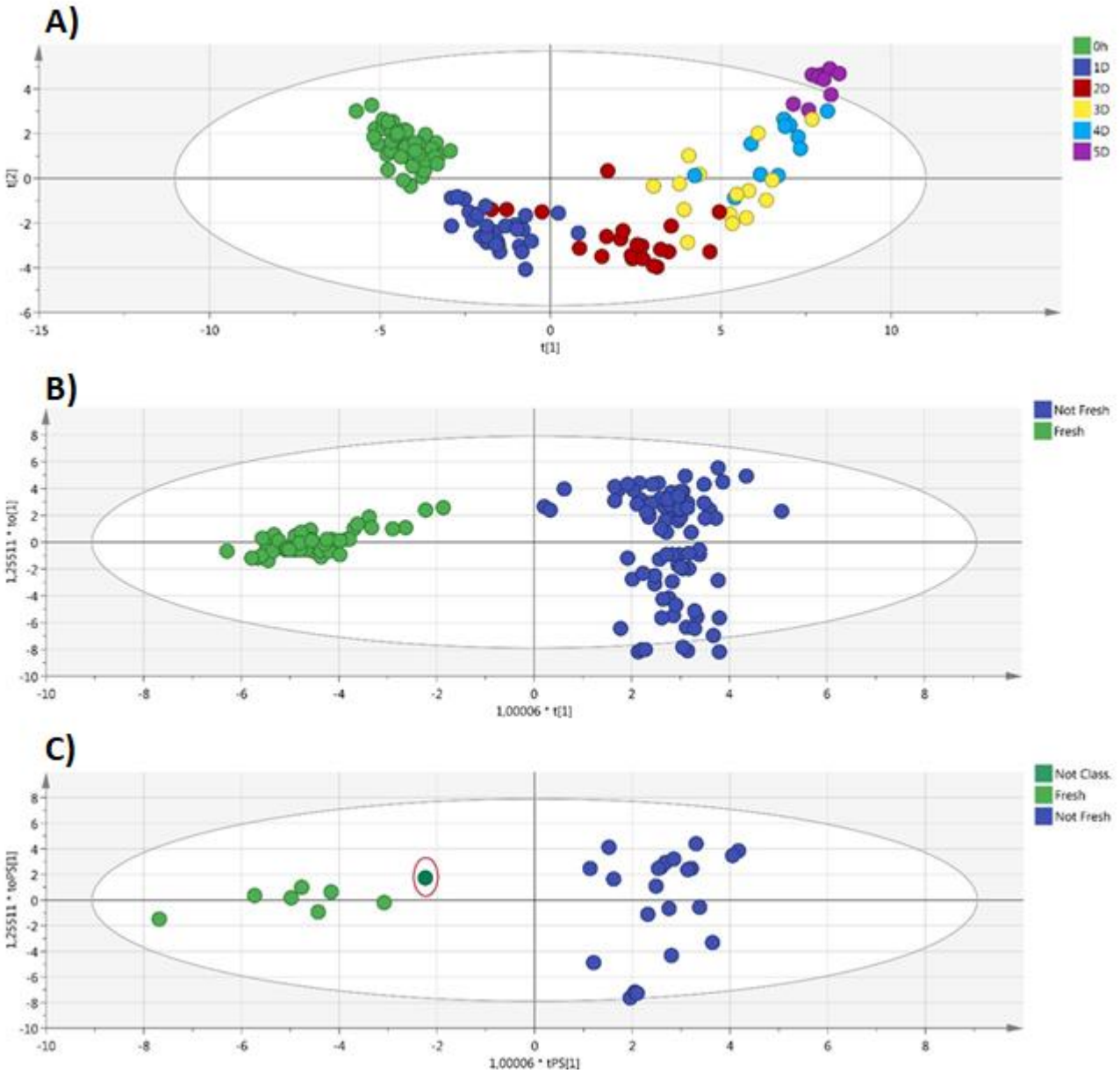


Figure 2: GC-IMS data analysis. A) PCA scores plot of the egg products samples, colored according to the time spent at room temperature. Leftmost area dots: 0 hours; 1D dots: 1 day; 2D dots: 2 days; 3D dots: 3 days; 4D dots. 4 days; 5D dots: 5 days; B) OPLS-DA scores plot of the egg products samples, colored according to the groups. Left area dots: “fresh samples”; Right area dots: “not fresh samples”; C) visual prediction of the validation samples; Left: fresh, Circled-left: not classified, Right: not fresh.

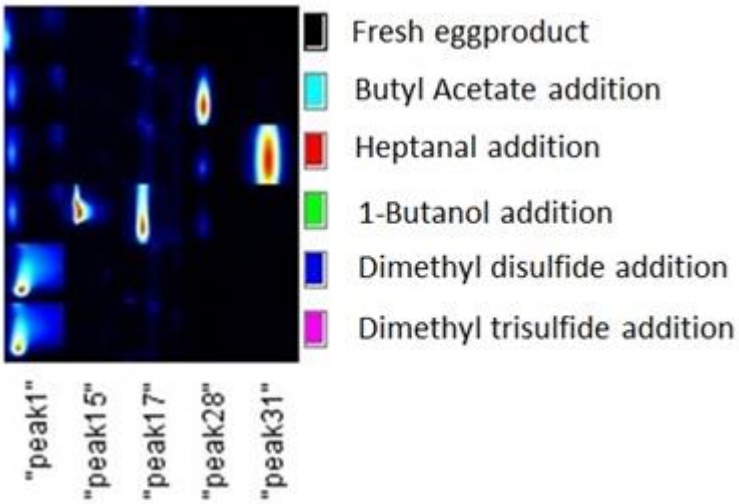


Figure 3: Correlation between spots and reference standards addition.

Supplementary Material

Section 1 – Sampling Plan

The full sampling scheme applied within this study, is reported in details in Table S1 and Table S2.

SAMPLE	0 hours	1 day	2 days	3 days	4 days	5 days
1	X*	X	X	X	X	
2	X	X*	X	X	X	
3	X	X	X*	X		
4	X	X	X	X*		
5	X	X	X			X*
6	X*	X	X			X
7	X	X	X			X
8	X	X			X*	X
9	X	X			X	X
10	X	X*	X	X	X	
11	X	X	X*	X	X	
12	X	X	X	X		
13	X	X	X*	X		
14	X	X*	X	X		
15	X	X	X*			X
16	X	X			X*	X
17	X*	X			X	X
18	X	X			X*	X
19	X	X*	X	X	X	
20	X	X	X	X	X*	
21	X	X	X*	X		
22	X*	X	X	X		
23	X	X*	X	X		
24	X	X	X	X*		
25	X	X*	X			
26	X*	X	X			
27	X	X				
28	X	X*				
29	X*	X				
30	X					
31	X					
32	X					
33	X					
34	X					
35	X					
36	X					
37	X					
38	X*					
39	X					
40	X*					
41	X					
42	X					
43	X					
44	X*					
45	X					
46	X*					
47	X					
48	X*					
49	X					
50	X					
51	X					
52	X					

= double sample prep

Table S1: Sampling plan of egg product batches, used for chemometric model set up

SAMPLE	0 hrs	1 day	2 days	3 days	4 days	5 days
Val_1	X	X	X			X
Val_2	X	X	X			X
Val_3	X	X	X	X	X	
Val_4	X	X	X	X	X	
Val_5	X	X	X			X
Val_6	X	X	X			X
Val_7	X	X	X	X		

Table S2: Sampling plan for the model external validation

Section 2 – Instrumental set up

The IMS instrumental set up is reported in Figure S1 and in Table S3, while the parameters applied for GC-IMS analysis are reported in Table S4.

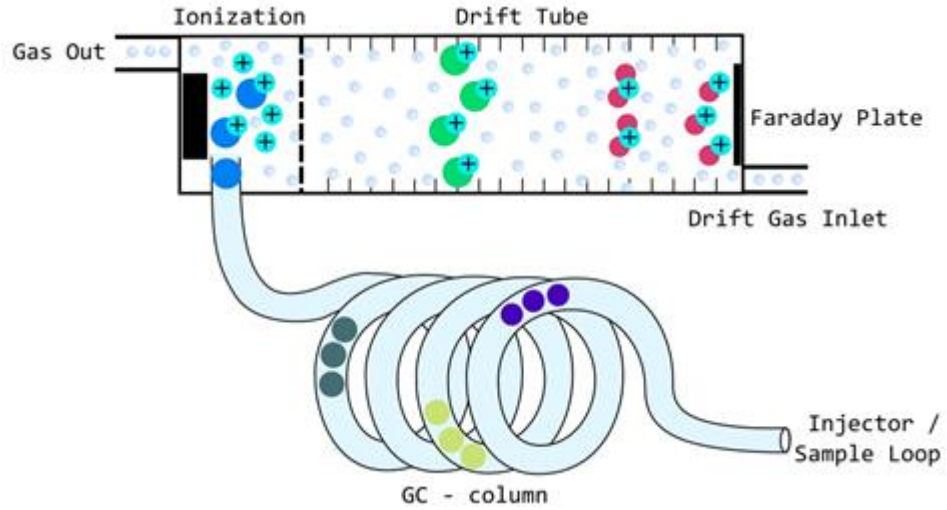


Figure S1: Schematic representation of the Flavourspec® instrument (from the Gas- Dortmund website)

Drift tube length	**9.8 cm**
Drift gas flow rate	150 mL min^{-1}
Drift time	30 ms
Drift Voltage	5 kV
Drift tube temperature	45 °C
Ionization mode	Positive

Table S3: IMS instrumental parameters

Injector temperature	**230 °C**
Carrier gas	Helium
Flow rate	1 ml/min
Transfer line temperature	230 °C
Ion source temperature	230 °C
Ionization mode	Electronic impact (positive)
Mass Range	30-350 m/z

Table S4: GC instrumental parameters

Section 3 – GC-IMS data treatment

The model was challenged with a validation set. Affinity data obtained are reported in Table S5. According to the established rules, values higher than 70% correspond to a certain belonging of the sample to that group.

Sample name	Affinity % with "fresh" class	Affinity % with "not fresh" class
val1_0h	77.97	22.03
val1_1D	18.16	81.84
val1_2D	2.09	97.91
val1_5D	12.39	87.61
val2_0h	95.57	4.43
val2_1D	67.13	32.87
val2_2D	4.45	95.55
val2_5D	22.31	77.69
val3_0h	92.25	7.75
val3_1D	23.28	76.72
val3_2D	3.07	96.93

val3_3D	5.48	94.52
val3_4D	0.00	100.00
val4_0h	100.00	0.0
val4-1D	0.00	100.00
val4_2D	0.00	100.00
val4_3D	7.75	92.25
val4_4D	1.65	98.35
val5_0h	100.00	0.00
val5_1D	16.74	83.26
val5_2D	4.82	95.18
val5_5D	11.27	88.73
val6_0h	100.00	0.00
val6_1D	0.86	99.14
val6_2D	0.00	100.00
val6_5D	10.70	89.30
val7_0h	100.00	0.00
val7_1D	0.00	100.00
val7_2D	0.00	100.00
val7_3D	0.00	100.00

Table S5: Predicted group affinity values for the validation samples.

Section 4 – Marker Identification

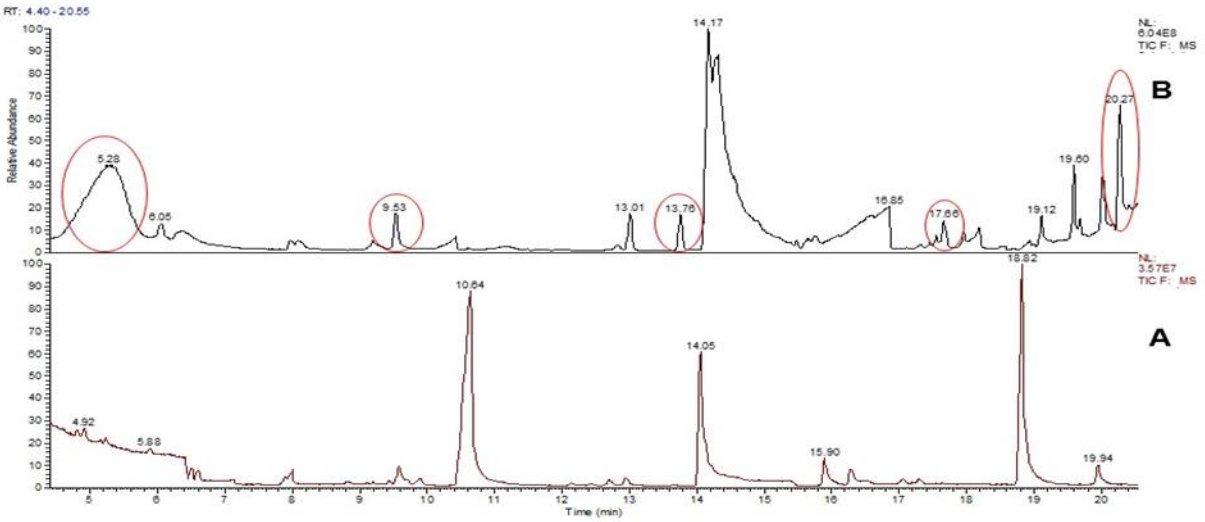


Figure S2: SPME-GC-MS chromatograms (partial) of the fresh egg product (A) and of the egg product left 5 days at room temperature (B). Peaks highlighted: RT 5.28 min – 1-Butanol; RT 9.53 min – Dimethyl disulfide; RT 13.76 min. – Butyl Acetate; RT 17.66 min. – Heptanal; RT 20.27 min. – Dimethyl trisulfide

Bibliographic references

Costanzo, A., Panseri, S., Giorgi, A., Romano, A., Caprioli, M., & Saino, N. (2016). The Odour of Sex: Sex-Related Differences in Volatile Compound Composition among Barn Swallow Eggs Carrying Embryos of Either Sex. *PLoS ONE, 16*(11), e0165055.

D.L. no. 65, February 1993 Italian Legislation reception of European Union.

Dragone, G., Mussatto, S., Oliveira, J., & Teixeira, J. (2009). Characterisation of volatile compounds in an alcoholic beverage produced by whey fermentation. *Food Chemistry, 112*, 929–935.

Eiceman, G., Karpas, Z., & Hill, Jr, H. (2016). *Ion Mobility Spectrometry, Third Edition.* Boca Raton: Taylor & Francis Group.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006). *Multi- and Megavariate Data Analysis - part 1- Basic Principles and Applications* (2 ed.). Umea: Umetrics AB.

Friedrich, J., & Acree, T. (1998). Gas Chromatography Olfactometry (GC/O) of Dairy Products. *International Dairy Journal, 8*(3), 235-241.

Garrido-Delgado, R., Arce, L., & Valcárcel , M. (2012). Multi-capillary column-ion mobility spectrometry: a potential screening system to differentiate virgin olive oils. *Analytical and Bioanalytical Chemistry, 402*(1), 489-498.

Garrido-Delgado, R., Dobao-Prieto, M. M., Arce, L., & Valcárcel, M. (2015). Determination of volatile compounds by GC–IMS to assign the quality of virgin olive oil. *Food Chemistry, 187*, 572-579.

Garrido-Delgado, R., Dobao-Prieto, M., Arce , L., Aguilar , J., Cumplido, J., & Valcárcel , M. (2015). Ion mobility spectrometry versus classical physico-chemical analysis for assessing the shelf life of extra virgin olive oil according to container type and storage conditions. *Journal Of Agricultural and Food Chemistry, 63*(8), 2179-2188.

Gas Chromatography (GC) coupled to Ion Mobility Spectrometry (IMS). (2018). Gas- Dortmund: https://www.gas-dortmund.de/index-gas.php?lan=1&spath=389 Accessed 12 June 2018

Gerhardt, N., Birkenmeier, M., Sanders, D., Rohn, S., & Weller, P. (2017). Resolution-optimized headspace gas chromatography-ion mobility spectrometry (HS-GC-IMS) for non-targeted olive oil profiling. *Analytical and Bioanalytical Chemistry, 409*, 3933-3942.

Giri, A., Osako, K., & Ohshima, T. (2010). Identification and characterisation of headspace volatiles of fish miso, a Japanese fish meat based fermented paste, with special emphasis on effect of fish species and meat washing. *Food Chemistry, 120*, 621–631.

Huang, Q., Qiu, N., Ma, M., Jin, Y., Yang, H., Geng, F., & Sun, S. (2012). Estimation of egg freshness using S-Ovalbumin as an indicator. *Poultry Science, 91*(3), 739-743.

Li, J., Zhu, S., Jiang, S., & Wang, J. (2017). Prediction of egg storage time and yolk index based on electronic nose combined with chemometric methods. *LWT - Food Science and Technology, 82*, 369-376.

Lin, H., Zhao, J., Sun, L., Chen, Q., & Zhou, F. (2011). Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis. *Innovative Food Science & Emerging Technologies, 12*(2), 182-186.

Reg. CE 853/2004. (n.d.).

Suman, M., Riani, G., & Dalcanale, E. (2007). MOS-based artifical olfactory system for the assessment of egg products freshness. *Sensors and Actuators B: Chemical, 125*(1), 40-47.

(2016). *U.S. Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection.*

Wang, Q., Zhou, K., Wang, C., & Ma, M. (2015). Egg Freshness Detection Based on Hyperspectral Image Technology. *Advance Journal of Food Science and Technology, 7*(8), 652-657.

Yimenu, S., Kim, J., & Kim, B. (2017). Prediction of egg freshness during storage using electronic nose. *Poultry Science, 96*(10), 3733-3746.

Zhang, L., Shuai, Q., Li, P., Zhang, Q., Ma, F., Zhang, W., & Ding, X. (2016). Ion mobility spectrometry fingerprints: A rapid detection technology for adulteration of sesame oil. *Food Chemistry, 192*, 60-66.

Zhang, W., Pan, L., Tu, S., Zhan, G., & Tu, K. (2015). Non-destructive internal quality assessment of eggs using a synthesis of hyperspectral imaging and multivariate analysis. *Journal of Food Engineering, 157*, 41-48.

Zhao, J., Li, H., Chen, Q., Huang, X., Sun, Z., & Zhou, F. (2010). Identification of egg's freshness using NIR and support vector data description. *Journal of Food Engineering, 98*(4), 408-414.

Author Contributions

Conceptualization: D.C., C.D.,M.S.; study execution: D.C., S.Z.; writing-original draft preparation: D.C..; writing -review and editing: all authors; supervision. C.D., M.S

Egg product freshness evaluation: a metabolomic approach

Daniele Cavanna[1,2], Dante Catellani[1], Chiara Dall'Asta[2], Michele Suman[1]

[1]Barilla Advanced Laboratory Research, Via Mantova 166 - 43122 Parma, Italy
[2]Department of Food and Drug, University of Parma, Parco Area delle Scienze 95/A- 43124 Parma, Italy

Introduction

Starting from the beginning of the new millennium, consumers' attention on the authenticity and quality of food commodities has strongly increased, leading to an increase in the demand of fit-for-purpose methods for detecting food fraud from both industries and research institutes.

Eggs, mostly in the egg products form, are largely used for the formulation of food products. Freshness is therefore a crucial parameter for ensuring the production of safe and high quality commodities. However, the assessment of egg freshness is still challenging, due to the lack of robust chemical markers.

In the past, according to the European legislation, two parameters typical of microbiological fermentation had to be monitored: Lactic acid (≤ 1000 mg/kg dry egg) and Succinic Acid (≤ 25 mg/kg dry egg) (D.L. no. 65, February 1993 Italian Legislation reception of European Union), by the current laws consider only Lactic Acid as reliable marker (Reg. CE 853/2004).

These two compounds could act sometimes as "tardive" markers of eggs ageing and their increase is often not linear during the time; moreover, the use of only one/two molecule(s) for the evaluation of such critical parameter can be not sufficient in some circumstances; at the same time more "not freshness" compounds should be monitored.

Up to now, major attention was paid to the development of rapid techniques able to assess egg products freshness, e.g. using electronic noses (Li, Zhu, Jiang, & Wang, 2017) (Suman, Riani, & Dalcanale, 2007) (Yimenu, Kim, & Kim, 2017) or spectroscopic analyses (Wang, Zhou, Wang, & Ma, 2015) (Zhao, et al., 2010) (Lin, Zhao, Sun, Chen, & Zhou, 2011). These techniques are particularly suitable for industrial screening purposes, due to the reduced time and costs.

However, the identification of robust "markers of freshness" (i.e. compounds that decrease their intensity during the egg products storage) could strongly support the evaluation of egg products aging, with a reduction of the risk for the final consumer.

In this context, the use of "non-targeted" methods based on a metabolomic approach – which are considered emerging methodologies for detecting food frauds (Riedl, Esslinger, & Fauhl-Hassek, 2015) – will offer the opportunity to identify and validate proper markers.

In consideration of the ageing process, chances in the volatile profile of egg products can be expected. Therefore, among possible analytical techniques, Gas Chromatography Mass Spectrometry (GC-MS) could be used for the identification of these changes, in agreement with previous applications in the field of food frauds, i.e. tomatoes cultivars discrimination (Figueira, Câmara, Pereira, & Câmara, 2014), honey authenticity (Silva, et al., 2017) or geographical origin of saffron (Aliakbarzadeh, Sereshti, & Parastar, 2016).

However, the non-volatile fraction of food may be of great interest for the discrimination of ageing. Scientific papers suggest that other rapid mass spectrometry approaches, as for example Direct Analysis Real Time Mass Spectrometry (DART-MS) (Hajslova, Cajka, & Vaclavik, 2011) or Rapid Evaporative Ionization Mass Spectrometry (REIMS) (Balog, et al., 2016), could lead to the characterization of raw materials, avoiding the chromatographic separation.

In any case, Liquid Chromatography High Resolution Mass Spectrometry (LC-HRMS) is the technique of choice for the execution of a metabolomic study: thanks to its versatility, a huge amount of information related to the non-volatile profile can be extracted from the chromatographic fingerprints.

Literature presents applications of this approach for the detection of frauds related to a wide range of raw materials, from fruit juices (Jandric, Islam, Singh, & Cannavan, 2017) to wheat (Righetti, et al., 2018), from extra virgin olive oil (Kalogiouri, Alygizakis, Aalizadeh, & Thomaidis, 2016) to spirit drinks (Collins, Zweigenbaum, & Ebeler, 2014).

Independently from the analytical technique used, the creation of robust chemometric models is a crucial step for the selection of new marker compounds responsible of the target fraud (Rubert, Zachariasova, & Hajslova, 2015).

On the basis of what described above, this work presents a metabolomic study on egg products samples able to select and identify a group of new compounds that can be used as "freshness" or "not freshness" markers.

LC-HRMS was exploited for fingerprints recording; subsequently, robust data elaboration was executed, creating different chemometric models able to select some marker compounds that were then studied and identified.

The final step of this workflow was the validation of these molecules, with the aim to assess their value as chemical markers, independently from the chemometric model used to select them.

Materials and Methods

Chemicals and samples description

Methanol, Acetonitrile, Formic Acid and Ammonium Formate were purchased from VWR International, Ltd (Poole, United Kingdom).

Analytical standards of Chloramphenicol, Lactic Acid, Succinic Acid, Phenyllactic Acid, Tyramine, Uracil, Asparagine, Glutamine, Guanosine, Guanosine Monophosphate, Serine, Uridine and Uridine Monophosphate were purchased from Sigma Aldrich (St. Louis, MO).

Water was purified using a Milli-Q system (Millipore, Bedford, MA).

Egg products samples were directly collected from batches which daily arrive at production plant sites; in order to increase the variability of the model, the sampling period lasted six weeks and samples coming from different batches, different suppliers and with different amounts of carotenoids (in the range of tenths of mg/kg, authorized for improving appearance and correspondent consumer preference in finished products) were selected.

Experimental Design

The first step of the adopted workflow, as suggested by US Pharmacopoeia (U.S. Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection, 2016), required the creation of chemometric models able to discriminate between fresh and not fresh samples; subsequently, the features responsible of this clusterization were selected and a tentative of compounds identification was performed.

Finally, marker compounds identified were validated, that means that the target molecules were searched in new egg products batches subject to the same ageing applied during the model creation (Jandric, Islam, Singh, & Cannavan, 2017).

For the creation of the model, 29 egg products batches were collected and extracted immediately after the reception at the production plant site; subsequently, these raw materials were left at room temperature and extracted again after 1 day and after 2 days, according to the sampling plan described in table 1. Summarizing, a global amount of 79 samples was used for the creation of the chemometric model.

For the validation of the marker compounds, 7 new egg products batches were collected, extracted and analyzed following the same experimental design described for the model creation: the total amount of samples was 21.

Sample preparation

300 µl of each egg product at each time point were spiked into a 1.5 ml tube, together with 30 µl of a 20 µg/ml Chloramphenicol solution (used as internal standard) and 900 µl of a mixture of Acetonitrile:water in the ratio 80:20 (stored at 2-8°C).

This mixture was vortexed for 1 minute, then stored for 5 minutes at -20°C and subsequently centrifuged for 10 minutes at 14000 rpm with a Rotina 380R (Hettic lab technology, Tuttlingen, Germany) maintained at 4°C.

The extract was filtered into an HPLC vial with a 0.22 µm PTFE syringe filter (Phenomenex, Torrance, CA) and stored at -20°C until analysis with LC-HRMS instrument.

For the evaluation of method reproducibility, 20% of the samples were double prepared, as detailed in table 1.

During each extraction session, the same procedure was performed also into empty tubes in which all the steps were executed without the egg product addition. These samples were labelled as "extraction blanks".

A 500 ng/ml Chloramphenicol solution in Acetonitrile: water 80:20 was prepared and named "standard solution".

Additionally, two quality control (QC) samples were prepared mixing 10 µl of each extract sample.

LC-HRMS analysis

HPLC analysis was performed with a Dionex UltiMate 3000 UHPLC (Thermo Fisher Scientific, Inc., Waltham, MA) equipped with a Luna Omega 150 mm x 2.1 mm, 1.6 µm particle size analytical column (Phenomenex, Torrance, CA, USA) with a pre column security guard ultra C18, both maintained at 30°C. Gradient elution was performed using Formic Acid (FA) and Ammonium Formate (AF) as mobile phase modifiers with a constant flow rate of 0.3 ml/min

Gradient conditions are the following: after 1 minute with 95 % of mobile phase A (0.1% FA and 5mM AF in Water) and 5% of mobile phase B (0.1% FA and 5mM AF in Methanol), the percentage of solvent B increased to 95% in 24 minutes and then was maintained at this percentage for 5 minutes before column re-equilibration (10 minutes).

Autosampler was maintained at 5 °C and the injection volume was 5 µl.

Mass spectrometry detection was performed with a benchtop Q Exactive™ Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific, Waltham, MA) equipped with a heated (H)

electrospray ionization (ESI) interface (Thermo Fisher Scientific,Waltam, MA). Two analytical sequences (one with positive and one with negative ionization mode) were executed with a "Full Scan-data dependent fragmentation" experiment.

Source conditions were as follows: sheath and auxiliary gas flow rates 40 and 20 arbitrary units, respectively; heater temperature 250 °C with a spray voltage of 3.2 kV (ESI pos) and -3.0 kV (ESI neg). The capillary was kept at 220 °C and the S-lens RF level was set at 55 AU for both the acquisition modes.

The full-scan accurate mass spectra from 75 to 1000 Da for both the ionization modes were obtained with a resolution of 70000 FWHM (*m/z* 200), automatic gain control (AGC) target 1e6 and maximum injection time (IT) 200 ms.

In all the experiments, the data-dependent MS/MS acquisition was executed with a resolution of 17500 FWHM (*m/z* 200) and intensity threshold 6e4. The quadrupole isolation window was kept at 2.0 m/z with a TopN value of 5. The scan range was from 50 to the fragmented mass m/z (m/z +25), AGC target 2e5, maximum IT 50 ms, normalized collision energy (NCE) 30% with ±50% step.

Samples, together with the "extraction blanks", were randomly injected in order to avoid systematic bias (Righetti, et al., 2018). The "standard solutions" and the QC samples were injected at the beginning of the sequences and every 10 sample injections.

Data treatments and statistics

UHPLC-HRMS raw data were acquired using Xcalibur software (version 3.0 Thermo Fisher Scientific,Waltam, MA); peaks alignment, "extraction blanks" subtraction and features extraction were performed using Compound Discoverer software (version 2.0 Thermo Fisher Scientific,Waltam, MA); the mass range inspected was between 75 *m/z* and 1000 *m/z*from 1 minute to 30 minutes of the chromatographic runs.

The values of the critical parameters for features extractions are the following: precursor ion deviation 5 ppm (for the positive runs) and 10 ppm (for the negative runs); maximum retention time shift 0.3 minutes; minimum peak intensity for a peak to be selected 1000000 AU; relative intensity tolerance used for isotope search 30%.

Structures prediction was also performed using "Chemspider" databases, setting a maximum mass shift of 5 ppm for positive acquisitions and 10 ppm for negative acquisitions.

For the "*m/z CLOUD™"* MS/MS library search, the precursor mass tolerance used was 0.05 Da while the fragments mass tolerance was 10 ppm.

The resulting two data matrixes (for positive and negative ionization modes), containing the area values provided by Compound Discoverer for all the features, were exported and processed with SIMCA software (version 14.1 Umetrics, Umea, Sweden) for chemometric data elaboration.

Data were log transformed and Pareto scaled, then a preliminary Principal Component Analysis (PCA) was executed in order to check the clusterization of the samples and the QCs positioning in the scores plot.

Subsequently, features were filtered and only the ones that had a CV% lower than 40% in the QC samples were selected.

A new PCA was executed in order to check the expected improvement in samples separation and QCs positioning; after the evaluation of the replicate samples placement in the scores plot, a final PCA was calculated without the QC samples and with the average values of the replicates.

Afterwards, supervised orthogonal partial least square discriminant analysis (OPLS-DA) models were built comparing the "fresh" samples against the "1 day" samples, against the "2 days" samples and against the union of "1 day" and "2 days" samples.

Thanks to the S-plots, statistically significant markers responsible of the clusterization were selected: the ions furthest away from the origin contribute more significantly to the separation between the groups and may therefore be regarded as the differentiating ones (Jandric, et al., 2014). This approach was used for the six OPLS-DA models in order to be sure that all the most discriminative features could be selected.

In addition, their VIP (variable influence on projection) values were evaluated in order to assess their relevance for the chemometric model (VIP values had to be > 1.4) (Akarachantachote, Chadcham, & Saithanu, 2014).

Features that survived this process were studied and a tentative of compounds identification was performed.

In all the chemometric models created, an internal leave 1/7 out cross validation was executed.

Compounds identification workflow

Compounds identification was performed according to the following steps:

1) Chemical formula hypothesis, taking into account also adducts, if present.
2) Retention time evaluation.
3) Evaluation of the mass shift between the experimental and the theoretical values.
4) Isotopic pattern evaluation and hypothesis of a chemical structure referring also to Chemspider databases.

5) MS/MS spectra study and comparison between the theoretical fragmentations (performed with Compound Discoverer) and the fragments detected.
6) MS/MS spectra comparison with the m/z CLOUD™ database (in which fragmentations with other Orbitrap instruments are available).
7) If available, Reference Standard injection as final confirmation.

The Standard Initiative in Metabolomics (Sumner, et al., 2007) and subsequent improvements (Schymanski, et al., 2014) suggest different ranks of compounds identification according to the number of steps completed: identification steps completed: 1-5, level of identification: 3; identification steps completed: 1-6, level of identification: 2; identification steps completed: 1-7, level of identification: 1; this classification was used for the selected features.

Markers validation procedure

The compounds identified with the workflow detailed above were searched in seven new egg products batches not used for the model creation; they were collected, extracted and analyzed with the same experimental design previously described.

The maximum reliability of these compounds as "freshness" or "not freshness" markers implies they have to be confirmed not only with their presence or absence in the new samples but also highlighting the same increasing or decreasing trend through the time points.

Results and Discussion

Extraction procedures and sequences evaluation

Before starting the real data elaboration, different evaluations on the internal standard results were performed.

Chloramphenicol is an exogenous compound that was added to every sample with the aim of monitoring the extraction procedures and the goodness of the analytical sequences (Khamis, Adamko, & El-Aneed, 2017).

The injection of the "standard solution" periodically during the sequence provided information about the sequence trends and helped in the detection of potential decrease in signal's intensity.

For this reason, at the end of each analytical sequence, the CV% of both areas and retention times of Chloramphenicol peak in the "standard solution" injections (11 for each sequence) were calculated.

For the sequence analyzed in positive ionization mode, the values obtained were respectively 10.8%

(Areas) and 2.2% (retention time); for the negative ionization mode, 10.1% (Areas) and 2.8% (retention time) respectively.

The results obtained highlight that the response of the instrument was acceptable during the whole sequences (126 injections each) for both positive and negative ionization modes.

An analogous evaluation of Chloramphenicol peaks intensity and retention times shift was performed for all the sample injections: in this way more information about the goodness of the analytical sequences were obtained; moreover, the CV% of the area values of the internal standard was used for the extraction reproducibility evaluation, assuring that no mistakes occurred during samples preparation.

For the sequence analyzed in positive ionization mode, the areas CV was 7.2% and the retention time CV was 1.6%; for the negative ionization mode, 10.5% (areas) and 3.1% (retention time) respectively.

According to the results obtained, we can deduce that no issues occurred during sample extraction procedures; moreover, the slight shift in the retention times is a further proof of the goodness of the analytical sequences.

Chemometric data elaboration

The preliminary PCA obtained with the areas of the entire data set highlighted a separation between the fresh samples and the not fresh ones with a clustering of the QC samples for both positive and negative ionization mode (data not shown); the first two components described more than 45% of the variance for both the models.

After data filtering (obtained by selecting in the QC samples the features with CV lower than 40% in the area values) an improvement in the clusterization of the samples was obtained, with a tight clustering of the QC samples (fig. 1 and fig. 2). In addition, the overlay in the scores plot of the replicate extracted samples was essentially complete.

The tight clustering of the QC samples in the center of the score plot is a further confirmation of the robustness of the analytical procedure; moreover this ensure that the separation through the groups is not random but due to a real variability (Godzien, Alonso-Herranz, Barbas, & Armitage, 2015).

The overlay in the scores plot of the replicate samples certify that the extraction procedure is repeatable not only for the Chloramphenicol compound but for the whole fingerprint of the egg product.

The final PCA, created with the average value of the replicates and without the QC samples, clearly highlight the separation between the fresh samples and the not fresh ones (fig. 3 and 4).

All the OPLS-DA supervised models, as expected, extraordinarily increase the separation between the two groups. As example, figures 5 and 6 present the OPLS-DA models created between the fresh samples and the "1 day" samples.

Table 2 and 3 summarize the variance of the x and y variables explained by the models previously described (R^2X (cum) and R^2Y (cum)) and the percent of variation of the training set predicted by the model (Q^2) according to cross validation.

The results obtained strongly indicate that the models created for both positive and negative ionization modes are reliable and robust.

Compounds identification

Features responsible of the clusterization were selected and a group of them was identified according to what described in the "compounds identification workflow" paragraph.

Both markers of "not freshness" (that drastically increase their intensity or even appear during the eggs ageing) and of "freshness" (that drastically decrease their intensity or even disappear during the eggs ageing) were identified.

Tables 4 and 5 summarize the compounds responsible of the "freshness" and "not freshness" classification, ranked according to the level of identification suggested by the Standard Initiative in Metabolomics (Sumner, et al., 2007) (Schymanski, et al., 2014).

12 compounds were identified with their respective reference standards and a total amount of 31 markers (15 of freshness and 16 of not freshness) were selected.

Table 6 resumes the number of features filtered out through each statistical step.

As highlighted in the markers lists, Lactic Acid and Succinic Acid were undoubtedly detected, but together with many other compounds. This could lead to a more complete evaluation of the freshness issue: potentially 31 m/z values can be simultaneously considered for this topic.

Markers validation

As also suggested by the US Pharmacopoeia (U.S. Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection, 2016), predictive chemometric models should be always validated with an external set of samples (possibly collected in different periods) that have to be treated as unknown; their classification should be predicted, in order to certify that the samples clustering is real and not related to some overfitting of the chemometric model.

This approach is hard to perform with metabolomic results: because of the UHPLC-HRMS intrinsic variability, only chromatograms acquired with the same analytical sequence can be aligned and

studied; the use of chemometric model to predict the classification of samples acquired with different sequences could lead to dangerous mistakes. Moreover, the goal of this work was the identification of new markers of freshness and the chemometric model was considered only a tool to reach this goal.

For these reasons, the marker compounds and not the model were validated, that means that all these 31 features were searched in the new samples and their presence or absence, together with their trends through the time points, were evaluated and compared with the results obtained during the model creation, as presented in tables 7 and 8.

All the target molecules were detected with almost the same trend through the time points (with exception of a couple of markers but only in terms of relative ratio): these results prove that the compounds identified are real markers of "freshness" or "not freshness" in egg products and are not related only to a misleading overfitting of the chemometric model.

Markers interpretation

Most of the markers reported in this work have not been previously identified in eggs shelf life studies, because of the analytical approach. While egg freshness is commonly followed by volatile profiling or by e-nose fingerprinting, the use of HR-LC/MS allowed to enlarge the identification to medium-polar non-volatile compounds.

Volatile markers are actually based on the formation of volatile aldehydes from fatty acids oxidative degradation (Yimenu, Kim, & Kim, 2017). In our study, markers deriving from the very first stage of fatty acid oxidative degradation are reported in the fresh products (namely, 4-hydroxybutyric acid, hydroxycaproic acid, 2-hydroxyvaleric acid). These are probably formed during thermal stabilization of egg products, accumulated over time and further degraded to volatile aldehydes along storage.

Interestingly, most of the compounds reported as markers of freshness are precursors of compounds listed as "markers of non freshness". In particular, it can be observed a strong activation of nitrogen and pyrimidine metabolism pathways, probably on account of the metabolic activity of residual microorganisms growing over time. This leads to the decrease of free amino acids, with the accumulation of biogenic amine (i.e. tyramine), amino acid and purine degradation products (i.e. phosphorylethanolamine, N-acetylhistidine, phenyllactic acid, uracil), and precursors of peptidoglycane (D-Alanyl-D-alanine). Consistently, an activation of microbial fermentation is also attested by the accumulation of succinic and lactic acid over time.

Conclusions

In the current study, new compounds related to freshness in egg products samples were identified using a non-targeted metabolomic approach. After model creation, the robustness of the identified markers of freshness was assessed through the analysis of a validation set of not previously used egg products.

Along with Lactic Acid and Succinic Acid, the reference compounds already considered in the EU legislation, other 29 compounds were detected as discriminant features, allowing a more robust evaluation of freshness.

Further improvements of the results presented in this study should lead to two directions: firstly, target methods on the identified compounds could be developed for a quantitative evaluation: quality control laboratories should only be able to detect these 31 molecules, avoiding the need of high resolution mass spectrometry and chemometric software, that are much more expensive than a simple single stage LC-MS instrument.

Secondly, this group of compounds, and generally the recorded fingerprints, could be helpful also for both the safety evaluations on microbial growth perspectives and the detection of other frauds related to egg products food chain, as for example the illegal use of incubated eggs.

Acknowledgement

Authors would like to thank Angelo Restori and all the Barilla's quality control laboratory team for the invaluable support.

This work was supported by the European Union, Seventh Framework Program (FP7/2007-2013), under Grant Agreement 613688 (Ensuring the Integrity of the European food chain FoodIntegrity).

Tables

SAMPLE	0 hrs	1 day	2 days
1	X*	X	X
2	X	X*	X
3	X	X	X*
4	X	X	X
5	X	X	X
6	X*	X	X
7	X	X	X
8	X	X	
9	X	X	
10	X	X*	X
11	X	X	X*
12	X	X	X
13	X	X	X*
14	X	X*	X
15	X	X	X*
16	X	X	
17	X*	X	
18	X	X	
19	X	X*	X
20	X	X	X
21	X	X	X*
22	X*	X	X
23	X	X*	X
24	X	X	X
25	X	X*	X
26	X*	X	X
27	X	X	
28	X	X*	
29	X*	X	

Table 1: sampling plan for the model creation. *= double sample prep

Model	PCA (preliminary)	PCA (filtered)	PCA (final)	OPLS-DA ("Fresh" vs "1D+2D")	OPLS-DA ("Fresh" vs "1D")	OPLS-DA ("Fresh" vs "2D")
R^2X (cum)	0.691	0.750	0.725	0.618	0.589	0.680
R^2Y (cum)	/	/	/	0.992	0.997	0.999
Q^2	0.544	0.643	0.590	0.989	0.990	0.994

Table 2: global value of R^2X (cum), R^2Y (cum) and Q^2 parameters for positive ionization mode

Model	PCA (preliminary)	PCA (filtered)	PCA (final)	OPLS-DA ("Fresh" vs "1D+2D")	OPLS-DA ("Fresh" vs "1D")	OPLS-DA ("Fresh" vs "2D")
R^2X (cum)	0.750	0.769	0.777	0.617	0.627	0.698
R^2Y (cum)	/	/	/	0.994	0.997	0.999
Q^2	0.584	0.577	0.546	0.991	0.991	0.996

Table 3: global value of R^2X (cum), R^2Y (cum) and Q^2 parameters for negative ionization mode

Name	Pseudomolecular ion	Detected m/z	RT (min)	Predicted Formula	Mass error (ppm)	CV IN QCs (%)	VIP Value	ID type
Lactic acid	$[M-H]^-$	89.0244	1.73	C3 H6 O3	12.4*	27	2.59[c]	1
Phenyllactic acid	$[M-H]^-$	165.0555	11.95	C9 H10 O3	5.45	11	2.64[b]	1
Succinic acid	$[M-H]^-$	117.0193	3.22	C4 H6 O4	8.55	11	2.49[b]	1
Tyramine	$[M+H]^+$	138.0913	3.65	C8H11NO	0.72	30	2.21[b]	1
	$[M+H-NH3]^+$	121.0650			1.65			
Uracil	$[M-H]^-$	111.0196	1.70	C4 H4 N2 O2	6.31	15	2.88[c]	1
4-Hydroxybutyric acid	$[M-H]^-$	103.0398	3.90	C4 H8 O3	8.74	23	1.83[b]	2
6-Methylquinoline	$[M+H]^+$	144.0807	8.39	C10 H9 N	0.69	7	2.27[b]	2
D-Alanyl-D-alanine	$[M-H]^-$	159.0773	1.34	C6 H12 N2 O3	5.66	12	1.88[b]	2
Hydroxycaproic acid	$[M-H]^-$	131.0712	11.25	C6 H12 O3	7.63	17	3.57[c]	2
Phosphorylethanolamine	$[M-H]^-$	140.0118	1.14	C2 H8 N O4 P	7.85	18	3.00[c]	2
3-Thiomorpholinecarboxylic acid	$[M+H]^+$	148.0426	1.70	C5 H9 N O2 S	0.68	29	2.09[a]	3
Arginine-Proline dipeptide	$[M+H]^+$	272.1716	1.30	C11 H21 N5 O3	0.37	13	2.55[c]	3
Lysine Leucine dipeptide	$[M-H]^-$	258.1818	1.70	C12 H25 N3 O3	2.32	16	1.81[b]	3

N-Acetylhistidine	$[M-H]^-$	196.0725	1.43	C8 H11 N3 O3	4.08	17	2.20[b]	3
2-Hydroxyisovaleric acid or 2-Hydroxymethylbutyric acid	$[M-H]^-$	117.0557	7.15	C5 H10 O3	9.39	18	2.15[c]	3
Mannose 6- phosphate or Glucose 6-phosphate	$[M-H]^-$	259.0226	1.08	C6 H13 O9 P	5.01	32	2.99[c]	3

* This value is higher than the limit (10 ppm) but is due to the low molecular weight of Lactic Acid. The injection of the reference standard certified its identity beyond any reasonable doubt

Table 4: resume of "not freshness" markers identified. a= VIP value obtained from the OPLS-DA model 0h vs all; b= VIP value obtained from the OPLS-DA model 0h vs 2D; c= VIP value obtained from the OPLS-DA model 0h vs 1D

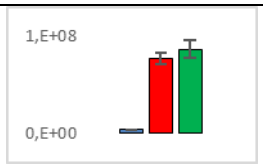
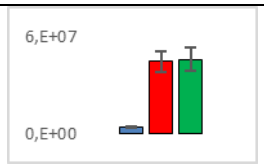
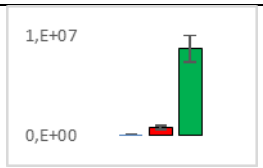
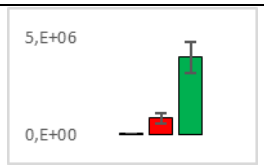
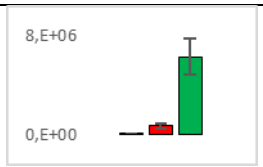
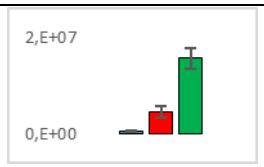
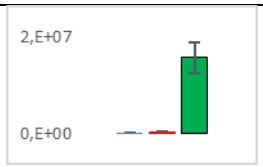
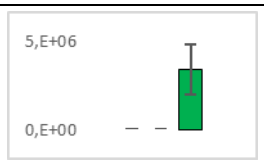
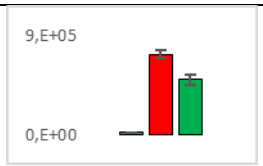
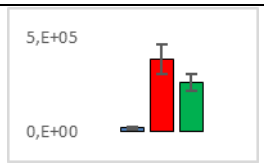
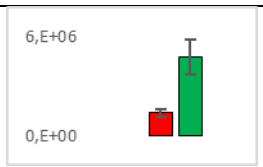
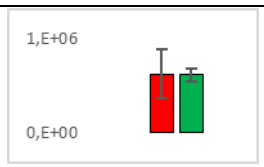
Name	Pseudomolecular ion	Detected m/z	RT (min)	Predicted Formula	Mass error (ppm)	CV IN QCs (%)	VIP Value	ID type
Asparagine	$[M+H]^+$	133.0608	1.21	C4 H8 N2 O3	0.30	14	2.35[c]	1
Glutamine	$[M+H]^+$	147.0762	1.23	C5 H10 N2 O3	1.36	8	2.75[c]	1
	$[M-H]^-$	145.0615	1.17		4.83	12	2.55[c]	
Guanosine	$[M-H]^-$	282.0844	4.85	C10 H13 N5 O5	3.90	33	2.88[c]	1
Guanosine monophosphate (GMP)	$[M-H]^-$	362.0506	1.66	C8 H19 N3 O9 P2	1.93	28	3.10[c]	1
Serine	$[M+H]^+$	106.0502	1.23	C3 H7 N O3	2.83	8	2.67[c]	1

Uridine	$[M-H]^-$	243.0619	1.68	C9 H12 N2 O6	2.88	16	2.85[c]	1
Uridine monophosphate	$[M-H]^-$	323.0285	1.37	C9 H13 N2 O9 P	3.10	17	3.20[c]	1
Arginine	$[M+H]^+$	175.1189	1.22	C6 H14 N4 O2	0.22	16	2.15[b]	2
	$[M-H]^-$	173.1038	1.17		2.89	9	2.38[b]	
Aspartic Acid	$[M+H]^+$	134.0448	1.22	C4 H7 N O4	0.03	10	1.72[c]	2
Histidine	$[M+H]^+$	156.0768	1.20	C6 H9 N3 O2	0.02	20	1.69[c]	2
Methionine sulfoxide	$[M+H]^+$	166.0530	1.28	C5 H11 N O3 S	1.20	12	1.68[c]	2
Methylhistidine	$[M+H]^+$	170.0924	1.22	C7 H11 N3 O2	0.06	37	2.50[c]	2
Threonine	$[M+H]^+$	120.0656	1.25	C4 H9 N O3	0.80	8	1.42[b]	2
Guanosine 5'-diphospho-D-mannose or Guanosine 5'-diphospho-D-glucose	$[M-H]^-$	604.0694	1.30	C16 H25 N5 O16 P2	0.17	27	2.77[c]	3
N-Acetyl-α-D-galactosamine or N-Acetyl-α-D-glucosamine	$[M+H]^+$	222.0973	1.38	C8 H15 N O6	0.45	17	1.87[c]	3

Table 5: resume of "freshness" markers identified. a= VIP value obtained from the OPLS-DA model 0h vs all; b= VIP value obtained from the OPLS-DA model 0h vs 2D; c= VIP value obtained from the OPLS-DA model 0h vs 1D

DATA ANALYSIS STEP	NUMBER OF FEATURES SELECTED	
	POSITIVE IONIZATION MODE	NEGATIVE IONIZATION MODE
Peak alignment	3452	4893
Filtration according to CV% in QC	2189	3407
Extracted features from S-plots	207	240
Final list	14	19

Table 6: resume of the features selected after each statistical step

COMPUND NAME	MODEL CREATION	MARKER VALIDATION
LACTIC ACID		
PHENYLLACTIC ACID		
SUCCINIC ACID		
TYRAMINE		
URACIL		
2-HYDROXYISOVALERIC ACID OR 2-HYDROXYMETHYLBUTYRIC ACID		

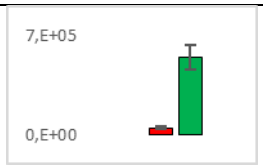
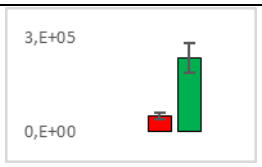
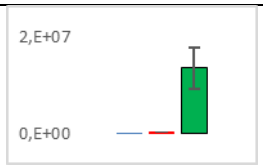
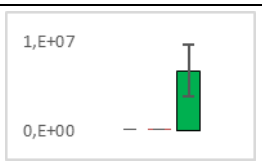
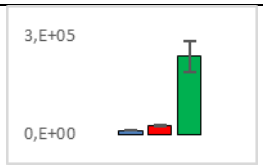
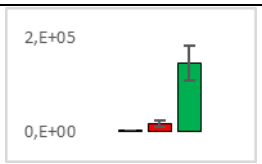
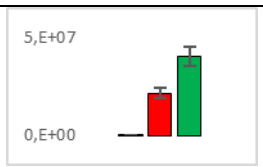
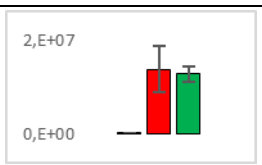
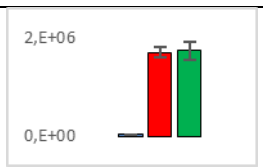
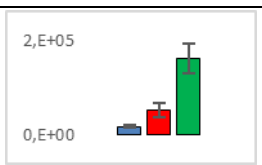
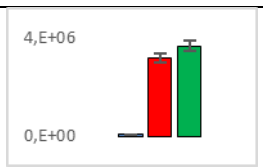
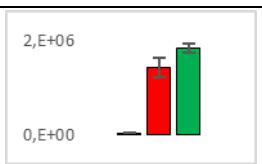
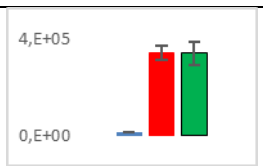
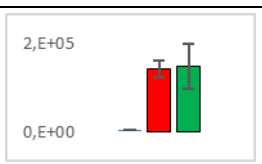
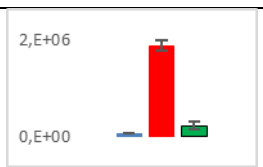
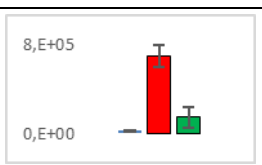
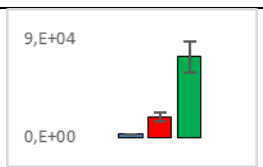
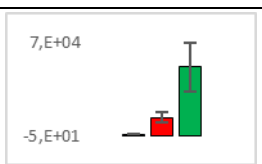
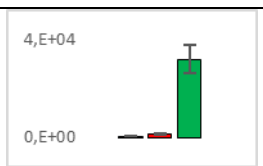
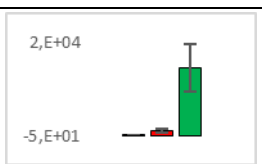
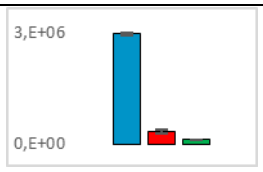
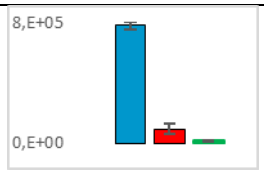
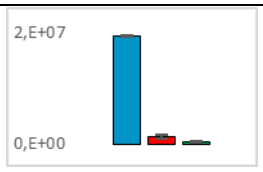
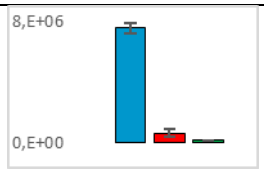
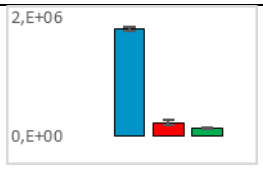
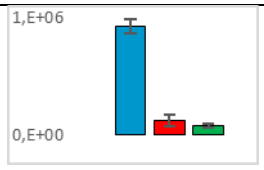
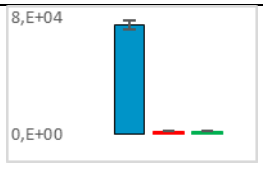
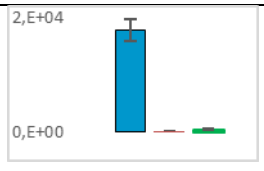
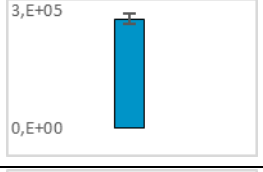
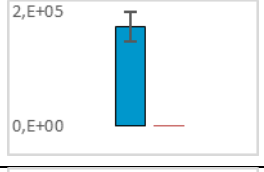
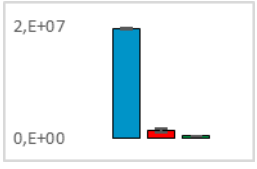
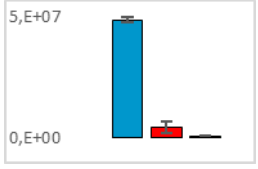
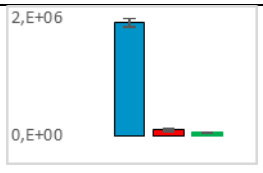
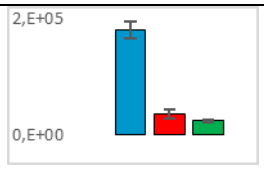
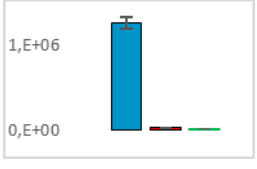
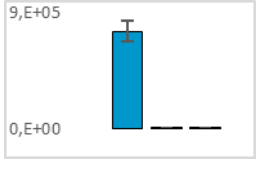
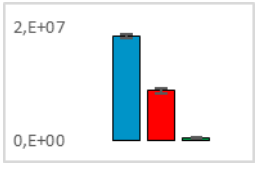
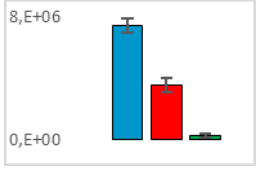
COMPUND NAME	MODEL CREATION	MARKER VALIDATION
4-HYDROXYBUTYRIC ACID		
6-METHYLQUINOLINE		
D-ALANYL-D-ALANINE		
HYDROXYCAPROIC ACID		
MANNOSE 6- PHOSPHATE OR GLUCOSE 6- PHOSPHATE		
PHOSPHORYLETHANOLAMINE		
3-THIOMORPHOLINECARBOXYLIC ACID		
ARGININE-PROLINE DIPEPTIDE		
LYSINE LEUCINE DIPEPTIDE		
N-ACETYLHISTIDINE		

Table 7: comparison of the mean area values (+/- Standard Error) of not freshness markers between the model creation and the markers validation. Time points: 1st bar, “t zero”; 2nd bar, “1 day”; 3rd bar, “2 days”

COMPUND NAME	MODEL CREATION	MARKER VALIDATION
ASPARAGINE		
GLUTAMINE (ESI+)		
GLUTAMINE (ESI-)		
GUANOSINE		
GUANOSINE MONOPHOSPHATE		
SERINE		
URIDINE		
URIDINE MONOPHOSPHATE		
ARGININE (ESI+)		

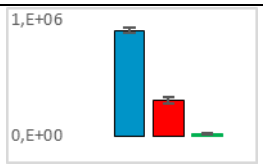
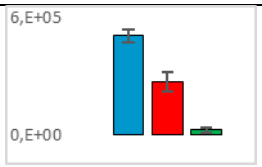
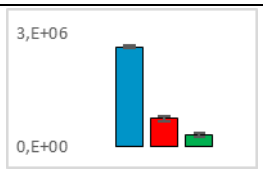
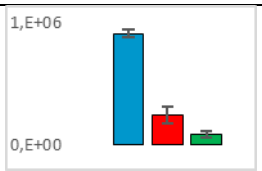
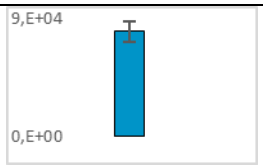
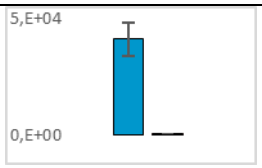
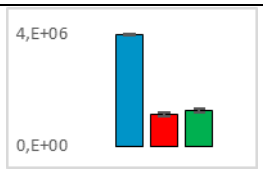
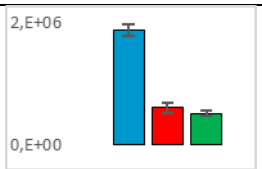
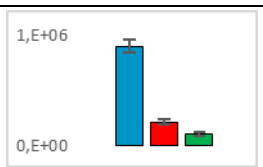
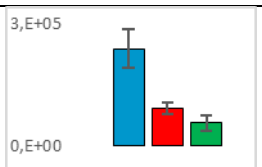
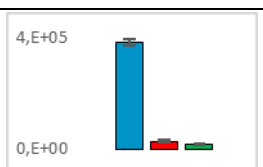
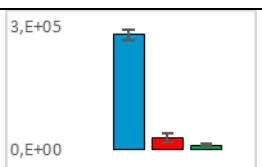
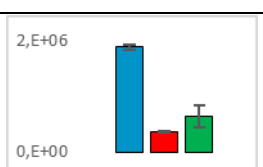
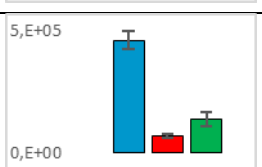
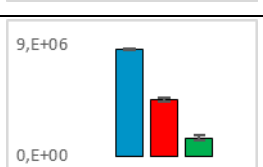
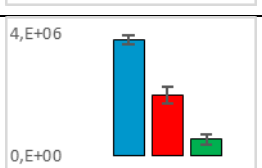
COMPUND NAME	MODEL CREATION	MARKER VALIDATION
ARGININE (ESI-)		
ASPARTIC ACID		
GUANOSINE 5'-DIPHOSPHO-D-MANNOSE OR GUANOSINE 5'-DIPHOSPHO-D-GLUCOSE		
HISTIDINE		
METHIONINE SULFOXIDE		
METHYLHISTIDINE		
N-ACETYL-α-D-GALACTOSAMINE OR N-ACETYL-α-D-GLUCOSAMINE		
THREONINE		

Table 8: comparison of the mean area values (+/- Standard Error) of freshness markers between the model creation and the markers validation. Time points: 1st bar, “t zero”; 2nd bar, “1 day”; 3rd bar, “2 days”

Figures

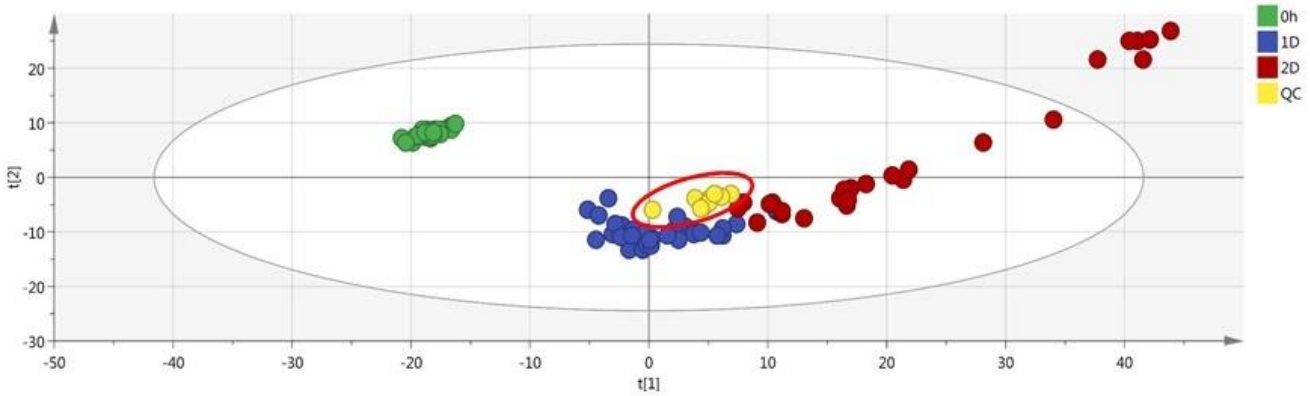


Figure 1. ESI + PCA Scores plot of the samples (n=107) after features filtering. Leftmost area dots (0h): fresh samples; 1D dots: “1 day” samples; 2D dots: “2 days” samples; QC dots: “QC samples” (circled). Explained variance of the first two PCs: 55.6%

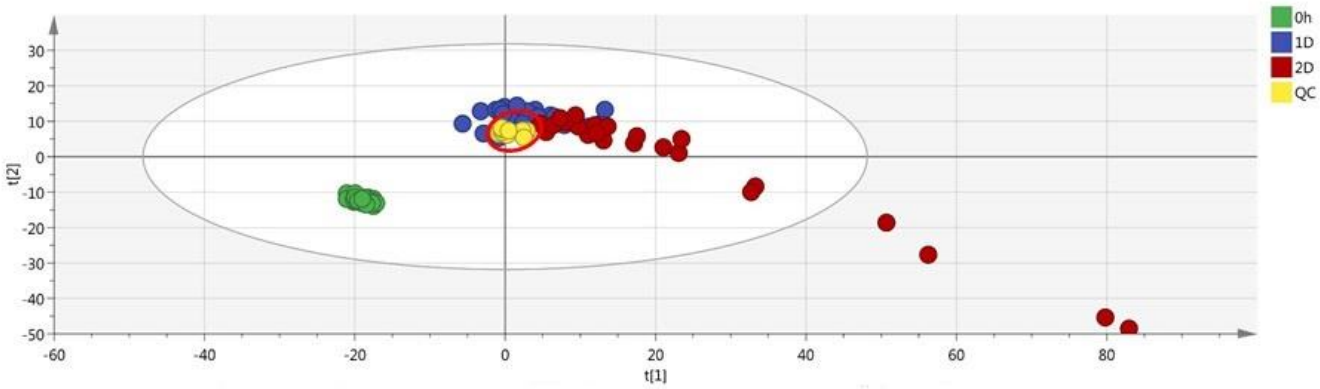


Figure 2. ESI - PCA Scores plot of the samples (n=107) after features filtering. Leftmost area dots (0h): fresh samples; 1D dots: “1 day” samples; 2D dots: “2 days” samples; QC dots: “QC samples” (circled). Explained variance of the first two PCs: 55.9%

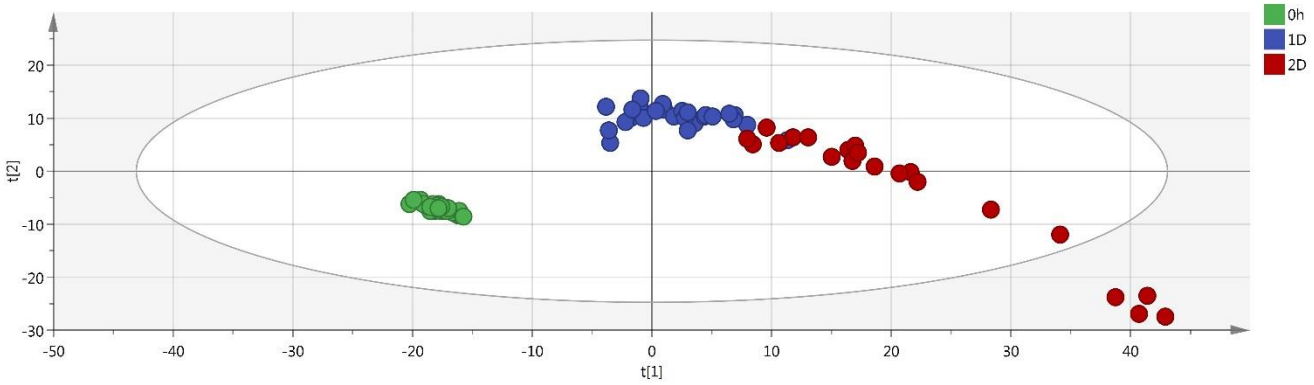


Figure 3. Final ESI + PCA Scores plot of the samples. Leftmost area dots (0h): fresh samples; 1D dots: “1 day” samples; 2D dots: “2 days” samples. Explained variance of the first two PCs: 57.3%

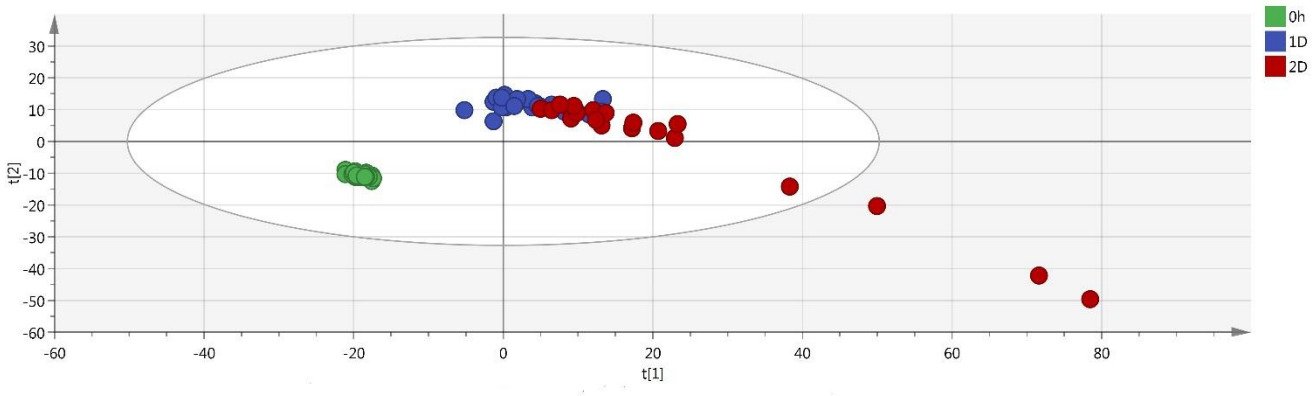


Figure 4. Final ESI - PCA Scores plot of the samples. Leftmost area dots (0h): fresh samples; 1D dots: "1 day" samples; 2D dots: "2 days" samples. Explained variance of the first two PCs: 58.2%

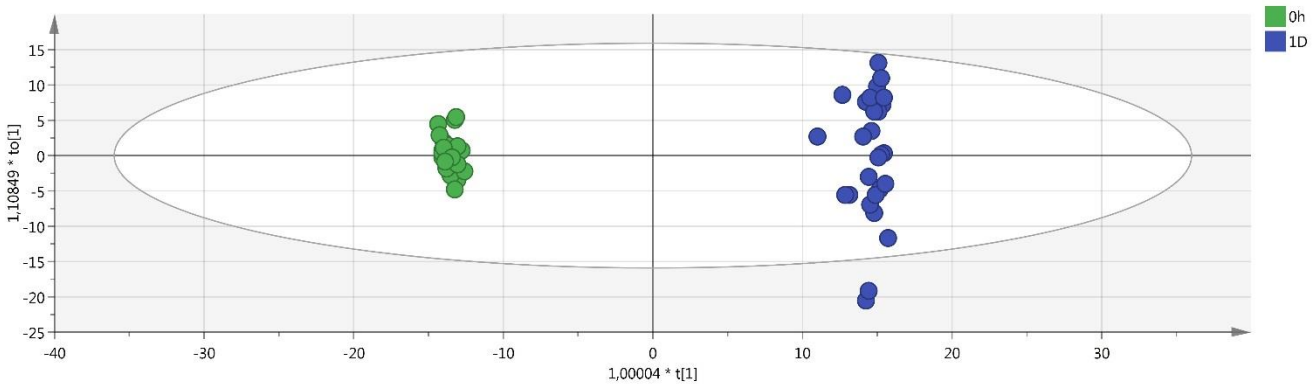


Figure 5. ESI + OPLS-DA Scores plot of the fresh samples against the "1 day" samples. Left area dots (0h): fresh samples; right area dots (1D): "1 day" samples

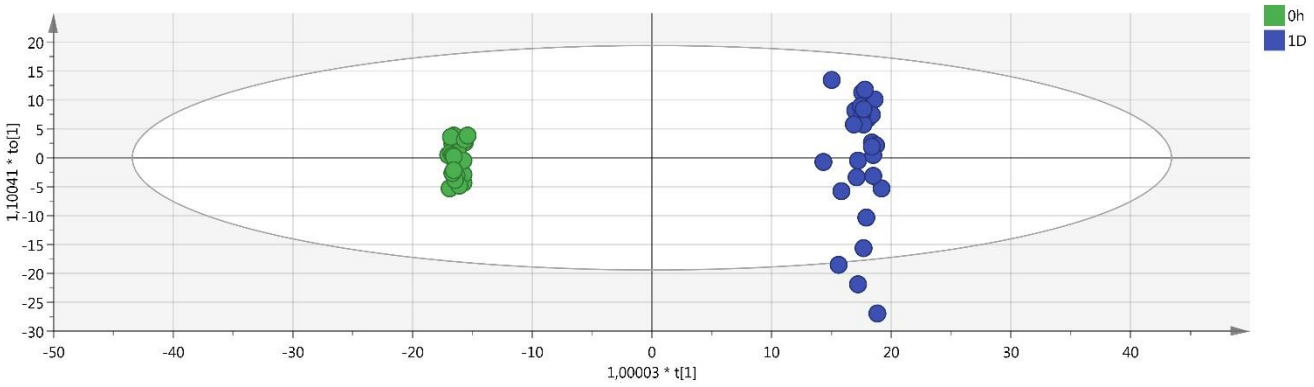


Figure 6. ESI - OPLS-DA Scores plot of the fresh samples against the "1 day" samples. Left area dots (0h): fresh samples; right area dots (1D): "1 day" samples

Bibliographic references

Akarachantachote, N., Chadcham, S., & Saithanu, K. (2014). Cutoff Threshold of Variable Importance In Projection For Variable Selection. *International Journal of Pure and Applied Mathematics, 94(*3*),* 307-322.

Aliakbarzadeh, G., Sereshti, H., & Parastar, H. (2016). Pattern recognition analysis of chromatographic fingerprints of Crocus sativus L. secondary metabolites towards source identification and quality control. *Analytical and Bioanalytical Chemistry, 408*(12), 3295-3307.

Balog, J., Perenyi, D., Guallar-Hoyas, C., Egri, A., Pringle, S., Stead, S., . . . Takats, Z. (2016). Identification of the Species of Origin for Meat Products by Rapid Evaporative Ionization Mass Spectrometry. *Journal of Agricultural and Food Chemistry, 64*(23), 4793-4800.

Collins, T., Zweigenbaum, J., & Ebeler, S. (2014). Profiling of nonvolatiles in whiskeys using ultra high pressure liquid chromatography quadrupole time-of-flight mass spectrometry (UHPLC-QTOF MS). *Food Chemistry, 163*, 186-196.

(s.d.). *D.L. no. 65, February 1993 Italian Legislation reception of European Union.*

Figueira, J., Câmara, H., Pereira, J., & Câmara, J. (2014). Evaluation of volatile metabolites as markers in Lycopersicon esculentum L. cultivars discrimination by multivariate analysis of headspace solid phase microextraction and mass spectrometry data. *Food Chemistry, 145*, 653-663.

Godzien, J., Alonso-Herranz, V., Barbas, C., & Armitage, E. (2015). Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample. *Metabolomics, 11*(3), 518-528.

Hajslova, J., Cajka, T., & Vaclavik, L. (2011). Challenging applications offered by direct analysis in real time (DART) in food-quality and safety analysis. *Trends in Analytical Chemistry, 30*(2), 204-218.

Jandric, Z., Islam, M., Singh, D., & Cannavan, A. (2017). Authentication of Indian citrus fruit/fruit juices by untargeted and targeted metabolomics. *Food Control, 72*, 181-188.

Jandric, Z., Roberts, D., Rathor, M., Abrahim, A., Islam, M., & Cannavan, A. (2014). Assessment of fruit juice authenticity using UPLC–QToF MS: A metabolomics approach. *Food Chemistry, 148*, 7-17.

Kalogiouri, N., Alygizakis, N., Aalizadeh, R., & Thomaidis, N. (2016). Olive oil authenticity studies by target and nontarget LC–QTOF-MS combined with advanced chemometric techniques. *Analytical and Bioanalytical Chemistry, 408*, 7955-7970.

Khamis, M., Adamko, D., & El-Aneed, A. (2017). Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrometry Reviews, 36*(2), 115-134.

Li, J., Zhu, S., Jiang, S., & Wang, J. (2017). Prediction of egg storage time and yolk index based on electronic nose combined with chemometric methods. *LWT - Food Science and Technology, 82*, 369-376. doi:10.1016/j.lwt.2017.04.070

Lin, H., Zhao, J., Sun, L., Chen, Q., & Zhou, F. (2011). Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis. *Innovative Food Science & Emerging Technologies, 12*(2), 182-186. doi:10.1016/j.ifset.2011.01.008

Reg. CE 853/2004. (s.d.).

Riedl, J., Esslinger, S., & Fauhl-Hassek, C. (2015). Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Analytica Chimica Acta, 885*, 17-32.

Righetti, L., Rubert, J., Galaverna, G., Hurkova, K., Dall'Asta, C., Hajslova, J., & Stranska-Zachariasova, M. (2018). A novel approach based on untargeted lipidomics reveals differences in the lipid pattern among durum and common wheat. *Food Chemistry, 240*, 775-783.

Rubert, J., Zachariasova, M., & Hajslova, J. (2015). Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review. *Food Additives & Contaminants: Part A, 32*(10), 1685-1708.

Schymanski, E., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology, 48*, 2097-2098.

Silva, P., Freitas, J., Silva, C., Perestrelo, R., Nunes, F., & Câmara, J. (2017). Establishment of authenticity and typicality of sugarcane honey based on volatile profile and multivariate analysis. *Food Control, 73*, 1176-1188.

Suman, M., Riani, G., & Dalcanale, E. (2007). MOS-based artifical olfactory system for the assessment of egg products freshness. *Sensors and Actuators B: Chemical, 125*(1), 40-47. doi:10.1016/j.snb.2007.01.031

Sumner, L., Amberg, A., Barrett, D., Beale, M., Beger, R., Daykin, C., . . . Viant, M. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics, 3*, 211–221. doi:10.1007/s11306-007-0082-2

(2016). *U.S. Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection.*

Wang, Q., Zhou, K., Wang, C., & Ma, M. (2015). Egg Freshness Detection Based on Hyperspectral Image Technology. *Advance Journal of Food Science and Technology, 7*(8), 652-657. doi:10.19026/ajfst.7.1623

Yimenu, S., Kim, J., & Kim, B. (2017). Prediction of egg freshness during storage using electronic nose. *Poultry Science, 96*(10), 3733-3746. doi:10.3382/ps/pex193

Zhao, J., Li, H., Chen, Q., Huang, X., Sun, Z., & Zhou, F. (2010). Identification of egg's freshness using NIR and support vector data description. *Journal of Food Engineering, 98*(4), 408-414. doi:10.1016/j.jfoodeng.2010.01.018

Author Contributions

Conceptualization: D.Cav., C.D.,M.S.; study execution: D.Cav., D.Cat.; writing-original draft preparation: D.Cav.; writing -review and editing: all authors; supervision. C.D., M.S

Egg products: industrial benefits and future perspectives

Rapid approach

The model developed with the GC-IMS instrument for a rapid freshness evaluation was tested for approximately six weeks in the quality control laboratories, analyzing 70 batches.

The aim was compare the results obtained with this new approach with the ELISA immunoassays quantification of Lactic Acid, that is the typical method used in the production sites: the cost of the analysis is 10 time higher and, according to a risk assessment evaluation, it is executed only on 30% of the batches weekly received.

According to the GC-IMS results six batches were at the limit of the acceptance threshold, with only 70% of affinity to the "fresh" group.

Due to the reduced amount of samples analyzed with the ELISA approach, only one of these six batches was controlled with this technique and the Lactic Acid amount was largely below the legal limit. The other five samples would have been used in the plants without any other control.

This study highlighted the advantages that this new technique could provide in a routine approach: all the batches could be analyzed with an increase in method sensitivity and a reduced cost per samples. A comparison of the two approaches is presented in table 1.

	ELISA APPROACH	NOVEL APPROACH
BATCHES WEEKLY ANALYZED	30%	100%
ANALYSIS TIME	30 minutes per sample	40 minutes per sample
TECHNICIAN TIME SPENT IN SAMPLE PREPARATION	25 minutes per sample	2 minutes per sample
ANALYSIS COST	≈ 5/10 euro per sample	≈ 50 cent per sample

Table 1: comparison of the "ELISA" and the "GC-IMS" approaches for egg products freshness evaluation

Confirmatory study

The trends of the identified markers was studied for 10 batches of egg products stored at 2-8°C and analyzed at different time points (5, 8 ,12 ,15 ,20 and 22 days of storage). The most interesting

compound seems the “Guanosine 5'-diphospho-D-Mannose or Glucose”, which after only five days of storage in the fridge showed a dramatic decrease in its intensity, as presented in figure 1.

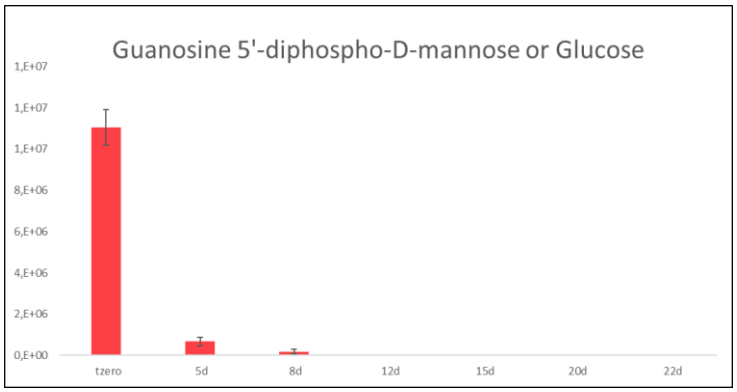


Figure. 1: target compound average area values (± Standard Error) through the time points

Further studies will be executed in order to confirm this trend that is really promising.
In addition, a novel targeted method will be developed by a third part laboratory, with the aim to set a concentration limit for each of the 31 “freshness” and “not freshness” markers identified.

CHAPTER 2

EXTRA VIRGIN OLIVE OIL

Extra Virgin Olive Oil

General overview of the product

Extra Virgin Olive oil (EVOO) is the core of the Mediterranean diet. It is the olive oil with the highest value and, as the other virgin olive oils, it is obtained from the fruit of olive tree using mechanical and physical approaches that do not alter its properties. The only treatments allowed are washing, decantation, centrifugation and filtration (Garcia-Gonzalez, Aparicio, & Aparicio-Ruiz, 2018).

Different chemical characteristics should be satisfied in order to classify a virgin olive oil as "extra virgin": according to the IOC (International Olive Council) rules, it is "*a virgin olive oil which has a free acidity, expressed as oleic acid, of not more than 0.8 grams per 100 grams, and the other characteristics of which correspond to those fixed for this category in this standard*" (IOC, 2016).

In addition to the chemical characteristics, also the organoleptic properties are an important parameter; a virgin olive oil can be considered "extra virgin" only if the median of defect is 0 and the median of the fruity attribute is higher than 0 (IOC, 2016).

Product Identity

Current standards of identity or related legislation (Codex, EU, US, ISO, Trade associations…)

Olive oils have to comply with different rules and standards depending on where they are traded (Conte, et al., 2019); the three most important regulations are:

Reg (EEC) 136/66 - a regulation for the establishment of a common organisation of the market of fats and oil. It details different type of oils but does not report any analytical parameter that should be satisfied.

Reg (EEC) 2568/91 – it details all the quality and purity parameters of the different oil categories, with a description of the analytical methods that have to be used.

COI/T.15/NC No 3/Rev. 11 – The regulation that details all the criteria that allow the classification of different type of olive oils outside the EU area.

Authenticity issues

Most of the common adulterations that were executed in the past (i.e. the addition of refined oils to virgin olive oils) are now easily detected with the official chemical parameters required by the legislation; however, other kind of illegal activities are sometimes put in place by dishonest farmers. The **most risky factors** that can have impact on the EVOO authenticity are:

- **Addition of soft-deodorized virgin olive oil to extra virgin olive oil**

  The soft deodorization is a thermal process at low temperature (<100°C). It is done in order to remove sensory defects but, according to the legislation, this product can no longer be considered "virgin" (Garcia-Gonzalez, Aparicio, & Aparicio-Ruiz, 2018). For this reason, mixtures of pure EVOO with this deodorized samples are illegal, but an official analytical strategy is not available up to now (Aparicio-Ruiz, Romero, García-González, Oliver-Pozo, & Aparicio, 2017) and this is the reason why this activity is one of the most diffused fraud.

- **Geographical provenance**

  The origin is sometimes presented as an added value of the product but for the oils there are no standard methods that are able to assess this parameter.

Aim of the work

In the following chapters, a rapid approach and a non-targeted method will be presented as two parallel new tools able to identify the fraudulent addition of soft refined virgin olive oils to extra virgin olive oils.

The first chapter was submitted to "*Food Analytical Methods"*, while the second was submitted to "*Journal of Agricultural and Food Science"*. For additional details see section "Author".

Bibliographic references

Aparicio-Ruiz, R., Romero, I., García-González, D., Oliver-Pozo, C., & Aparicio, R. (2017). Soft-deodorization of virgin olive oil: Study of the changes of quality and chemical composition. *Food Chemistry, 220*, 42-50.

Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., . . . Gallina Toschi, T. (2019). Olive oil quality and authenticity: A review of current EU legislation, standards, relevant methods of analyses, their drawbacks and recommendations for the future. *Trends in Food Science & Technology (in press)*.

Garcia-Gonzalez, D., Aparicio, R., & Aparicio-Ruiz, R. (2018). Olive oil. In J. Morin, & M. Lees, *FoodIntegrity Handbook: A guide to food authenticity issues and analytical solutions* (p. 336-357). Nantes.

IOC, I. O. (2016). *COI/T.15/NC No 3/Rev. 11 - Trade standard applying to Olive Oils and Olive-Pomace Oils.* Madrid, Spain.

Rossi, M. (1998). Proprietà funzionali degli ovoprodotti. *Rivista di Avicoltura, 67*, 28-34.

Head Space-based profiling techniques for screening purposes to assess authenticity issues of the EVOO sector

Tito Damiani[1#], Daniele Cavanna[1,2#], Andrea Serani[3], Chiara Dall'Asta[1], Michele Suman[2]*

[1]Department of Food and Drug, University of Parma, Parma, Italy
[2]Advanced Research Laboratory, Barilla G. e R. Fratelli S.p.A., Parma, Italy
[3]COTECA Srl Consulenze Tecniche agroindustriali, Pisa, Italy

[#] T.D. and D.C. equally contributed to this study.

Submitted to "Food Analytical Methods"

Introduction

Olive oil is part of the Mediterranean culture and due to its unique nutritional and organoleptic properties, is one of the most appreciated and expensive edible oils (Barjol, 2013). The growing consumption and high price represent explicit fraud drivers for potential offenders (van Ruth, Huisman, & Luning, 2017).

The olive oil sector is one of the most regulated markets worldwide, due to the coexistence of international (International Olive Council, Codex standards) and European standards (European Economic Community). Moreover, in the last decade, the globalization of this sector led to new national legislation in the emerging olive oil producing countries (i.e. USA, Australia, California, New Zealand). In spite of slight differences, all these regulations basically group the olive oils in different quality grades, based on specific chemical and organoleptic criteria (Bajoub, Bendini, Fernández-Gutiérrez, & Carrasco-Pancorbo, 2018). According to the European Regulation No. 2568/91, and subsequent amendments, "virgin olive oil" (VOO) must be obtained using solely mechanical or other physical means, in conditions which do not alter the oil in any way. It has not undergone any treatment other than washing, decanting, centrifuging, and filtering (EU, 1991). Within the VOO category, "extra virgin olive oil" (EVOO) represents the top-quality grade and, therefore, it is more vulnerable to adulteration. In the other hand, oil which does not fulfil the minimum VOO quality criteria, named "lampante olive oil" (LOO), is not intended for direct consumption but rather to technical use or refining process.

Although the olive oil sector is highly regulated in Europe, many problems related to possible fraud are still open, and fats and oils, including olive oils, are ranked third in the 2018 EU Food Fraud

Network report on non-compliances per product category. Among them, the large majority of non-compliant products were affected by mislabelling (59%), followed by replacement/dilution/addition/removal (18%), and manipulated documents (17%). To further complicate the European olive oil situation, over the last years Europe has endured poor olive harvests. Low harvests tend to encourage further manipulation of olive oil in the form of adulteration with non-olive oils and/or use of damaged or over-ripe olives. This results in olive oil that does not meet the classification standards of extra virgin olive oil.

One of the major issue reported for olive oil, was the blending of EVOO with refined olive oil, often obtained by soft refinement procedures based on the use of mild technologies.

Within soft refinement process, the lower temperatures do not seem to produce great chemical modification in the bulk and it stands to reason that malefactors employ these soft-refined olive oils (SROO) to create mixtures no longer detectable by regular methods (Garcia, Martins, & Cabrita, 2013) (Aparicio-Ruiz R. , Romero, García-González, Oliver-Pozo, & Aparicio, 2017).

Soft-deodorization consists of a steam distillation, carried at lower temperature respect to the conventional process (i.e. 100°C instead of 180-200°C) (Aparicio-Ruiz R. , Romero, García-González, Oliver-Pozo, & Aparicio, 2017). Deacidification is designed to remove the free fatty acids (FFA), which speed up the oxidation process and are involved in the development of rancid flavour. It is normally achieved through the addition of alkali (such as sodium hydroxide) to the oil. This results in the precipitation of the FFA as a insoluble soap dreg, subsequently removed by centrifugation and/or filtration (Vaisali, Charanyaa, Belur, & Regupathi, 2014).

The main changes due to soft refining are expected to occur in the volatile fraction, rather than the bulk. Therefore, targeting volatile compound profile may be a successful approach to highlight possible addition of refined oil to EVOO. Among methodologies for fingerprinting volatiles in food, headspace (HS) based techniques, such as gas-chromatography ion mobility spectrometry (GC-IMS) and gas chromatography electronic nose (GC-Enose), present several advantages for industrial use, i.e. time- and cost-affordability.

IMS basically involves the ionization of aroma compounds and subsequent gas-phase separation by their relative mobility under a weak electric field at atmospheric pressure. Originally, it has been widely employed in the rapid detection of explosives and chemical warfare agents but, nowadays, it is an emerging technology in food control (Eiceman, 2002) (Garrido-Delgado, et al., 2011) (Zhang L. , et al., 2016). In order to increase its selectivity, IMS can be coupled with a prior GC separation (Contrerasa, Arroyo-Manzanaresa, Arce, & Arce, 2019). E-nose is intended to mimic the human olfactory system and consists of an array of non-selective sensors which, through the interaction with

the volatile organic compounds (VOCs), produce electronic signals (Ye, et al., 2016). Similar to the GC-IMS, E-nose has been frequently coupled to a GC separation in order to achieve a selectivity improvement (Melucci, et al., 2016) (Rottiers, et al., 2019).

The present work focused on the capability evaluation of the two aforementioned techniques as untargeted rapid screening tools for the detection of EVOOs adulteration with SROO.

Similar works were carried out in this field, reaching remarkable results (Garrido-Delgado, Arce, & Valcárcel, Multi-capillary column-ion mobility spectrometry: a potential screening system to differentiate virgin olive oils, 2012) (Garrido-Delgado, Dobao-Prieto, Arce, & Valcárcel, 2015). Our attempt was to move a step forward by getting a preliminary evaluation on the sensitivity towards soft refining of these analytical approaches. Therefore, SROO produced in-house under control conditions were mixed, at various percentage, with authentic EVOOs, reproducing a potential counterfeiting action. To the best of our knowledge, this is the first study focused on EVOOs blends with SROO, coupling the two above-mentioned screening techniques.

Materials and Methods

Chemicals

For the soft-refining process, Sodium Hydroxyde and Sodium Sulphate were purchased from VWR International, Ltd (Poole, United Kingdom), while water was purified using a Milli-Q system.

Samples set

A total of 43 commercial EVOO were collected, 40 of which were from different Italian regions, while the remaining 3 were declared to come from the EU. Three harvesting seasons were included in the sample collection: 2015/2016 ($n = 18$), 2016/2017 ($n = 8$) and 2017/2018 ($n = 17$). Henceforth, the EVOOs group will be referred as 15/16, 16/17 and 17/18, respectively. All the samples were stored in the dark and at 4 °C until the analysis. Soft-deodorization and deacidification were carried out on commercial VOOs and LOOs. Both the refining processes were replicated on different oils in order to check the procedure repeatability. At the end, 3 soft-deodorized (DEO) and 2 deacidified (DEA) oils were obtained. Furthermore, two aliquots of the DEA oils underwent also to the soft-deodorization process, gaining two olive oil samples both deacidified and soft-deodorized (DEA+DEO). The official EVOO physico-chemical quality parameters (EU E. U., 2016) were analysed in these refined oils.

Thereafter, 21 illegal mixtures were prepared, at different percentage (from 10 to 60%), by blending the so-obtained refined oils with EVOOs randomly chosen from the sample set. A detailed sample

list is provided in the Supplementary material (Table S1). Moreover, 2 commercial EVOOs suspected to be frauded with SROOs, were included in the sample set.

Soft Deodorization and deacidification

In authenticity studies, a large number of authentic and adulterated samples are needed for the development and validation of a reliable and robust classification model. In this regard, the sample collection is often the most demanding part of the study. Finding adulterated samples from the market is not an easy task and, to overcome this problem, many researchers tried to recreate the fraud practice at laboratory scale (Cavanna, Righetti, & Elliot, 2018). Accordingly, a soft-refining process was in-house reproduced and applied to non-EVOO samples (i.e. virgin and lampante olive oil). Afterwards, illegal blends were prepared by mixing commercial EVOOs and the so-obtained SROOs.

The soft-deodorization process was carried out partially modifying the protocol proposed by Aparicio-Ruiz et al. (Aparicio-Ruiz R. , Romero, García-González, Oliver-Pozo, & Aparicio, 2017). The oil sample (400 mL) was introduced into a triple-neck round-bottom flask containing a magnetic stir bar. Two necks were used to connect, respectively, the vacuum tubing and a dropping funnel containing Milli-Q grade water. A thermometer was introduced in the left neck in order to constantly check the system temperature and, finally, the apparatus was placed on a magnetic stirring hotplate (VELP Scientifica, Italy) operating at 300 rpm. In order to start the distillation, 0.2 mL of water were added when the oil temperature reached an initial temperature of 60 °C. Thereafter, the temperature was raised to 100 °C and maintained for 60 minutes. Once the initial water was evaporated, 0.2 mL more were added to carry on the distillation process. The time/temperature parameters were chosen according to the results achieved by Aparicio-Ruiz et al. (Aparicio-Ruiz R. , Romero, García-González, Oliver-Pozo, & Aparicio, 2017). At the end of the procedure, the sample was cooled down and stored in the dark at 4 °C until the analyses.

For the deacidification, 600 mL of defective oil were weighted into a beaker, and the latter placed on a magnetic stirring plate (VELP Scientifica, Italy) operating at room temperature. The neutralization was achieved adding NaOH and Na_2SO_4 directly to the oil. The volume of NaOH to add was computed according to the following formula while the amount of Na_2SO_4 was twice the so-calculated value.

$$NaOH\ (mL) = \left[\frac{(A-B)\cdot C\cdot D}{(\frac{E}{100})}\right]\cdot\left[\frac{(\frac{100}{F})}{G}\right]$$

Where:

- A and B are, respectively, the starting and intended value of FFA% of the oil (expressed as percentage of oleic acid equivalent). In the present work $B = 0.4\%$;
- C is the sample weight (g);
- D and E are, respectively, the NaOH and oleic acid molecular weight ($D = 40\ g/mol;\ E = 282\ g/mol$);
- F and G are, respectively, the concentration and the density of the NaOH solution concentration ($F = 50\%,\ G = 1.52$ g/mL);

After the addition of the salts and 5 minutes of stirring, the sample was centrifuged (5000 rpm) and the supernatant (neutralized oil) collected and stored at 4 °C before the analyses.

An aliquot of the so-obtained deacidified oil underwent to the soft-deodorization process described above.

Instrumental Analysis

The GC-IMS analyses were performed on an IMS commercial instrument (FlavourSpec® - G.A.S., Dortmund, Germany), equipped with a 2.5 mL Hamilton syringe, an automatic headspace sampling unit and a gas chromatograph.

For the analysis, 0.2 g of olive oil was placed into a 20 mL headspace vial and incubated at 60 °C. After 10 minutes of incubation, 500 µL of HS sample were sampled and injected into the GC–IMS equipment, with a speed rate of 30 mL min^{-1} and a syringe temperature of 80 °C. The GC was equipped with a SE-54-CB-1 capillary column (5%phenyl-1%vinyl-94% methylpolysiloxane), 15 m × 0.53 mm × 1 µm film thickness. The separation was achieved in isothermal conditions (40 °C), using N_2 as carrier gas with a flow ramp starting with 2 mL min^{-1} for 5 min, increasing up to 70 mL min^{-1} in 20 min and finally maintaining 70 mL min^{-1} for 5 min. The total GC runtime was 30 min (the eluted analytes were driven into the ionization chamber prior to the IM separation and detection (instrumental details in Supplementary Table 2).

Flash GC-Enose (FGC E-nose) analysis were performed by means of Heracles II (Alpha MOS, Toulouse, France). The instrument was equipped with two parallel GC columns: a non-polar column (MXT5: 5% diphenyl, 95% methylpolysiloxane, 10 m x 0.180 mm x 0.4 □m film thickness) and a slightly polar column (MXT1701: 14% cyanopropylphenyl, 86% methylpolysiloxane, 10 m x 0.180 mm x 0.4 □m film thickness). For the analysis, 0.2 g of olive oil was placed into a 20 mL HS vial. Prior to the injection, each vial was incubated for 20 min at 60 °C in a shaker oven. A HS volume of 5 mL was automatically sampled and injected in the GC-Enose equipment, using H_2 as carrier gas.

In order to remove excess air and moisture, and concentrate the analytes, the sampled volume was adsorbed on a CARBOWAX trap (40 °C for 65 s) prior to the chromatographic separation. Afterwards, the trap temperature was increased up to 240 °C in 30 s, and the volume driven into the chromatographic unit. The total chromatographic separation time was 110 s, employing the following chromatographic temperature programs: starting temperature of 40 °C (kept for 2 s), increasing up to 270 °C (at 3 °C s^{-1}) and then the final temperature was held for 21 seconds. A flame ionization detector (FID), operating at 270 °C, was present at the end of each column, thus recording two chromatograms for each sample.

All the GC-IMS and GC-Enose analysis took place over a 7-months period (from October 2017 to May 2018), each sample was analysed in duplicate for the GC-IMS and in triplicate for the GC-Enose. The replicates' average was used for the data processing.

Data treatment and statistics

GC-IMS and FGC E-nose data

The GC-IMS output consists of a 3D array, graphically represented in a topographic plot. Due to the twofold separation, the analytes occur as 2D-signals (spots) characterized by a chromatographic RT (y axis), an ion mobility Drift Time (x axis) and the intensity values represented by a colour scale. LAV software was used to visualize the data (i.e. topographic plots), pinpoint the relevant spots and manually box the map areas to carry the integration on, taking the peak height as intensity value (Cavanna, Zanardi, Dall'Asta, & Suman, 2019). GC-IMS spectra alignment respect to the reactant ion peak (RIP) position was automatically performed by the software and, after autoscaling, the obtained dataset underwent to chemometric computation.

FGC-Enose is equipped with two GC columns, and the respective FID, working in parallel. Therefore, two different chromatograms are provided by a single run. As proposed by Melucci et al., the obtained GC traces were aligned and arranged consecutively. All the detected peaks above a established intensity threshold were selected and automatically integrated by the Alphasoft software (Melucci, et al., 2016). Logarithmic transformation (Log10), followed by autoscaling, was found to be the optimal data preprocessing combination and applied to the dataset prior to the multivariate data analysis (MDA).

Multivariate data analysis

In order to compare the performance of the employed devices, GC-IMS and FGC E-nose datasets underwent to chemometric analysis, separately. Data visualization and exportation was carried out by means of the commercial software LAV 2.2.1 (G.A.S., Dortmund, Germany) and Alphasoft V14

(Alphamos, Tolouse, France), respectively for GC-IMS and FGC E-nose instruments. The obtained datasets were exported and processed using Matlab R2019a (The Mathworks Inc., Natick, MA, USA, 2007) and PLS Toolbox 8.7 (Eigenvector Research, Inc., Manson, WA, USA). Principal component analysis (PCA) was used for the preliminary data exploration in a reduced dimension space. Afterwards, partial least squares discriminant analysis (PLS-DA) algorithm was calculated in order to maximize groups separation and develop a classification model.

Results and discussion

The main purpose of the present study was to test the employed HS-based techniques in the detection of EVOOs adulteration with SROO. In order to overcome the lack of real counterfeited samples, a soft-refining process was recreated at a laboratory scale and applied to non-EVOOs. Illicit blends were prepared by mixing EVOO samples with decreasing amount (up to 10%) of SROO.

GC-IMS is a rapid and easy-to-use technology designed for in-line quality control testing (Vautz, et al., 2011). The data processing can be performed by means of the LAV software, manually boxing the spots considered to be relevant and exporting the resulting dataset. Accordingly, 52 markers were selected and included in the dataset (73 objects x 52 variables) for the following MDA. A detailed list of the considered spots is reported in Supplementary Material (Supplementary Table 3). However, the above-described approach might present weakness in term of reliability, since it is affected by the operator's subjectivity in the marker selection. In this regard, a more systematic method already proposed in literature was also tested and compared (Garrido-Delgado, Arce, & Valcárcel, 2012). Briefly, a prior data preprocessing, consisting of baseline correction, Savitzky-Golay smoothing and peak alignment using the spline interpolation algorithm, was performed (Gerhardt N. , Birkenmeier, Sanders, Rohn, & Weller, 2017). Afterwards. the three-dimensional GC-IMS data were unfolded and the IMS spectra of each samples arranged consecutively in a row, thus obtaining a final matrix of 73 objects x 1698780 variables. Basically, the whole informative spectrum region was kept and used as sample's fingerprint. Comparable and consistent results were achieved by the described data processing approaches and, actually, no significant improvements were provided by the second strategy (data not shown). It should be pointed out that the latter is more demanding in term of time, computing power and, besides, operator's expertise, thus less suitable for industrial implementation.

Electronic olfaction is a time- and labour-saving technology intended to supersede human olfactory sense in flavour analysis (Majchrzak, Wojnowski, Dymerski, Gębicki, & Namieśnik, 2018). Similarly to GC-IMS, a fingerprint approach was adopted by merging the two chromatograms obtained for each sample. The aim was to reduce the risk of discarding useful information by selection of relevant RT

ranges and possible errors during the automatic peak-area integration. On the other hand, it must be taken into account that the same analytes could generate peaks in both chromatograms, leading to redundant information in the dataset (Melucci, et al., 2016). At the end, 74 chromatographic peaks were found to be significant (73 objects x 74 variables).

As expected, strong differences were found between GC-IMS topographic plots of EVOOs and SROOs, with all the VOC's eluting between 1.89 and 14.1 minutes, and drift time ranging from 7.8 to 15.35 ms. Likewise, GC-Enose chromatograms of EVOOs showed a larger number of peaks (higher in intensity) respect to the SROOs traces (Fig. 1). This is not surprising as almost all the VOC's are supposed to be stripped out during the deodorization process.

As noteworthy aspect of the present work, all the analyses were carried out in a 7-months period, thereby covering the equipment fluctuations over the time, which is critical for fingerprinting method development. Indeed, in a manufacturing quality control context, it is reasonable that samples collection and analysis take place within a long-term monitoring programme. External influences (i.e. environmental condition. instrumental drift. etc.) might play a relevant role in fingerprinting measures, therefore, the experimental design should account also for these sources of variation in order to avoid an overestimation of the method performance (Esslinger, Riedl, & Fauhl-Hassek, 2014).

In this regard, the GC-IMS measurement intra-day and inter-day repeatability were assessed, by analysing several replicates of the same reference sample within the same day and over the whole analysis period. The highest peak was taken as reference and the recorded intensity values used to calculate the relative standard deviation (RSD) (Garrido-Delgado, et al., 2011). The intra-day and inter-day repeatability were 1.3% and 4.7%, respectively. These results are slightly higher respect to what found by Garrido-Delgado et al., however, it should be pointed out that the present work was carried over a far longer period, i.e. 7 months. The same computation was performed on the FGC E-nose data, after Log10 transformation. FGC E-nose exhibited intra-day and inter-day repeatability of 2.16% and 3.36%, respectively.

Therefore, both the equipment showed excellent stability over the time, which is a great advantage for fingerprinting measures.

Unsupervised pattern recognition

EVOO is known to be a non-stable food matrix, prone to chemical changes such as lipid peroxidation, triacylglycerols hydrolysis, minor compounds degradation, etc. Indeed, EVOO saleability, in ordinary shelf-life conditions (i.e. room temperature and exposed to light), is hardly ever longer than

one year since the physico-chemical quality descriptors risk to be no longer compliant with the legislation (Gutierrez & Fernandez, 2002) (Di Serio, Giansante, Di Loreto, & Di Giacinto, 2018). For these reasons, the authors preferred to exclude 15/16 samples from the first steps of MDA.

Concerning the PCA, in both cases the first four PCs accounted for more than 60% of the total variation (68.78% and 62.71% for GC-IMS and FGC E-nose, respectively). Plotting PC1 versus PC2 resulted in a good clustering among EVOOs and the illicit mixtures, for both the techniques (Fig. 2). The SROO replicates (i.e. three replications of DEO, two of DEA and DEA+DEO) are fairly close to each other, which might denote that the refinement protocol was rather reproducible and properly performed. At a preliminary view, FGC E-nose objects are tighter clustered with sharper confidence ellipses, which might be related to the slight higher reproducibility. However, this aspect should be further investigated.

At this point, 15/16 sample were projected onto the PCA models generated in the previous stage. A clear trend can be observed among the authentic EVOO samples, which fell closer to the adulterated admixtures, according to the harvesting year (Fig. 2c and 2d). PCA scores plot generated using all the samples (i.e. including EVOOs 2015/2016) is reported in Supplementary Materials (Supplementary Fig. 1).

Notably, almost all the 15/16 samples are entirely located within the illicit blends' 95% confidence ellipse. Such a trend suggests that VOCs undergoes to significant changes, even in proper storage conditions, and the volatile fraction tends to "converge" with that of SROOs. This is a mere qualitative evaluation but, nevertheless, it might be cause for a broader consideration. In several works, samples from different harvesting seasons are collected, stored in proper condition and finally analysed in a short time period. However, to properly apply such approach. the foodstuff's chemical stability during the storage should be previously ensured, otherwise, such non-controlled variations might lead to weak results (Riedl, Esslinger, & Fauhl-Hass, 2015) (Cavanna, Righetti, & Elliot, 2018).

Supervised pattern recognition

PLS-DA was performed considering only the two classes of interest, namely illicit mixtures and EVOOs (16/17 and 17/18 samples modelled as single class). At first, the sample set was split in training and test set, used for model generation and external validation purpose, respectively. A subset of objects was selected to assemble a test set for the external model validation. The two suspected commercial frauded mixtures were included in the test set, together with four EVOOs selected using the Kennard-Stone algorithm.

As can be seen from the training set's PLS-DA scores plot (Fig. 3), excellent segregation was achieved regardless the analytical techniques. In both cases, the first two LVs were suitable to clearly separate the two groups and correctly assign all the samples to their belonging class, and even the mixtures at lower adulteration percentage were correctly classified. Therefore, the generated classification model was used to predict the authenticity of the test set samples. In both cases, all the objects were assigned to their actual class, achieving a 100% correct classification rate.

Furthermore, a rigorous model validation enlarging the external test set, has to be performed in order to avoid possible overfitting and method performance overestimation (Riedl, Esslinger, & Fauhl-Hass, 2015). However, the results shown herein confirmed the high potential of the employed techniques as rapid screening tools for the detection of SROO-EVOO blending and no particularly better performance was shown by a technique respect to the other.

Concerning the marker identification, the main goal of the present work was the development of an untargeted screening method, thus, it was not carried out. Nevertheless, a significant step forward would be the unambiguous elucidation of the discriminant markers by means of confirmatory techniques (i.e. GC-MS/GC-HRMS). Thereafter, a twofold analytical platform, based on both screening and confirmatory analysis, can be developed and validated according to legal standards (International Organization for Standardization, 2005).

Conclusions

Nowadays, fast and reliable analytical methods are highly claimed in the food industry for screening purposes (e.g. raw materials quality control). In the present work, the authors dealt with one of the main authenticity issues of the EVOO sector: the undeclared blending with SROO. To this end, two HS-based profiling techniques, namely GC-IMS and GC-Enose, were tested as alternative to the classical targeted approaches. Both the analytical techniques exhibited notable robustness and stability over the time, with FGC E-nose apparently slightly more reproducible than GC-IMS, but without a significant preference for one of the two approaches.

The required sample preparation was minimal, providing global information about VOCs profile in roughly 30 minutes. The chemometric processing aimed to discern the authentic EVOOs from the illegal admixtures with SROOs. A 100% correct classification rate was achieved by both the techniques and even the lower-percentage illicit mixtures were recognized to be non-authentic EVOO.

These results proved that both GC-IMS and FGC-Enose are fast and suitable tools for EVOOs authenticity testing even using non-target VOCs signals of the sample's HS fraction. Future

perspectives, are unambiguous identification of the discriminant markers in order to develop and validate a targeted method as confirmatory platform.

Figures

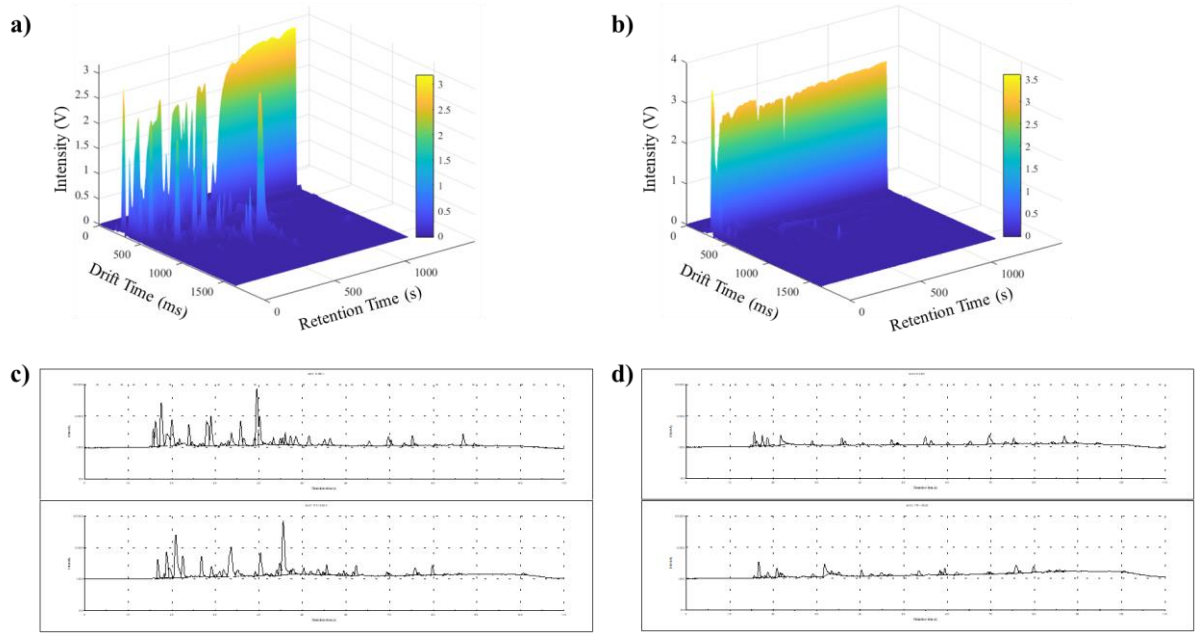


Figure 1: GC-IMS 3D topographic plot of EVOO (a) and soft deodorized oil (b); FlashGC chromatograms of EVOO (c) and soft deodorized oil (d);

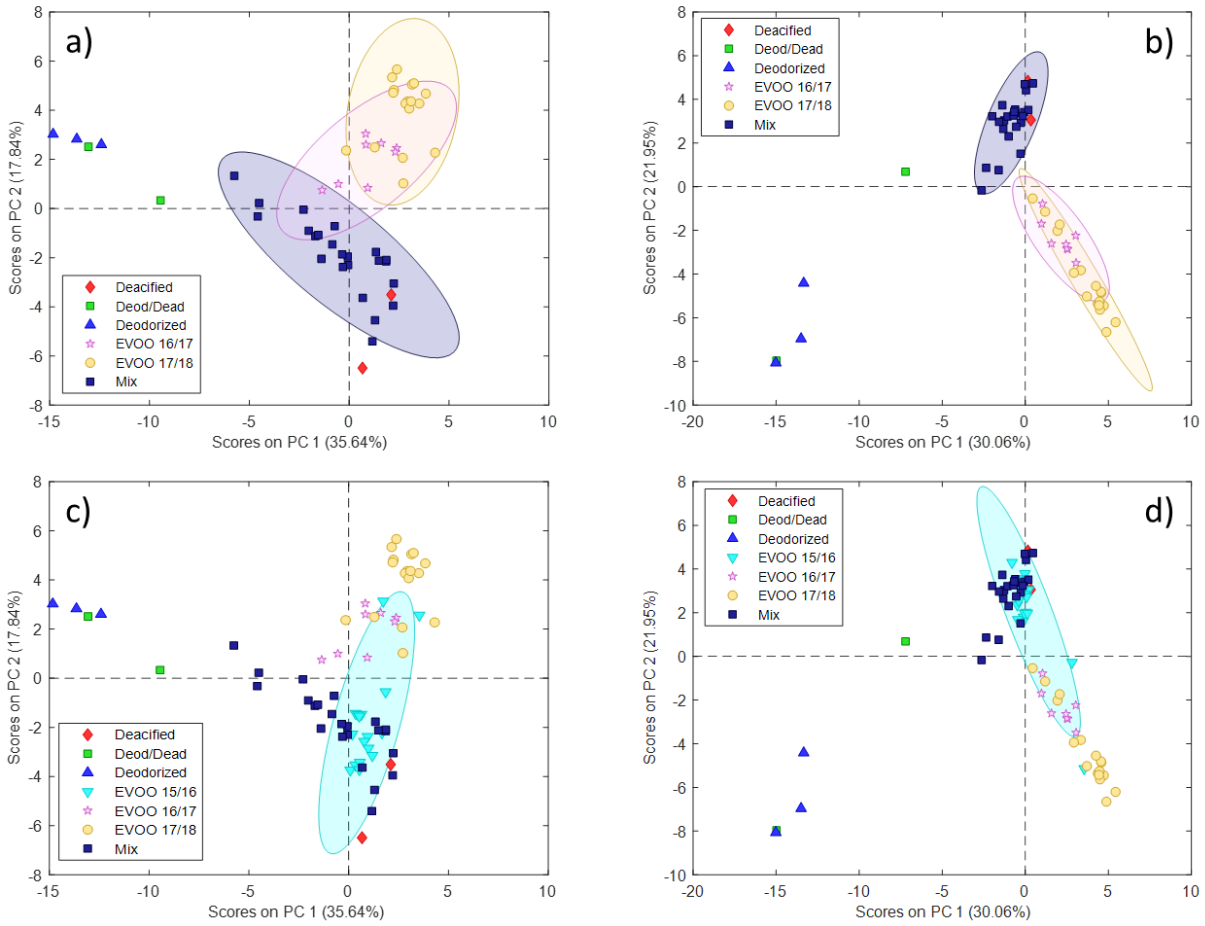


Figure 2: Score plots resulting from PCA performed on GC-IMS (a) and FGC E-nose (b) data, with 95% confidence ellipses for each group. Scores plot (c) and (d) include the projection of the 15/16 samples for the GC-IMS model and for the FGC E-nose model respectively

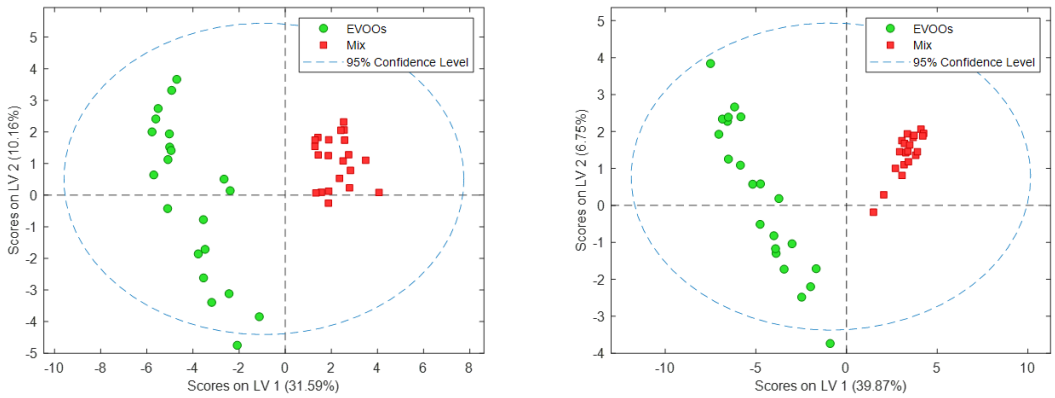


Figure 3: Score plots resulting from PLS-DA performed on GC-IMS (left) and FGC E-nose (right) data.

Supplementary Material

Category	Harvest Year	*n*
EVOO	2015/2016	18
EVOO	2016/2017	8
EVOO	2017/2018	17
Deodorized	/	3
Deacidified	/	2
Deodorized + Deacidified	/	2
Mix	/	21
Suspect frauded	/	2

Table S1: Sample list.

Drift tube length	**9.8 cm**
Drift gas flow rate	150 mL min^{-1}
Drift time	30 ms
Drift Voltage	5 kV
Drift tube temperature	45 °C
Ionization mode	Positive

Table S2: IMS instrumental parameters

Spot	RT min (min)	RT max (min)	DT min (ms)	DT max (ms)
Peak1	1.89	2.38	7.80	9.58
Peak2	2.32	2.69	8.78	9.99
Peak3	2.43	3.37	7.96	8.97
Peak4	2.59	2.91	8.54	9.31
Peak5	3.19	3.70	10.22	11.18
Peak6	3.24	3.76	8.41	9.32
Peak7	3.50	3.88	9.05	9.82
Peak8	3.52	4.75	7.98	9.06
Peak9	4.18	4.58	9.02	10.16
Peak10	3.98	4.44	10.37	12.35
Peak11	3.98	4.24	9.43	10.12
Peak12	5.04	5.46	9.69	12.14
Peak13	4.75	5.07	10.06	10.72
Peak14	4.75	5.10	10.27	11.09
Peak15	4.98	5.76	8.38	9.67
Peak16	5.96	6.36	9.10	10.77
Peak17	5.37	6.10	7.82	9.18
Peak18	5.94	6.21	11.29	12.84
Peak19	3.86	4.44	9.68	11.10
Peak20	6.08	6.18	10.38	11.46
Peak21	6.05	6.21	11.02	11.63

Peak22	6.24	6.44	10.36	11.53
Peak23	6.15	6.50	8.50	9.47
Peak24	6.37	6.58	8.85	10.40
Peak25	6.33	6.52	11.34	12.30
Peak26	6.37	6.54	12.23	12.79
Peak27	6.99	7.29	9.68	10.85
Peak28	6.97	7.14	11.58	12.18
Peak29	6.98	7.18	12.07	12.42
Peak30	6.98	7.38	12.20	12.81
Peak31	7.57	7.79	9.80	11.02
Peak32	7.92	8.42	8.45	10.40
Peak33	8.00	8.60	10.86	13.50
Peak34	8.58	9.07	9.77	11.34
Peak35	12.58	14.00	10.13	11.39
Peak36	8.28	8.73	9.46	10.63
Peak37	2.45	3.06	10.54	11.13
Peak38	7.38	7.95	12.90	13.38
Peak39	12.86	13.66	14.11	14.78
Peak40	6.57	6.80	9.63	10.40
Peak41	6.75	6.97	8.98	9.77
Peak42	6.58	6.85	10.42	11.64
Peak43	6.59	6.87	11.17	12.46
Peak44	6.60	6.85	12.66	13.49

Peak45	10.84	11.76	9.17	10.74
Peak46	2.90	3.20	9.53	10.30
Peak47	2.53	2.81	9.33	10.77
Peak48	7.54	7.99	11.29	12.36
Peak49	8.50	8.78	13.59	15.35
Peak50	8.64	9.00	12.38	14.20
Peak51	7.11	7.49	9.38	9.98
Peak52	3.48	3.80	10.45	11.52

Table S3*:* GC-IMS global area set integration parameters.

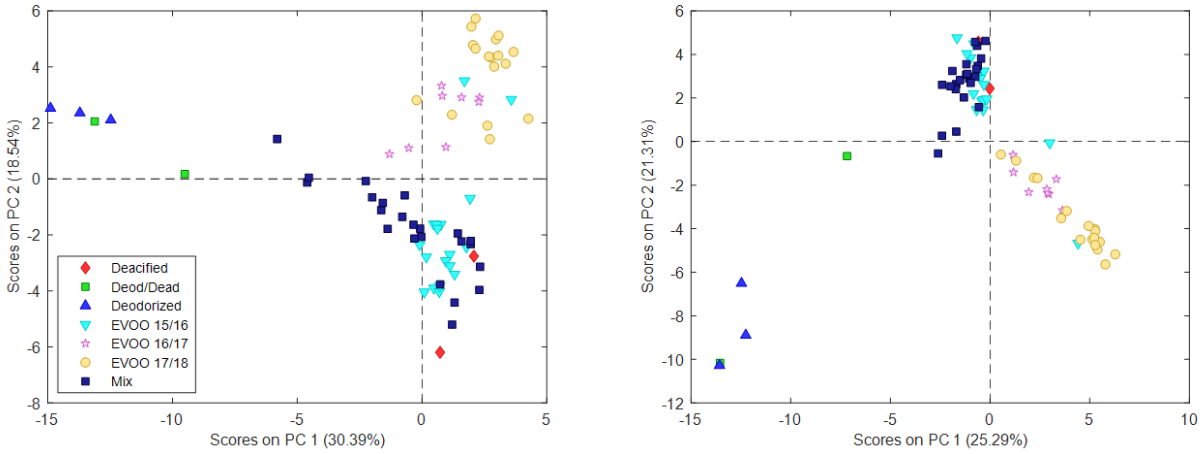


Figure S1: Score plots resulting from PCA performed on GC-IMS (left) and FGC E-nose (right) data of all the samples (i.e. EVOOs 2015/2016 included).

Bibliographic references

van Ruth, S., Huisman, W., & Luning, P. (2017). Food fraud vulnerability and its key factors. *Trends in Food Science & Technology*, *67*, 70-75.

Aparicio-Ruiz, R., Romero, I., García-González, D., Oliver-Pozo, C., & Aparicio, R. (2017). Soft-deodorization of Virgin Olive Oil: Study of the Changes of Quality and Chemical Composition. *Food Chemistry*, *220*, 42-50.

Bajoub, A., Bendini, A., Fernández-Gutiérrez, A., & Carrasco-Pancorbo, A. (2018). Olive oil authentication: A comparative analysis of regulatory frameworks with especial emphasis on quality and authenticity indices, and recent analytical techniques developed for their assessment. A review. *Critical Reviews in Food Science and Nutrition*, *58*(5), 832–857.

Barjol, J.-L. (2013). Introduction. In R. Aparicio, & J. Harwood, *Handbook of Olive Oil* (p. 1-17). Boston: Springer.

Cavanna, D., Righetti, L., & Elliot, C. (2018). The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach. *Trends in Food Science & Technology*, *80*, 223-241.

Cavanna, D., Zanardi, S., Dall'Asta, C., & Suman, M. (2019). Ion mobility spectrometry coupled to gas chromatography: A rapid tool to assess eggs freshness. *Food Chemistry*, *271*, 691–696.

Contrerasa, M., Arroyo-Manzanaresa, N., Arce, C., & Arce, L. (2019). HS-GC-IMS and chemometric data treatment for food authenticity assessment: Olive oil mapping and classification through two different devices as an example. *Food Control*, *98*, 82-93.

Di Serio, M., Giansante, L., Di Loreto, G., & Di Giacinto, L. (2018). Shelf life of extra- virgin olive oils: First efforts toward a prediction model. *Journal of Food Processing and Preservation*, 13663.

Eiceman, G. (2002). Ion-mobility spectrometry as a fast monitor of chemical composition. *Trends in Analytical Chemistry*, *21*(4), 259-275.

Esslinger, S., Riedl, J., & Fauhl-Hassek, C. (2014). Potential and limitations of non-targeted fingerprinting for authentication of food in official control. *Food Research International*, *60*, 189-204.

EU. (1991, July 11). Official Journal of the European Communities. *Commission Regulation (EEC) No. 2568 /91 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis*.

EU, E. U. (2016). Commission Delegated Regulation (EU) 2016/2095 of 26 September 2016 amending Regulation (EEC) No 2568/91 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis. *Official Journal Of the European Union*, p. L326/1-6.

Garcia, R., Martins, N., & Cabrita, M. (2013). Putative markers of adulteration of extra virgin olive oil with refined olive oil: Prospects and limitations. *Food Research International*, *54*(2), 2039–2044.

Garrido-Delgado, R., Arce, L., & Valcárcel, M. (2012). Multi-capillary column-ion mobility spectrometry: a potential screening system to differentiate virgin olive oils. *Analytical and Bioanalytical Chemistry*, *402*(1), 489–498.

Garrido-Delgado, R., Dobao-Prieto, M., Arce, L., & Valcárcel, M. (2015). Determination of volatile compounds by GC–IMS to assign the quality of virgin olive oil. *Food Chemistry*, *187*, 572–579.

Garrido-Delgado, R., Mercader-Trejo, F., Sielemann, S., de Bruyn, W., Arce, L., & Valcárcel, M. (2011). Direct classification of olive oils by using two types of ion mobility spectrometers. *Analytica Chimica Acta*, *696*(1-2), 108-115.

Gerhardt, N., Birkenmeier, M., Sanders, D., Rohn, S., & Weller, P. (2017). Resolution-optimized headspace gas chromatography-ion mobility spectrometry (HS-GC-IMS) for non-targeted olive. *Analytical and Bioanalytical Chemistry*, *409*(16), 3933–3942.

Gutierrez, F., & Fernandez, J. (2002). Determinant parameters and components in the storage of Virgin Olive Oil. Prediction of storage time beyond which the oil is no longer of "Extra" quality. *Journal of Agricultural and Food Chemistry*, *50*(3), 571-577.

International Organization for Standardization. (2005). *ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories*.

Majchrzak, T., Wojnowski, W., Dymerski, T., Gębicki, J., & Namieśnik, J. (2018). Electronic noses in classification and quality control of edible oils: A review. *Food Chemistry*, *246*, 192-201.

Melucci, D., Bendini, A., Tesini, F., Barbieri, S., Zappi, A., Vichi, S., Conte, L., Toschi, T. (2016). Rapid direct analysis to discriminate geographic origin of extra virgin olive oils by flash gas chromatography electronic nose and chemometrics. *Food Chemisty*, *204*, 263–273.

Riedl, J., Esslinger, S., & Fauhl-Hass, C. (2015). Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Analytica Chimica Acta*, *885*, 17-32.

Rottiers, H., Tzompa Sosa, D., Van de Vyver, L., Hinneh, M., Everaert, H., De Wever, J., Messens, K., Dewettinck, K. (2019). Discrimination of Cocoa Liquors Based on Their Odor Fingerprint: a Fast GC Electronic Nose Suitability Study. *Food Analytical Methods*, *12*(2), 475-488.

Vaisali, C., Charanyaa, S., Belur, P., & Regupathi, I. (2014). Refining of edible oils: a critical appraisal of current and potential technologies. *International Journal of Food Science and Technology*, *50*, 13-23.

Vautz, W., Zimmermann, D., Hartmann, M., Baumbach, J., Nolte, J., & Jung, J. (2011). Ion mobility spectrometry for food quality and safety. *Food Additives & Contaminants*, *23*(11), 1064-1073.

Ye, J., Wang, W., Ho, C., Li, J., Guo, X., Zhao, M., Jang, Y.,Tu, P. (2016). Differentiation of two types of pu-erh teas by electronic nose and ultrasound-assisted extraction-dispersive liquid-Liquid microextraction-gas chromatography-mass spectrometry. *Analytical Methods*, *8*(3), 593-604.

Zhang, L., Shuai, Q., Li, P., Zhang, Q., Ma, F., Zhang, W., & Ding, X. (2016). Ion mobility spectrometry fingerprints: A rapid detection technology for adulteration of sesame oil. *Food Chemistry*, *192*, 60-66.

A novel multi-platform High Resolution Mass Spectrometry non-targeted approach facing extra virgin olive oil adulteration with soft refined oils

Daniele Cavanna[1,2], Kamila Hurkova[3], Zbyněk Džuman [3], Andrea Serani[4], Matteo Serani[4], Chiara Dall'Asta[2], Monika Tomaniova[3], Jana Hajslova[3], Michele Suman[1]

[1]Advanced Laboratory Research, Barilla G. e R. Fratelli S.p.A., Via Mantova 166-43122 Parma (Italy)
[2]Department of Food and Drug, University of Parma, Parco Area delle Scienze 95/A-43124 Parma,(Italy)
[3]University of Chemistry and Technology, Prague, Department of Food Analysis and Nutrition, Technicka 3, 166 28 Prague 6 (Czech Republic)
[4]COTECA Srl Consulenze Tecniche agroindustriali - 56121 Pisa (Italy)

Submitted to "Journal of Agricultural and Food Chemistry"

Introduction

Extra Virgin Olive Oil (EVOO) is one of the most important and expensive edible oils, and therefore it is also one of the most adulterated food commodity over the global market.

The international Olive Council has clearly defined the different categories of olive oils (IOC, 2016) and, together with the European Union, the limits of the specific chemical parameters able to protect EVOO against potential adulterations with other edible oils (EU, 2016).

Typical frauds like the addition of other type of oils (i.e. seed oils) or mixtures with other refined oils can be easily detected with standards methods. For this reason, the fraudster are now focused on developing more sophisticated adulterations that would allow the creation of mixtures that cannot be discovered with regular methods (Garcia-Gonzalez, Aparicio, & Aparicio-Ruiz, 2018).

Although an official recognized method does not exist up to now, literature suggests different approaches, like Gas Chromatographic (Aparicio & Alonso, 1994), Nuclear Magnetic Resonance (Alonso-Salces, et al., 2010) or Isotopic Fingerprint studies (Faberi, et al., 2014) (Camin, et al., 2010) (Bontempo, et al., 2019) and the results presented are encouraging.

The use of soft deodorized (<100°C) or soft deacidified virgin or lampante olive oils can be considered one of the most critical point within this scenario: the products obtained at the end of the soft-refiniment processes cannot be considered "virgin" according to the current legislations, so any mixture of EVOO with soft refined oils is considered a fraud (Garcia-Gonzalez, Aparicio, & Aparicio-Ruiz, 2018).

As clearly described by Aparicio-Ruiz *et al*, this adulteration process is able to remove unpleasant volatile components or reduce the acidity of the oils without any macroscopic changes in the other chemical parameters (Aparicio-Ruiz, Romero, García-González, Oliver-Pozo, & Aparicio, 2017). The lack of changes in the chemical bulk of the EVOO during soft refinement is actually the reason that makes particularly challenging the detection of soft-refined oil addition with the official analyses established by the European Union.

The content of Fatty acid alkyl esters (Pérez-Camino, Cert, Romero-Segura, Cert-Trujillo, & Moreda, 2008), the lignans profile (Cecchi, et al., 2017), the amount of Pyropheopytin ''A'' (Serani & Piacenti, Sistema analytico per l'dentificazione di oli deodorati in oli vergini di oliva – Nota 1 – Analisi dei pigmenti clorofilliani in oli vergini di oliva, 2001) and the kinetic of diacyl glycerols isomerization (Serani, Piacenti, & Staiano, Sistema analytico per l'dentificazione di oli deodorati in oli vergini di oliva – Nota 2 – Cinetica di isomerizzazione dei digliceridi in oli vergini di oliva, 2001) were proposed in the past as possible tools to detect this fraud.

In addition, also the study of the trans- and cis-phytol isomer distribution (Vetter, Schröder, & Lehnert, 2012) is a possibility presented in literature, together with the evaluation of the global volatile profile (Cerretani, Bendini, Barbieri, & Lercker, 2008), but a robust analytical solution is not available up to now, especially when the aim is detect mixtures of soft refined oils with pure EVOO (Aparicio-Ruiz, Romero, García-González, Oliver-Pozo, & Aparicio, 2017).

Within this scenario, non-targeted analyses by LC-HRMS could represent an interesting approach for the identification of specific compounds responsible for the fraudulent process (Rubert, Zachariasova, & Hajslova, 2015).

Untargeted LC-HRMS methods were recently applied for the detection of defective olive oils (Kalogiouri, Alygizakis, Aalizadeh, & Thomaidis, 2016) or for the assessment of the geographical origin of the product (Gil-Solsona, et al., 2016) but, to the best of our knowledge, this is the first proposal of application for the detection of soft refined oils additions to extra virgin olive oils.

Due to the intrinsic nature of these non-targeted approaches, one critical point is their reliability when applied across different laboratories (Cavanna, Righetti, Elliott, & Suman, 2018); although ring-tests are highly recommended, they have not been largely applied yet for food protection purposes. Only some attempts of "metabo-ring tests" were reported recently in literature in other fields (Martin, et al., 2015) (Cajka, Smilowitz, & Fiehn, 2017).

On the basis of what described above, this work presents a non-targeted Liquid Chromatography – High Resolution Mass Spectrometry inter-laboratory study between two laboratories (equipped with two different types of mass spectrometers, a Q-Orbitrap and a Q-TOF) for the detection of new

chemical markers able to identify the addition of soft deodorized and soft deacidified low quality virgin or lampante olive oils to EVOO.

"In-house" soft refined oils were created and analyzed together with a group of pure Extra Virgin Olive Oils; in addition, different mixtures of pure EVOO and adulterated oils were included in the samples set.

After a robust data elaboration, the markers selected as discriminant in the two laboratories were compared with the aim to assess the reproducibility of the inter-laboratory study, and a tentative identification of compounds was performed.

Materials and Methods

Chemicals and samples description

Methanol, Ammonium Formate, Sodium Hydroxyde and Sodium Sulphate were purchased from VWR International, Ltd (Poole, United Kingdom).

Isopropyl Alcohol and Formic Acid were purchased from Sigma Aldrich (St. Louis, MO).

Water was purified using a Milli-Q system (Millipore, Bedford, MA).

43 pure Extra Virgin Olive Oil samples were collected trying to include the highest variability in their characteristic (i.e. coming from different suppliers, related to different production years, with different storage conditions etc.). 40 of them were produced with Italian olives and 3 with olives coming from European Union.

Moreover, 3 "soft deodorized", 2 "soft deacidifed" and 2 "soft deacidified and then deodorized" samples were prepared in a laboratory scale starting from virgin and lampante olive oils.

These adulterated samples were analyzed for the official EVOO parameters (EU, 2016); subsequently, according to the results obtained, six mixtures were prepared at different percentages of adulteration (with the aim to create frauded samples that could be in compliance with the current legislation) and were analyzed together with two suspect commercial blends with deodorized oils obtained from the bulk market.

Samples list is detailed in the Supplementary material (Section 1 – table S1).

Soft Deodorization

The process was executed partially according to a previous work presented in literature (Aparicio-Ruiz, Romero, García-González, Oliver-Pozo, & Aparicio, 2017).

400 mL of a defective oil were introduced into a round-bottom flask with triple-neck containing a magnetic stir bar.

A thermometer was inserted into the right neck, while a dropping funnel containing Milli-Q grade water was connected to the middle one. Finally, the left neck was used to connect the vacuum tubing to the system.

This apparatus was placed on a magnetic stirring hotplate (VELP Scientifica, Italy) operating at 300 rpm.

When the oil reached the temperature of 60°C, the first 0.2 mL of water were added to the oil in order to start the distillation. When the oil reached the temperature of 100 °C, it was kept for 60 minutes at that temperature with other additions of 0.2 mL of water when the previous ones evaporated.

After deodorization, the sample was cooled down to room temperature and stored at 4 °C before the analyses.

Three different samples were prepared following this procedure and were labelled as "DEO", "DEO2" and "DEO3".

Soft Deacidification

This process was executed on two defective olive oils with an acidity of 1.4% and 1.0% respectively. 600 mL of defective oil were weighed and transferred into a beaker containing a magnetic stir bar and put on a magnetic stirrer working at room temperature.

A specific amount (in mL) of a 50% (w\w) Sodium Hydroxide solution was added, followed by a double amount (in mL) of a 10% (w\w) Sodium Sulphate solution, according to the following formula:

$$NaOH\ (mL) = [(A-B)*C*D/(E/100)]* [(100/F)/G]$$

Where A is the starting acidity (Free Fatty Acids - FFA) of the oil (% Oleic acid); B is the target % FFA of the oil (0.4% in our work); C is the oil weight in grams; D is the Sodium Hydroxide molecular weight (40 g/mol); E is the Oleic Acid molecular weight (282 g/mol); F is the concentration of the Sodium Hydroxide solution and G is the density of the 50% (w\w) Sodium Hydroxide solution (1.52 g/mL).

After 5 minutes of stirring, the oil was centrifuged at 5000 rpm and 200 mL of the supernatant (the neutralized oil) was stored at 4 °C before the analyses.

The remaining 400 mL of the neutralized oil were subjected to the soft deodorization process described above.

At the end we created 2 “soft deacidified” samples (named “DEA” and “DEA2”) and 2 “soft deacidified and then deodorized” samples (named “DEO_DEA” and “DEO_DEA2”).

Sample extraction

The extraction procedure was based on a previous work of the University of Chemistry and Technology in Prague (Vaclavik, Cajka, Hrbek, & Hajslova, 2009). Both the laboratories applied the same extraction procedure: 1 mL of sample was transferred into a PTFE cuvette (15 mL) and 3 mL of a mixture methanol:water (80:20, v/v) were added for isolation of the polar fraction. Sample was then shaken for 20 min at 240 RPM using a horizontal laboratory shaker (IKA Labortechnik, Germany) and then centrifuged at 10 000 RPM for 5 min (Hettich, Germany).

1 mL of the hydro–alcoholic layer was transferred into an amber glass vial and stored at -20 °C until analysis with LC-HRMS instrument.

For the evaluation of method reproducibility, 20 % of the samples were double prepared.

The same procedure was performed also into empty tubes in which all the steps were executed without the oil addition. These samples were labelled as “extraction blank”.

Additionally, a quality control (QC) sample was prepared mixing 10 μL of each extract sample.

LC-HRMS analysis

In *Laboratory #1* HPLC analysis was performed with a Dionex UltiMate 3000 UHPLC (Thermo Fisher Scientific, Inc., Waltham, MA) equipped with a BEH C18 150 x 2.1 mm, 1.7 μm particle size analytical column (Waters, Milford, MA, USA) maintained at 45 °C. Gradient elution was performed using Formic Acid (FA) and Ammonium Formate (AF) as mobile phase modifiers with a constant flow rate of 0.3 mL/min.

Mobile phase A was constituted by a 5mM AF and 0.1% FA solution of a mixture of water and methanol (95:5 v\v), while mobile phase B was constituted by a 5mM AF and 0.1% FA solution of a mixture of Isopropyl Alcohol, Methanol and Water (65:30:5, v/v/v).

Gradient conditions are the following: after 1 minute with 75 % of mobile phase A and 25 % of mobile phase B, the percentage of solvent B increased to 80 % in 2.5 minutes, then to 100 % in 4 minutes and then was maintained at this percentage for 3.5 minutes before column re-equilibration (5 minutes).

Autosampler was maintained at 5 °C and the injection volume was 4 μL.

Mass spectrometry detection was performed with a benchtop Q Exactive™ Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific, Waltham, MA) equipped with a heated (H)

electrospray ionization (ESI) interface (Thermo Fisher Scientific,Waltam, MA). Two analytical sequences (one with positive and one with negative ionization mode) were executed with a "Full Scan-data dependent fragmentation" experiment.

The resolution of the "full scan" experiment was 70000 full width at half maximum (FWHM) (*m/z* 200) and was 17500 FWHM (*m/z* 200) for the MS/MS experiment.

All the other ionization source and mass spectrometer conditions are the same as described in a previous work of Barilla research group (Cavanna, Catellani, Dall'Asta, & Suman, 2018).

Samples were randomly injected, while the QC sample was injected at the beginning and at the end of the sequence and every 10 sample injections.

In *Laboratory #2* HPLC analysis was performed with a Dionex UltiMate 3000 UHPLC (Thermo Fisher Scientific, Inc., Waltham, MA) equipped with a BEH C18 100 x 2.1 mm, 1.7 μm particle size analytical column (Waters,Milford, MA, USA) maintained at 45°C.

HPLC and chromatographic conditions were the same used in *Laboratory #1* with the exception of the gradient steps, that are the following: after 0.5 minute with 75 % of mobile phase A and 25 % of mobile phase B, the percentage of solvent B increased to 80 % in 1.5 minutes, then to 100 % in 3 minutes and then was maintained at this percentage for 2.5 minutes before column re-equilibration (2.5 minutes).

Mass spectrometry detection was performed with TripleTOF® 6600 quadrupole time-of-flight (TOF) mass spectrometer (SCIEX) equipped with a DuoSpray™ source with a separated ESI and APCI ion sources. ESI was used for the sample measurement and APCI was used for exact mass calibration of the TripleTOF instrument.

Two analytical sequences (one with positive and one with negative ionization mode) were executed with a "TOF-MS" and "information-dependent acquisition" (IDA) experiments in order to simultaneously collect the "Full Scan" spectrum and the "Product Ion" spectra of the most abundant ions.

The achieved resolving power was >40000 FWHM (*m/z* 829.5393 for positive ionization mode and *m/z* 933.637 for negative ionization mode), while >25000 FWHM (*m/z* 811.5288 for positive ionization mode and *m/z* 933.637 for negative ionization mode) for the MS/MS experiments.

All the other ionization source and mass spectrometer conditions applied to the TripleTOF® 6600 are the same as described in a previous work of the University of Chemistry and Technology group (Hurkova, Rubert, Stranska-Zachariasova, & Hajslova, 2017).

Samples were randomly injected, while the QC sample was injected at the beginning and at the end of the sequence and every 10 sample injections.

Data treatments and statistics

Laboratory #1 UHPLC-HRMS raw data were acquired using Xcalibur software (version 3.0 Thermo Fisher Scientific,Waltam, MA); peaks alignment, "extraction blanks" subtraction and features extraction were performed using Compound Discoverer software (version 2.1 Thermo Fisher Scientific,Waltam, MA) directly connected to "Chemspider" and "m/z CLOUD" databases and able to perform "in silico" fragmentations; the mass range inspected was between 100 *m/z* and 1000 *m/z* from 0.5 to 11 minutes of the chromatographic runs.

The values of the critical parameters for features extractions and identification are the same as described in a previous work (Cavanna, Catellani, Dall'Asta, & Suman, 2018).

Laboratory #2 UHPLC-HRMS raw data were acquired using Analyst 1.7.1 TF (SCIEX), and the qualitative analysis was performed using PeakView 2.2 (SCIEX) equipped with MasterView and Formula Finder and directly linked to the "ChemSpider" database.

MarkerView software (version 1.2.1, SCIEX) was used for data processing; the mass range inspected was between 100 *m/z* and 1000 *m/z* from 0.6 to 8 minutes of the chromatographic runs.

Peaks alignment was carried out with a tolerance of 0.2 min for the retention time and of 0.01 Da for the *m\z* values, while MS\MS spectra were studied using METLIN (https://metlin.scripps.edu/landing_page.php?pgcontent=mainPage) online database.

In both the laboratories, the resulting two data matrixes (for positive and negative ionization modes), containing the area values for all the features, were exported and processed with SIMCA software (version 14.1 Umetrics, Umea, Sweden) for chemometric data elaboration.

Data were log transformed and Pareto scaled, then a preliminary Principal Component Analysis (PCA) was executed in order to check the clusterization of the samples and the QCs positioning in the scores plot.

Subsequently, QCs were removed and the replicate samples were averaged; then explorative Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) models were created comparing the group of the pure "EVOO" samples against the group "OTHER" containing the adulterated samples.

In this work different kind of adulterations are studied ("DEO", "DEA" and "DEO_DEA") and the amount of samples is different in each group; for these reasons, multivariate models were used only as preliminary approaches for a global evaluation but markers selection was executed with univariate models.

For each data matrix, different "One way ANOVA" tests were applied using "Microsoft Excel 2016 professional plus" software (Microsoft Corporation) comparing the "EVOO" group with the pure "DEA", the pure "DEO" and the pure "DEO_DEA" samples.

The molecules with a P-Value lower than 0.05 were selected and a tentative identification of compounds was executed following the workflow described in a previous paper (Cavanna, Catellani, Dall'Asta, & Suman, 2018) and referring to the Standard Initiative on Metabolomics (Sumner, et al., 2007) and subsequent improvements (Schymanski, et al., 2014) for identification ranking.

The presence of the most interesting features was evaluated in the mixtures samples in order to check their value as discriminant markers in real fraudulent situations.

Results and Discussion

The results presented in this work are related to the polar fraction of the oils, despite usually the attention of the researchers is mostly focused on the lipophilic molecules, which are the main constituents of this food commodity.

In a preliminary phase of this research, also the non polar fraction of the oils was extracted and studied but the results obtained were not encouraging and for this reason the attention was totally focused on the polar component, that, on the contrary, provided interesting results also at the early stage of this study.

As example, the Total Ion Chromatograms of a pure EVOO sample and of different mixtures are presented in the Supplementary Material (Section 2, figure S1).

Multivariate studies

After peaks alignment, blank subtraction and isotopes merging, in *Laboratory #1* 3024 features were extracted in positive mode and 1019 in negative, while in *Laboratory #2* the software selected 5612 and 2800 features in positive and negative modes respectively.

The preliminary PCA models highlighted for both the laboratories a tight clustering of the QCs and a substantial overlay of the replicate samples (data not shown) certifying the goodness of the analytical procedure performed in both the laboratories. Therefore, the PCA models were obtained without the inclusion of QC samples and using the average values of the replicate extracts. In this case, a partial separation of the EVOO samples from the non-authentic ones was achieved.

The ESI+ scores plot is reported in Figure 1, while ESI- scores plot can be found in the Supplementary Material (Section 2, Figure S2).

The clustering of the two groups dramatically increase when we move to supervised PLS-DA models, as presented in Figure 2 for the ESI+ mode and in Figure S3 (Supplementary Material) for the ESI- mode. The variance of the x and y variables explained by the model (R^2X (cum) and R^2Y (cum)) and the cumulative predicted variation in the Y matrix (Q^2 (cum)) are reported in the plots as well.
This global differentiation of the groups encouraged the subsequent univariate study, aimed at the selection and identification of the molecules responsible for the fraudulent processes.

Compounds selection and identification

Different "One-way ANOVA" tests were applied on the data obtained in *Laboratory #1*, comparing the "EVOO" group with the "DEA", the "DEO" and the "DEO_DEA" groups: the molecules with a P-Value lower than 0.05 were selected for subsequent studies.
The compounds selected in *Laboratory #1* were searched in the data set from *Laboratory#2*, evaluating that also in these studies their P-Value was lower than 0.05.
Subsequently, a tentative identification of compounds was executed for the features that were visible also in the mixtures, taking into account that only these could represent a realistic tool for frauds detection.
The summary of the results obtained is detailed in Tables 1 and 2; when more than one molecule is potentially related to the same compound, this means that that features are in-source fragments of the precursor ion.
Table 3 presents, as example, the "Compound 1" supposed name, together with the identification level (Schymanski, et al., 2014) and with its mean area values in the different groups.
The complete list of the supposed names, identification levels and mean area values in the groups for all the selected compounds is presented in the Supplementary Material (Section 3 – Table S2).
All the selected molecules were detected at significant levels in the eight mixtures (the six ones created in the laboratory and the two suspect commercial blends). As example, the "Variables trends plot" chart of "Compound 12" is given in Figure 3. Other examples of "Variables trends plot" charts can be found in the Supplementary Material (Section 3 – Figure S4).

Markers Interpretation

All the selected compounds were also detected in the Virgin and Lampante Olive Oils used for the refinement processes, but they occur at lower amount in EVOO. Accordingly, it is likely that these molecules are characteristic of oils obtained with low quality olives.
Moreover, these compounds are not affected by the adulteration procedures. This fact turns them into a set of robust markers, because they are not originated during refinement, and thus subjected to

possible fluctuation according to the process applied, but they clearly indicate the use of low quality raw material.

This result is consistent with other markers presented in the past, like the Fatty acid alkyl esters, that were not affected by the soft deodorization procedure (Pérez-Camino, Cert, Romero-Segura, Cert-Trujillo, & Moreda, 2008).

A chemical interpretation of these features is not always easy to perform; for example, it seems that compounds 4 and 6 are produced by the degradation of the oxidized fatty acids.

The reason why these compounds are present mostly in the adulterated samples is not totally clear but a hypothesis was developed: virgin but especially lampante olive oils have a high amount of oxidized fatty acids, and this should justify the presence of more products of this chemical process.

Moreover, oxidation generates peroxides and hydroperoxides that subsequently bring to the creation of conjugated dienes with a loss of water; for this reason, for example, the Oleic Acid is transformed into the conjugated Linolenic Acid (Frankel, Review.Recent advances in lipid oxidations, 1991) (Frankel, Secondary products of lipid oxidation, 1987). This process is catalysed by thermal and physical stresses that happen in virgin and lampante olive oils more than in extra virgin olive oils.

Our hypothesis is that, in a minor percentage, the oxidized fatty acids could also degrade to unsaturated hydroxyl acids, instead of moving to the formation of a conjugated diene with a loss of water. This justify, for example, the presence of the Propyl-12-hydroxy-9-octadecenoate (named "Compound 6" in the tables) as one of the markers.

Compounds 2, 3 and 5 are part of the metabolic pathway of polyphenols formation and are probably connected to a lower quality of the olive used for the creation of the oil.

The identification process of these molecules will be undoubtedly reinforced in the future in order to increase the identification level of some of these markers.

In any case, these features are present exclusively (i.e. "Compound 10" and "Compound 11") or mostly (i.e. "Compound 1" and "Compound 2") in the adulterated samples and were identified also in the mixtures, allowing the detection of fraudulent samples that would be wrongly considered "legal" according to the official legislation.

Moreover, the fact that these features were detected also in the commercial mixtures (not prepared in the laboratory) reinforce the results obtained, avoiding the risk that these data are affected by some mistakes performed during the "in-house" adulteration processes.

If compared with the other works presented in literature, the results obtained seems encouraging, taking into account that the compounds presented in this study were detected also in different mixtures.

All the works presented in the past, indeed, studied pure samples (Cecchi, et al., 2017) (Serani & Piacenti, Sistema analytico per l'dentificazione di oli deodorati in oli vergini di oliva – Nota 1 – Analisi dei pigmenti clorofilliani in oli vergini di oliva, 2001) (Serani, Piacenti, & Staiano, Sistema analytico per l'dentificazione di oli deodorati in oli vergini di oliva – Nota 2 – Cinetica di isomerizzazione dei digliceridi in oli vergini di oliva, 2001) (Vetter, Schröder, & Lehnert, 2012), using also sometimes deodorization conditions stronger than the ones used in this work (Pérez-Camino, Cert, Romero-Segura, Cert-Trujillo, & Moreda, 2008).

Furthermore, the maximum concentration of fatty acid ethyl esters is now an official parameter (IOC, 2016), so the markers presented by Pérez-Camino *et al* (Pérez-Camino, Cert, Romero-Segura, Cert-Trujillo, & Moreda, 2008) should be considered a quality criterion rather than authentication markers.

Evaluation of the inter-laboratory study

The whole analytical process (from sample preparation to markers interpretation) was performed in two different laboratories with two different HRMS platforms, in order to evaluate the inter-laboratory applicability of a non-targeted mass spectrometry study.

The amount of features detected in *Laboratory#2* is higher than the ones obtained in *Laboratory#1*. This is partially explained with the fact that MarkerView software treated as separated features all the ionization adducts (i.e. ammonium, sodium etc.) of the same compound; on the contrary, Compound Discoverer software merged as unique feature all the adducts. This strongly contributed to the discrepancy in the total amount of compounds between the two laboratories.

The multivariate approaches (for both unsupervised PCA and supervised PLS-DA models) showed comparable results, showing that in both the laboratories the fingerprints detected were fit for the purpose of this work.

Moving to the univariate analysis, seven compounds were selected as discriminative in both the laboratories, with a similar trend throughout the samples.

"Compound 1" was detected in *Laboratory #2* only as the in-source fragmentations feature and not "as is". This is probably due to the different ESI ionization sources present in the mass spectrometers. Except for this discrepancy, for each of the common *m/z* values the results obtained are comparable not only for the information provided with the "Full scan" experiment, but also for the MS/MS spectra (see Supplementary Material – Section 4 - Figure S5 and S6 as examples).

These results put in evidence a similar behavior of the molecules in the two instrumentation, despite the two different approaches that are applied in the mass spectrometers for compounds fragmentation and for accurate mass measurement.

The results obtained in the two laboratories are similar despite the use of two different mass spectrometers (a Q-Orbitrap and a Q-TOF): it must be underlined that the good agreement was obtained with freshly prepared samples, and not injecting the same extracts in different platforms.

Conclusions

In summary, the results obtained within the inter-laboratory ring test and the two analytical processes are on the whole comparable, demonstrating the goodness of the untargeted approach for the detection of chemical markers introduced by the addition of soft refined oils (soft deodorized and soft deacidified) to extra virgin olive oils.

Seven compounds related to soft refined oils additions were selected as markers and tentatively identified. These markers are a valuable tool for unmasking this specific fraud, considering that they were detected in raw materials and in the final products, remaining unaffected by soft-refining. In addition, their robustness was proven independently over two laboratories and confirmed by the application to samples from the market.

In this work, the minimum adulteration percentage of the mixtures was 40%: although it may seem an high value, we know that it is a real condition because, as also reported in literature, all the methods presented in the past, including the official methods, were not able to detect additions of 50% of soft refined oils (Aparicio-Ruiz, Romero, García-González, Oliver-Pozo, & Aparicio, 2017). Moreover, the markers selected in our study were detected in the two commercial mixtures: we don't know the percentage of additions in these two samples but we know that the market considered these samples as extra virgin olive oils.

Further efforts will be performed for a better clarification of the chemical structures. This will lead to the development of a target method for a quantitative evaluation.

Acknowledgement

Authors would like to thank Raffaele Maranzoni and Evelyne Sasso from "Coppini Arte Olearia" for the invaluable support in EVOO authentic samples collection.

This work was supported by the European Union, Seventh Framework Program (FP7/2007- 2013), under grant agreement 613688 (Ensuring the Integrity of the European food chain FoodIntegrity).

This work has received funding from the Czech National Agency for Agriculture Research (NAZV-QJ1530272) the Operational Programme Prague Competitiveness (CZ.2.16/3.1.00/21537 and CZ.2.16/3.1.00/24503) and by the "National Program of Sustainability I" NPU I LO1601. This research has also received funding from the European Union Horizon 2020 research and innovation program under grant agreement No 692195 (MultiCoop).

Tables

	LABORATORY #1							LABORATORY #2						
Compound ID	**Formula**	**m/z $[M+H]^+$**	**Δ mass (ppm)**	**RT [min]**	**P-value DEA**	**P-value DEO**	**P-value DEO_DEA**	**Formula**	**m/z $[M+H]^+$**	**Δ mass (ppm)**	**RT [min]**	**P-value DEA**	**P-value DEO**	**P-value DEO_DEA**
1	$C_{21}H_{42}O_3$	343.3207	1.5	6.06	2.68E-04	2.00E-06	1.32E-07	N.D.	/	/	/	/	/	/
	$C_{18}H_{36}O_3$	301.2737	1.7	6.06	6.29E-05	1.05E-07	6.41E-09	$C_{18}H_{36}O_3$	301.2707	11.9	3.99	1.30E-49	1.87E-23	5.48E-53
	$C_{18}H_{34}O_2$	283.2632	1.8	6.06	2.68E-04	2.00E-06	1.32E-07	$C_{18}H_{34}O_2$	283.2625	4.2	3.99	2.07E-24	1.61E-22	2.26E-28
	$C_{18}H_{32}O$	265.2526	1.9	6.06	6.29E-05	1.05E-07	6.41E-09	$C_{18}H_{30}$	247.2407	7.7	3.99	3.24E-33	2.25E-24	6.72E-33
2	$C_{10}H_{12}O_2$	165.0909	4.2	3.65	6.57E-06	4.06E-05	1.24E-07	N.D.	/	/	/	/	/	/
3	$C_{11}H_{15}NO_3$	210.1128	0.9	2.10	2.16E-10	1.72E-08	1.00E-15	N.D.	/	/	/	/	/	/
4	$C_{10}H_{16}O_2$	169.1223	3.0	4.57	4.53E-07	3.28E-01	1.00E-15	N.D.	/	/	/	/	/	/
5	$C_{12}H_{16}O_5$	241.107	2.5	2.58	1.13E-02	2.28E-05	4.82E-05	N.D.	/	/	/	/	/	/
6	$C_{21}H_{40}O_3$	341.3049	2.1	6.19	1.36E-04	1.27E-01	1.16E-06	N.D.	/	/	/	/	/	/
	$C_{18}H_{34}O_3$	299.2579	2.3	6.19	1.89E-04	1.35E-02	1.28E-07	N.D.	/	/	/	/	/	/

Table 1: comparison of the adulterations markers detected in positive ionization mode. N.D. = Not Detected

	LABORATORY #1							LABORATORY #2						
Compound ID	**Formula**	**m/z [M-H]⁻**	**Δ mass (ppm)**	**RT [min]**	**P-value DEA**	**P-value DEO**	**P-value DEO_DEA**	**Formula**	**m/z [M-H]⁻**	**Δ mass (ppm)**	**RT [min]**	**P-value DEA**	**P-value DEO**	**P-value DEO_DEA**
7	$C_{23}H_{29}ClO_4$	403.1682	1.5	5.67	2.03E-10	1.00E-15	1.17E-08	$C_{23}H_{29}ClO_4$	403.1663	3.2	3.74	5.19E-37	1.52E-37	1.22E-20
8	$C_{23}H_{27}ClO_4$	401.1534	3.7	5.60	1.00E-15	1.00E-15	1.09E-05	$C_{23}H_{27}ClO_4$	401.151	2.2	3.67	7.33E-56	7.16E-43	n.a.
9	$C_{25}H_{33}ClO_6$	463.1903	3.5	5.35	9.10E-10	1.00E-15	1.07E-07	$C_{25}H_{33}ClO_6$	463.182	14.5	3.49	2.70E-24	2.26E-22	1.04E-19
10	$C_{25}H_{31}ClO_6$	461.1748	3.9	5.41	1.00E-15	1.00E-15	1.00E-15	$C_{25}H_{31}ClO_6$	461.1723	1.5	3.53	4.33E-21	2.67E-13	6.85E-24
11	$C_{23}H_{33}ClO_5$	423.1955	4.0	5.40	1.00E-15	1.00E-15	1.00E-15	$C_{23}H_{33}ClO_5$	423.1927	2.6	3.52	2.06E-42	2.75E-50	2.69E-24
12	$C_{23}H_{32}O_4$	371.2231	2.4	5.50	1.29E-05	3.34E-09	8.05E-06	$C_{23}H_{32}O_4$	371.222	0.5	3.58	3.59E-26	4.23E-28	1.21E-15

Table *2*: comparison of the adulterations markers detected in negative ionization mode

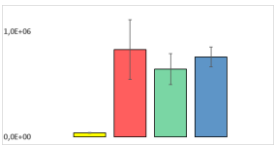
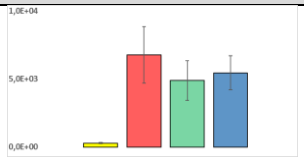
Compound ID	Name	ID Level	*Lab#1* Area values	*Lab#2* Area values
1	Propylene glycol - 1 Stearate	3		

Table *3* "Compound 1" description and comparison of its mean area values (+/− standard error) through the groups (yellow bar: EVOO; red bar: DEO; green bar: DEA, blue bar: DEO+DEA samples)

Figures

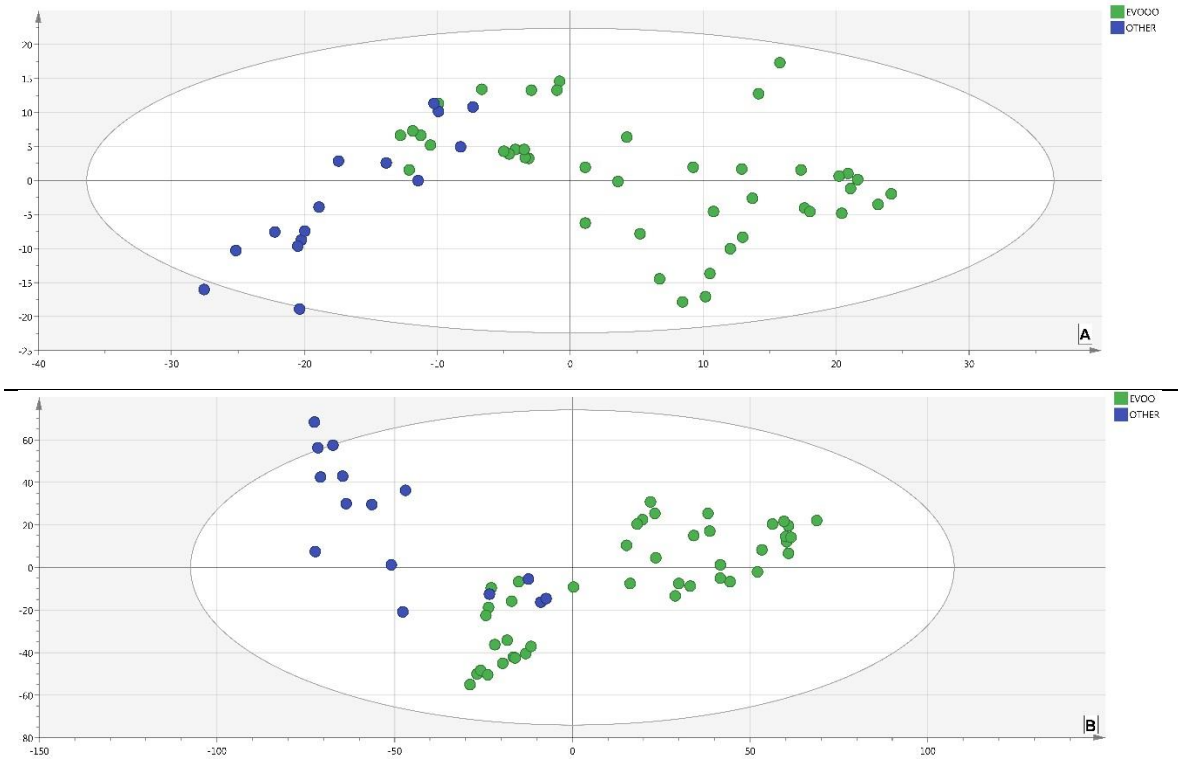


Figure 1. ESI + PCA Scores plots of the samples. A: *Laboratory #1*; number of PCs: 8; explained variance of the first two PCs: 46.2%. B: *Laboratory #2; n*umber of PCs: 7; explained variance of the first two PCs: 18.6%. Green dots: "EVOO" samples; blue dots: "OTHER" samples.

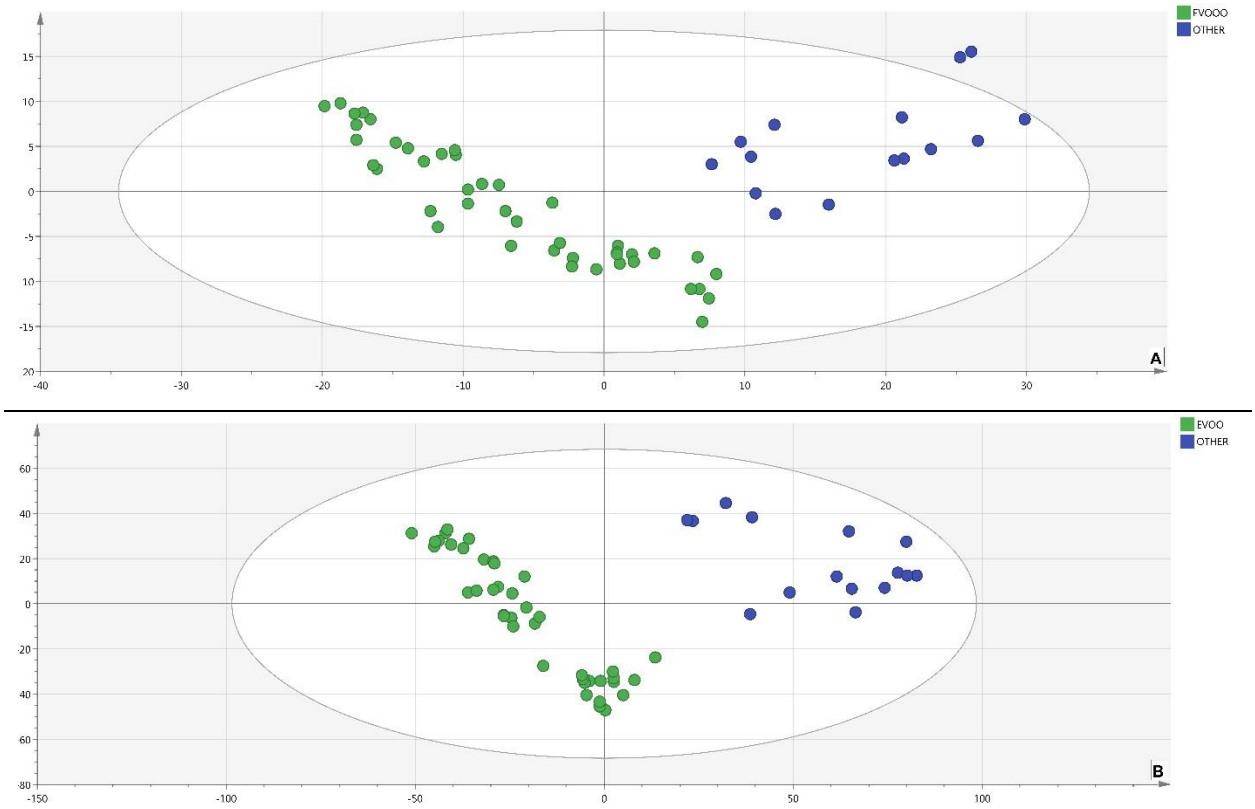


Figure 2. ESI + PLS-DA Scores plots of the samples. A: *Laboratory #1*; R^2X (cum) = 0.564; R^2Y (cum) = 0.964; Q^2 (cum) = 0.903. B: *Laboratory #2*; R^2X (cum) = 0.208; R^2Y (cum) = 0.991; Q^2 (cum) = 0.902. Green dots: "EVOO" samples; blue dots: "OTHER" samples.

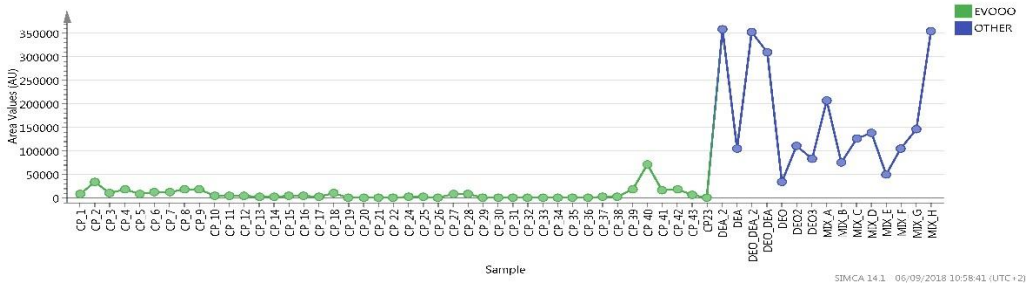
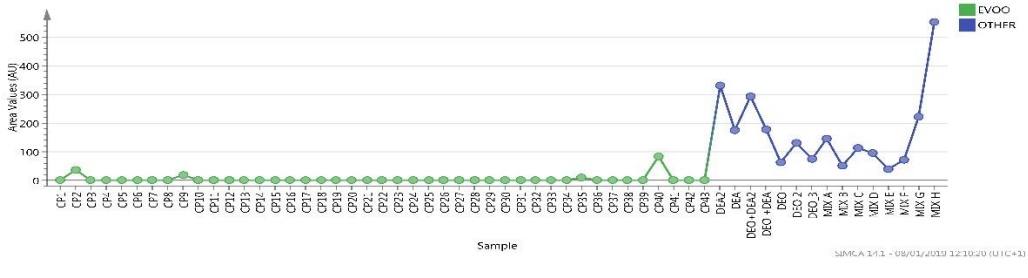
LABORATORY #1	
LABORATORY #2	

Figure 3. Variable trend plots of "Compound 12". Green dots: area values of the marker in "EVOO" group; blue dots: area values of the marker in "OTHER" group.

Supplementary Material

Section 1 – Sampling Plan

The global list of the samples used in this study is detailed in table S1.

NAME	TYPE	YEAR	SUPPLIER	STORAGE
CP_1	EVOO	2015/2016	1	2-8°C
CP_2	EVOO	2015/2016	1	2-8°C
CP_3	EVOO	2015/2016	2	2-8°C
CP_4	EVOO	2015/2016	2	2-8°C
CP_5	EVOO	2015/2016	2	2-8°C
CP_6	EVOO	2015/2016	2	2-8°C
CP_7	EVOO	2015/2016	2	2-8°C
CP_8	EVOO	2015/2016	2	2-8°C
CP_9	EVOO	2015/2016	1	2-8°C
CP_10	EVOO	2015/2016	3	2-8°C
CP_11	EVOO	2015/2016	3	2-8°C
CP_12	EVOO	2015/2016	3	2-8°C
CP_13	EVOO	2016/2017	2	2-8°C
CP_14	EVOO	2016/2017	2	2-8°C
CP_15	EVOO	2016/2017	2	2-8°C
CP_16	EVOO	2016/2017	2	2-8°C
CP_17	EVOO	2016/2017	2	2-8°C
CP_18	EVOO	2016/2017	2	2-8°C
CP_19	EVOO	2016/2017	2	2-8°C
CP_20	EVOO	2017/2018	2	2-8°C
CP_21	EVOO	2017/2018	2	2-8°C
CP_22	EVOO	2017/2018	2	2-8°C
CP_23	EVOO	2017/2018	2	2-8°C
CP_24	EVOO	2017/2018	2	2-8°C
CP_25	EVOO	2017/2018	2	2-8°C
CP_26	EVOO	2017/2018	2	2-8°C

NAME	TYPE	YEAR	SUPPLIER	STORAGE
CP_27	EVOO	2017/2018	2	2-8°C
CP_28	EVOO	2017/2018	2	2-8°C
CP_29	EVOO	2017/2018	2	2-8°C
CP_30	EVOO	2017/2018	2	2-8°C
CP_31	EVOO	2017/2018	2	2-8°C
CP_32	EVOO	2016/2017	2	2-8°C
CP_33	EVOO	2015/2016	2	2-8°C
CP_34	EVOO	2017/2018	2	2-8°C
CP_35	EVOO	2017/2018	2	2-8°C
CP_36	EVOO	2017/2018	2	2-8°C
CP_37	EVOO	2017/2018	2	2-8°C
CP_38	EVOO	2017/2018	2	2-8°C
CP_39	EVOO	2015/2016	1	Room Temperature
CP_40	EVOO	2015/2016	1	Room Temperature
CP_41	EVOO	2015/2016	1	Room Temperature
CP_42	EVOO	2015/2016	1	Room Temperature
CP_43	EVOO	2015/2016	1	Room Temperature
DEO	DEODORIZED	/	/	2-8°C
DEO2	DEODORIZED	/	/	2-8°C
DEO3	DEODORIZED	/	/	2-8°C
DEA	DEACIDIFIED	/	/	2-8°C
DEA2	DEACIDIFIED	/	/	2-8°C
DEO_DEA	DEODORIZED AND DEACIDIFIED	/	/	2-8°C
DEO_DEA2	DEODORIZED AND DEACIDIFIED	/	/	2-8°C

NAME	TYPE	YEAR	SUPPLIER	STORAGE
MIX A	EVOO 25% + DEA-DEO 75%	/	/	2-8°C
MIX B	EVOO 25% + DEO 75%	/	/	2-8°C
MIX C	EVOO 55% + DEA-DEO 45%	/	/	2-8°C
MIX D	EVOO 60%+ DEA-DEO 40%	/	/	2-8°C
MIX E	EVOO 50% + DEO 50%	/	/	2-8°C
MIX F	EVOO 40% + DEO 60%	/	/	2-8°C
MIX G	Commercial DEO sample	/	/	2-8°C
MIX H	Commercial DEO sample	/	/	2-8°C

Table S2: list and description of the samples. CP_19, CP_26 and CP_31 were produced with olives coming from European Union

Section 2 – Multivariate studies

The Total Ion Chromatograms of a pure EVOO sample, of mixture A and of mixture H are presented in the figure below.

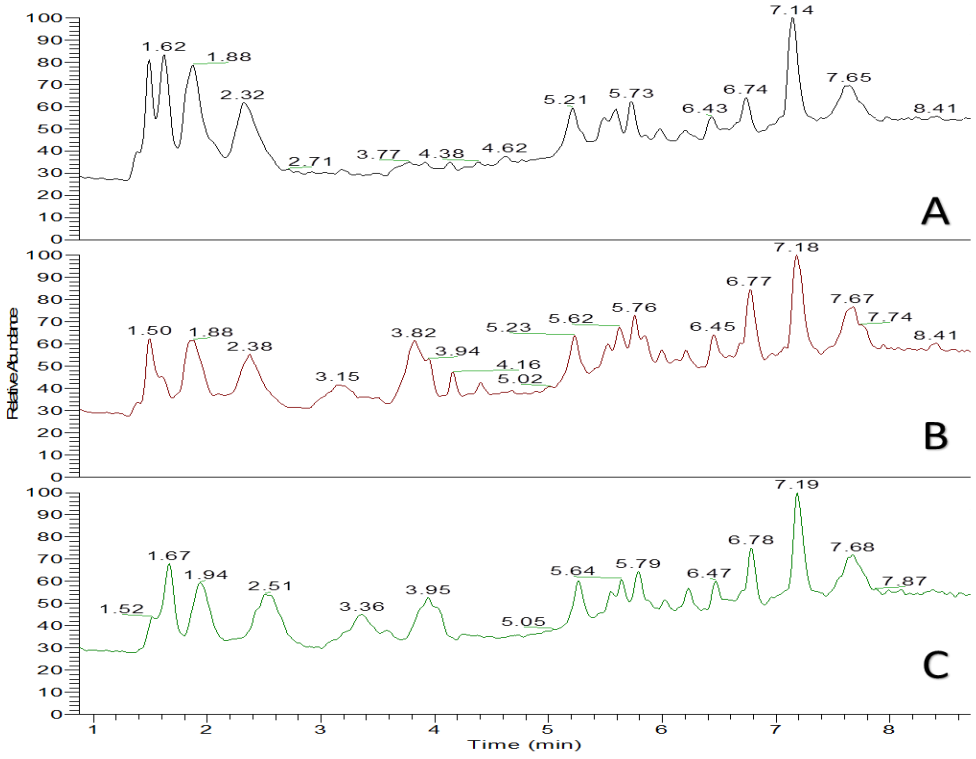


Figure S1. Total Ion Chromatograms (from 1 to 9 min) of (A) pure EVOO sample; (B) mixture A; (C) mixture H

The PCA and PLS-DA scores plot of the results obtained with the negative ionization mode are presented in figures S2 and S3.

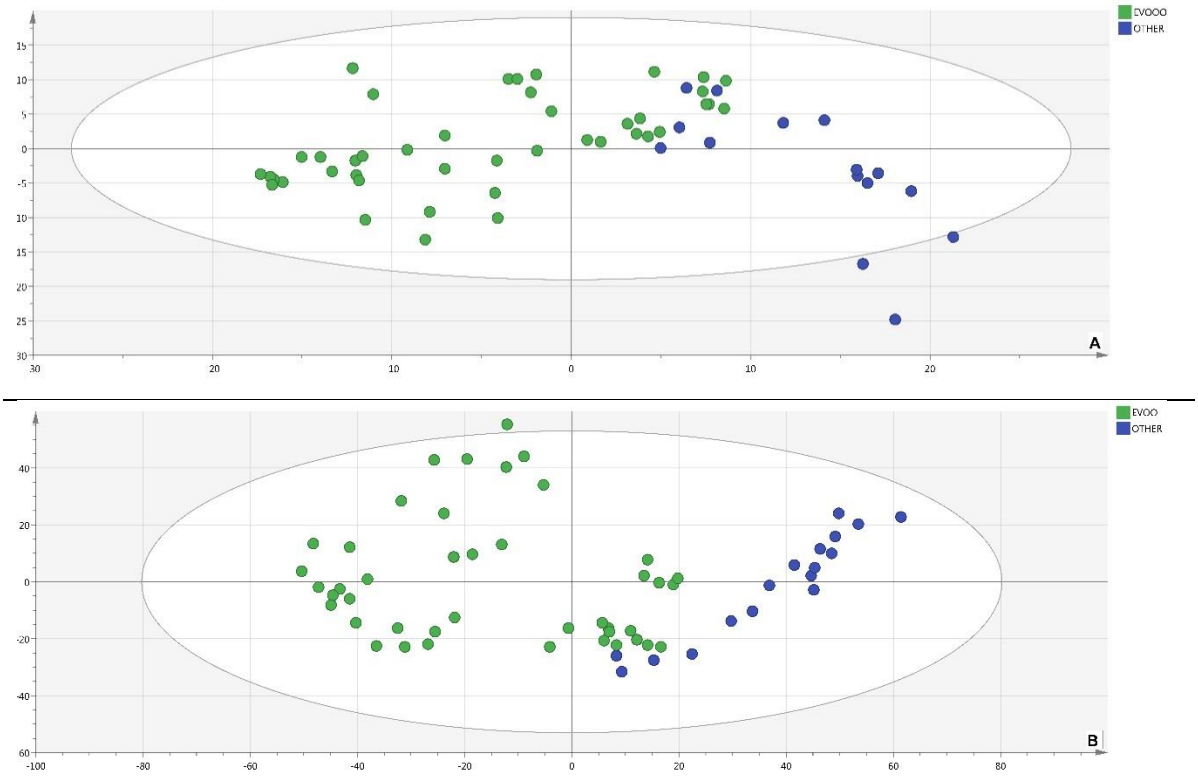


Figure S2. ESI - PCA Scores plots of the samples. A: *Laboratory #1*; number of PCs: 9; explained variance of the first two PCs: 49.4%. B: *Laboratory #2; n*umber of PCs: 7; explained variance of the first two PCs: 22.0%. Green dots: "EVOO" samples; blue dots: "OTHER" samples.

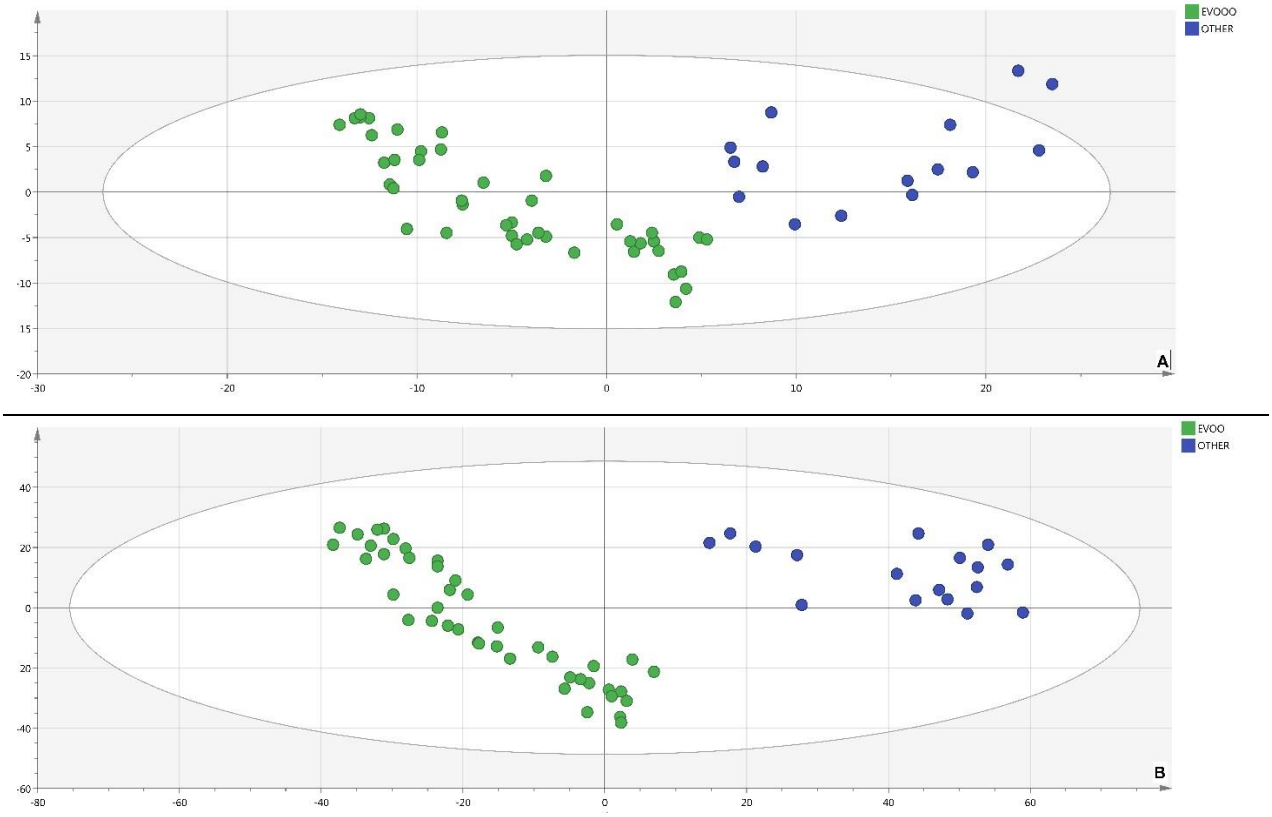
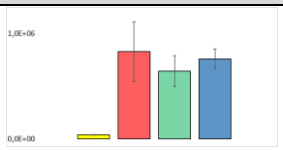
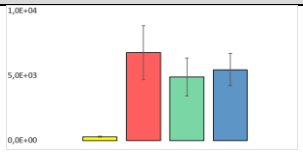
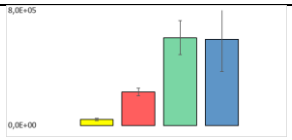
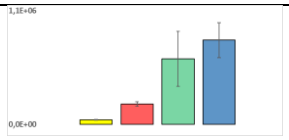
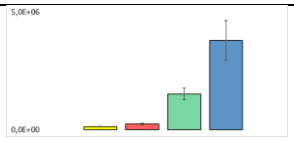
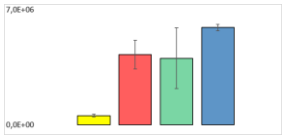
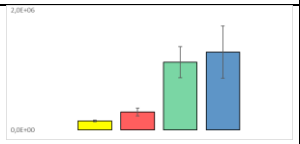
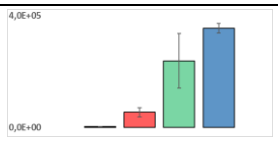
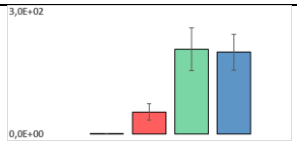
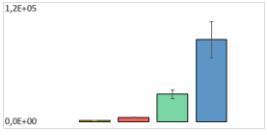
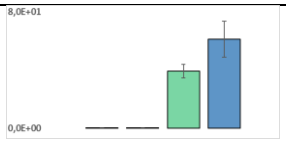
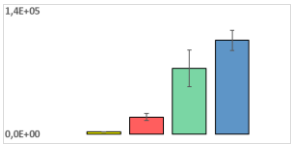
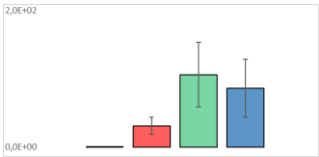
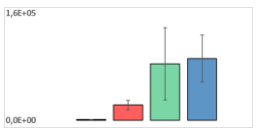
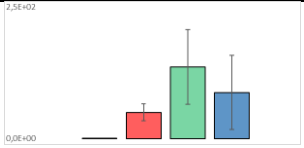


Figure S3. ESI - PLS-DA Scores plots of the samples. A: *Laboratory #1*; R^2X (cum) = 0.682; R^2Y (cum) = 0.954; Q^2 (cum) = 0.871. B: *Laboratory #2*; R^2X (cum) = 0.235; R^2Y (cum) = 0.989; Q^2 (cum) = 0.906. Green dots: "EVOO" samples; blue dots: "OTHER" samples.

Section 3 – Compound identification

The complete list of the supposed names of the 12 selected compounds is presented in Table S2, together with their identification level (Schymanski, et al., 2014) and their mean area values in the different groups. The "Variables trends plot" of "Compound 1" and of "Compound 7" are presented in figure S4.

Compound ID	Name	ID Level	*Lab#1* Area values	*Lab#2* Area values
1	Propylene glycol - 1 Stearate	3		
2	4-Phenylbutyric acid	2		N.D.
3	Tyrosine ethyl ester	3		N.D.
4	Geranic Acid	3		N.D.
5	3,4,5-trimethoxydihydrocinnamic acid	3		N.D.
6	Propyl-12-hydroxy-9-octadecenoate	3		N.D.
7	N.A.	4		
8	N.A.	4		
9	(2R,3E)-5-(3-Chloro-5-formyl-2,6-dihydroxy-4-methylphenyl)-3-methyl-1-[(1S,2R,6R)-1,2,6-trimethyl-3-oxocyclohexyl]-3-penten-2-yl acetate	2		
10	N.A.	4		

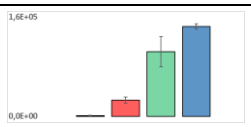
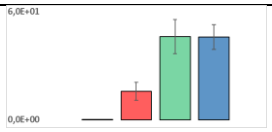
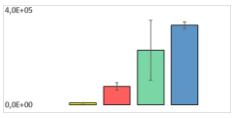
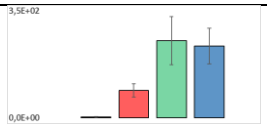
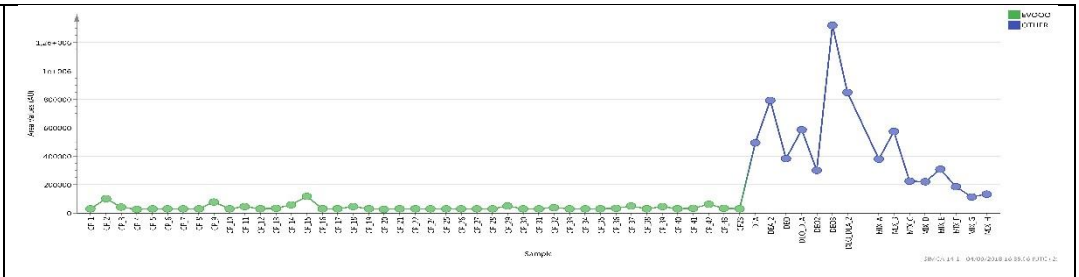
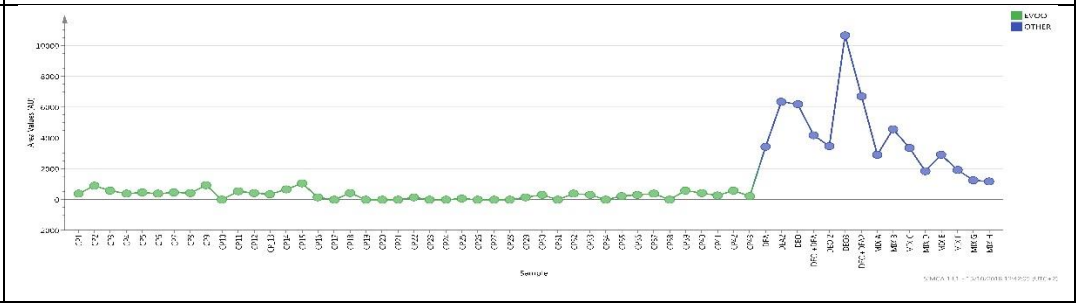
11	N.A.	4		
12	2,3,4-Trihydroxy-6-methyl-5-[(2E,6E)-3,7,11-trimethyl-2,6,10-dodecatrien-1-yl]benzaldehyde	3		

Table *S2*: compounds description and comparison of their mean area values (+/− standard error) through the groups (yellow bar: EVOO; red bar: DEO; green bar: DEA, blue bar: DEO+DEA samples) N.A.= Not available

COMPOUND 1

LABORATORY #1	
LABORATORY #2	

COMPOUND 7

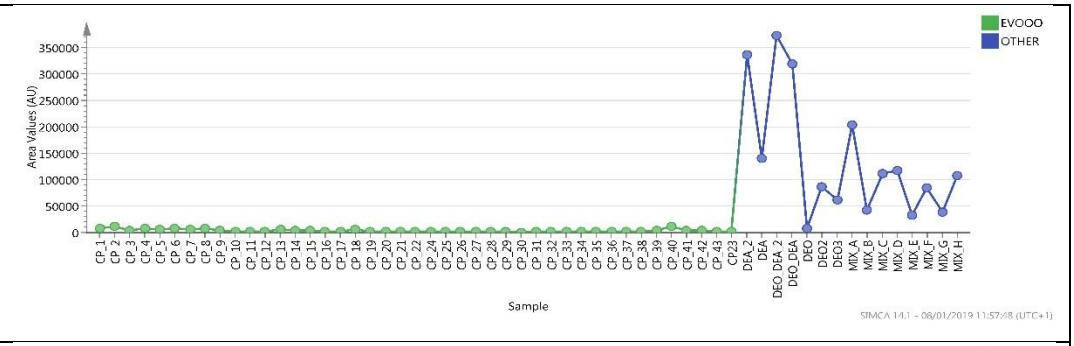
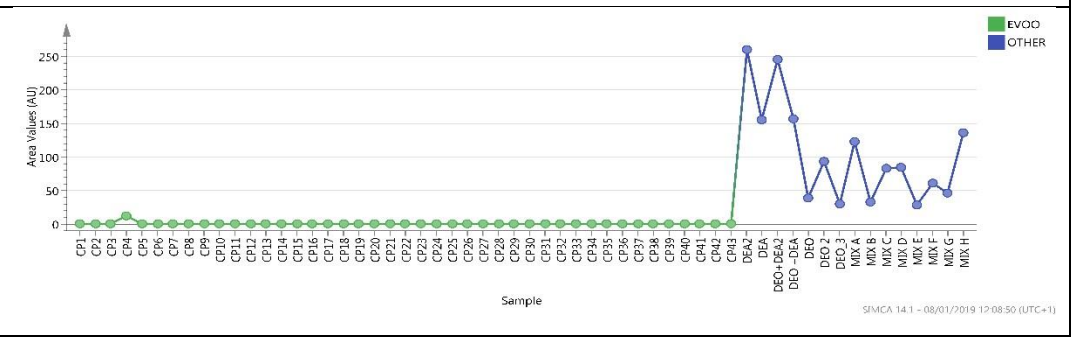
LABORATORY #1	
LABORATORY #2	

Figure S4. Variable trend plots of "Compound 1" and of "Compound 7". Green dots: area values of the marker in "EVOO" group; blue dots: area values of the marker in "OTHER" group.

Section 4 – Evaluation of the Inter-laboratory Study

The comparison of the MS/MS spectra of "Compound 1" and of "Compound 12" obtained with the two instrument are presented, as example, in figure S5 and figure S6.

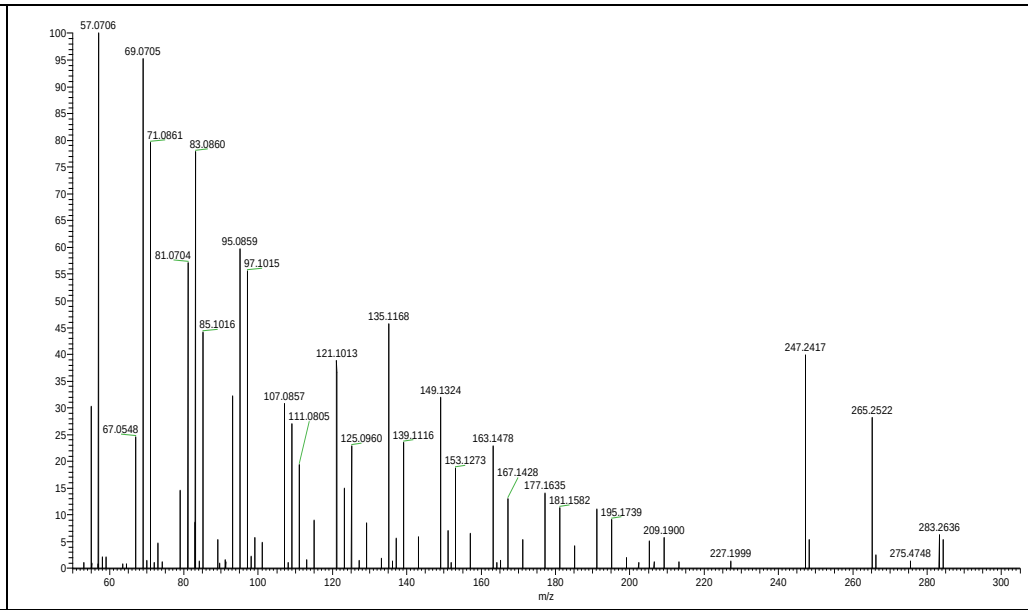
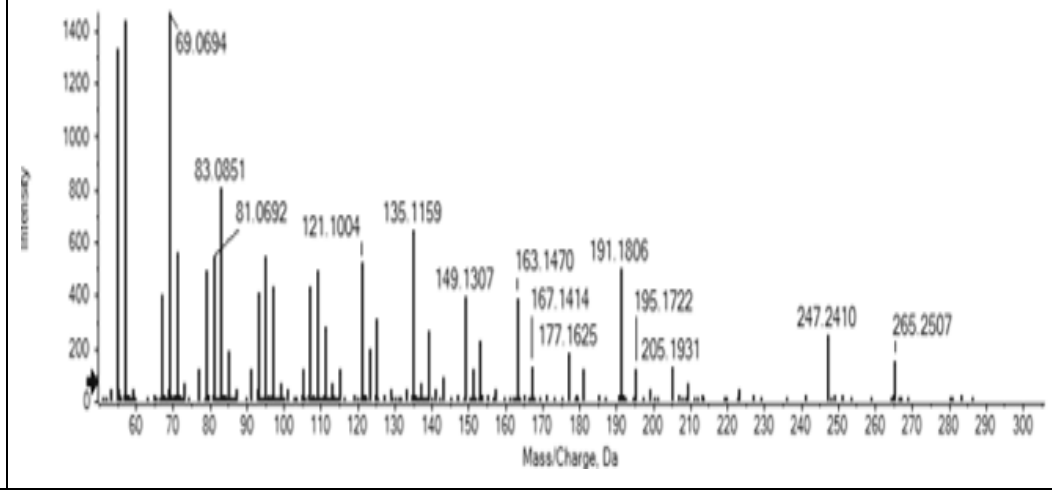
LABORATORY #1	
LABORATORY #2	

Figure S5 Comparison of the MS/MS spectra of the "Compound 1" obtained in *Laboratory #1* and in *Laboratory #2*. Instrumental conditions are detailed in the text

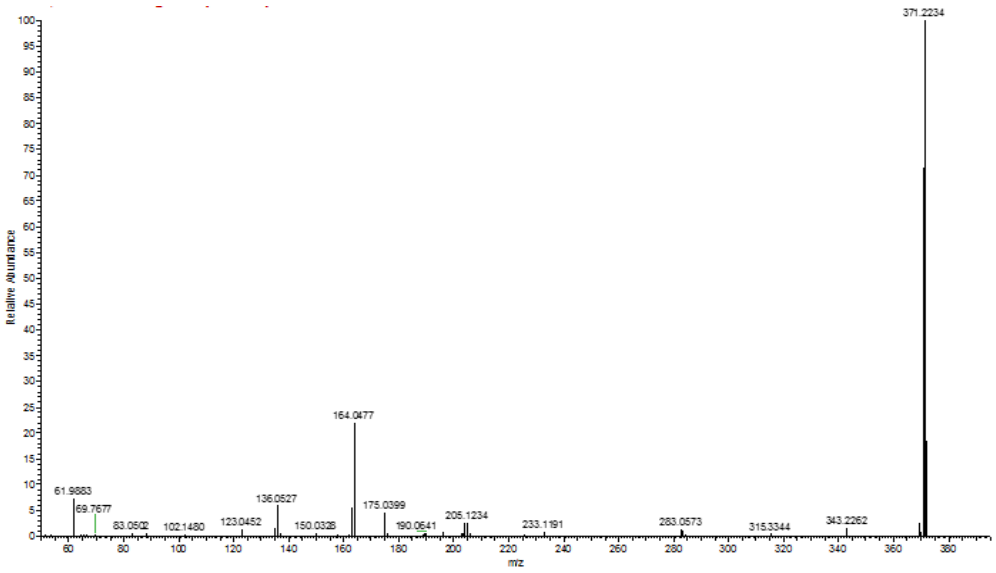
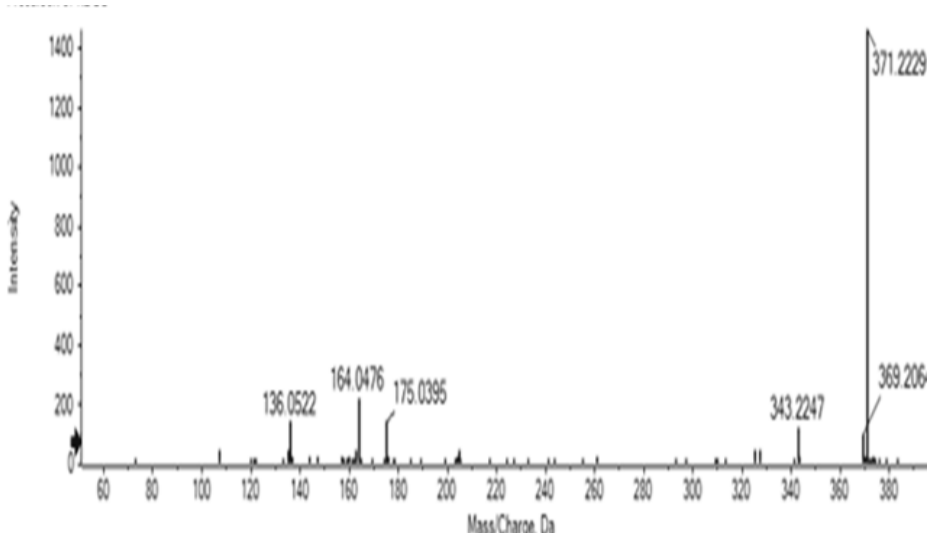
LABORATORY #1	
LABORATORY #2	

Figure S6 Comparison of the MS/MS spectra of the "Compound 12" obtained in *Laboratory #1* and in *Laboratory #2*. Instrumental conditions are detailed in the text

Bibliographic references

https://metlin.scripps.edu/landing_page.php?pgcontent=mainPage

Alonso-Salces, R., Moreno-Rojas, J., Holland, M., Reniero, F., Guillou, C., & Héberger, K. (2010). Virgin Olive Oil Authentication by Multivariate Analyses of 1H NMR Fingerprints and δ13C and δ2H Data. *Journal of Agricultural and Food Chemistry, 58*(9), 5586-5596.

Aparicio, R., & Alonso, V. (1994). Characterization of virgin olive oils by SEXIA Expert System. *Progrss in Lipid Research, 33*, 29-38.

Aparicio-Ruiz, R., Romero, I., García-González, D., Oliver-Pozo, C., & Aparicio, R. (2017). Soft-deodorization of virgin olive oil: Study of the changes of quality and chemical composition. *Food Chemistry, 220*, 42-50.

Bontempo, L., Paolini, M., Franceschi, P., Ziller, L., García-González, D., & Camin, F. (2019). Characterisation and attempted differentiation of European and extra-European olive oils using stable isotope ratio analysis. *Food Chemistry, 276*, 782-789.

Cajka, T., Smilowitz, J., & Fiehn, O. (2017). Validating quantitative untargeted lipidomics across nine liquid chromatography–high-resolution mass spectrometry platforms. *Analytical Chemistr, 89*(22), *12360-12368*.

Camin, F., Larcher, R., Perini, M., Bontempo, L., Bertoldi, D., Gagliano, G., . . . Versini, G. (2010). Characterisation of authentic Italian extra-virgin olive oils by stable isotope ratios of C, O and H and mineral composition. *Food Chemistry, 118*(4), 901-909.

Cavanna, D., Catellani, D., Dall'Asta, C., & Suman, M. (2018). Egg product freshness evaluation: a metabolomic approach. *Journal Of Mass Spectrometry 53*(9), *849-861*.

Cavanna, D., Righetti, L., Elliott, C., & Suman, M. (2018). The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach. *Trends in Food Science & Technology, 80*, 223-241.

Cecchi, L., Innocenti, M., Melani, F., Migliorini, M., Conte, L., & Mulinacci, N. (2017). New isobaric lignans from Refined Olive Oils as quality markers for Virgin Olive Oils. *Food Chemsitry, 219*, 148-157.

Cerretani, L., Bendini, A., Barbieri, S., & Lercker, G. (2008). Preliminary observations on the change of some chemical characteristics of virgin olive oils subjected to a "soft deodorization" process. *Rivista Italiana delle Sostanze Grasse, 85*(2), 75-82.

EU, E. U. (2016). Commission Delegated Regulation (EU) 2016/2095 of 26 September 2016 amending Regulation (EEC) No 2568/91 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis. *Official Journal Of the European Union*, p. L326/1-6.

Faberi, A., Marianella, R., Fuselli, F., La Mantia, A., Ciardiello, F., Montesano, C., . . . Compagnone, D. (2014). Fatty acid composition and δ13 C of bulk and individual fatty acids as marker for authenticating Italian PDO/PGI extra virgin olive oils by means of isotopic ratio mass spectrometry. *Journal of Mass Spectrometry, 49*, 840-849.

Frankel, E. (1987). Secondary products of lipid oxidation. *Chemsitry and physics of lipids, 44*, 73-85.

Frankel, E. (1991). Review.Recent advances in lipid oxidations. *Journal of the Science of Food and Agriculture, 54*(4), 495 - 511.

Garcia-Gonzalez, D., Aparicio, R., & Aparicio-Ruiz, R. (2018). Olive oil. In J. Morin, & M. Lees, *FoodIntegrity Handbook: A guide to food authenticity issues and analytical solutions* (p. 336-357). Nantes.

Gil-Solsona, R., Raro, M., Sales, C., Lacalle, L., Díaz, R., Ibáñez, M., . . . Hernández, F. (2016). Metabolomic approach for Extra virgin olive oil origin discrimination making use of ultra-high performance liquid chromatography – Quadrupole time-of-flight mass spectrometry. *Food Control, 70*, 350-359.

Hurkova, K., Rubert, J., Stranska-Zachariasova, M., & Hajslova, J. (2017). Strategies to Document Adulteration of Food Supplement Based on Sea Buckthorn Oil: a Case Study. *Food Analytical Methods, 10*, 1317-1327.

IOC, I. O. (2016). *COI/T.15/NC No 3/Rev. 11 - Trade standard applying to Olive Oils and Olive-Pomace Oils.* Madrid, Spain.

Kalogiouri, N., Alygizakis, N., Aalizadeh, R., & Thomaidis, N. (2016). Olive oil authenticity studies by target and nontarget LC–QTOF-MS combined with advanced chemometric techniques. *Analytical and Bioanalytical Chemistry, 408*, 7955-7970.

Martin, J., Maillot, M., Mazerolles, G., Verdu, A., Lyan, B., Migne´, C., . . . Rutledge, D. (2015). Can we trust untargeted metabolomics? Results of the metabo-ring initiative, a large-scale, multi-instrument inter-laboratory study. *Metabolomics, 11*(4), 807-821.

Pérez-Camino, M., Cert, A., Romero-Segura, A., Cert-Trujillo, R., & Moreda, W. (2008). Alkyl esters of fatty acids a useful tool to detect soft deodorized olive oils. *Journal of Agricultural and Food Chemistry, 56*(15), 6740-6744.

Rubert, J., Zachariasova, M., & Hajslova, J. (2015). Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review. *Food Additives & Contaminants: Part A, 32*(10), 1685-1708.

Schymanski, E., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology, 48*, 2097-2098.

Serani , A., & Piacenti, D. (2001). Sistema analytico per l'dentificazione di oli deodorati in oli vergini di oliva – Nota 1 – Analisi dei pigmenti clorofilliani in oli vergini di oliva. *Rivista italiana delle sostanze grasse, 78*, 459-463.

Serani, A., Piacenti, D., & Staiano, G. (2001). Sistema analytico per l'dentificazione di oli deodorati in oli vergini di oliva – Nota 2 – Cinetica di isomerizzazione dei digliceridi in oli vergini di oliva. *Rivista italiana delle sostanze grasse, 78*, 525-528.

Sumner, L., Amberg, A., Barrett, D., Beale, M., Beger, R., Daykin, C., . . . Viant, M. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics, 3*, 211–221. doi:10.1007/s11306-007-0082-2

Vaclavik, L., Cajka, T., Hrbek, V., & Hajslova, J. (2009). Ambient mass spectrometry employing direct analysis in real time (DART) ion source for olive oil quality and authenticity assessment. *Analytica Chimica Acta, 645*, 56-63.

Vetter, W., Schröder, M., & Lehnert, K. (2012). Differentiation of Refined and Virgin Edible Oils by Means of the trans- and cis-Phytol Isomer Distribution. *Journal of Agricultural and Food Chemistry, 60*, 6103-6107.

Extra Virgin Olive Oil: industrial benefits and future perspectives

Rapid approaches

At the moment, in the laboratories there are no rapid ways that should help in the detection of these type of adulterated oils.

The promising results presented in this chapter represent a first step of a process that could potentially lead to the implementation of one of the two rapid methods directly in the production plants, in order to intercept adulterated or extremely aged oils.

A larger set of samples will be collected and the number of adulterated samples will be increased, in order to strengthen the models and confirm the preliminary results obtained.

Subsequently, following the same approach presented in the "egg products" chapter, a test directly in the QC lab will be executed.

Confirmatory study

The routine research of the molecules that were identified as responsible of soft refinement processes represents an additional tool that the company will be able to put in place when suspect samples will be intercepted by Quality Control laboratories.

Further studies will be executed in order to identify without any reasonable doubt the selected markers; subsequently, a novel targeted method will be developed by a third part laboratory, with the aim to set a concentration limit for each of the compounds

CHAPTER 3

DURUM WHEAT

Durum Wheat

General overview of the product

According to the FAO's definition, the term "cereals" refers only to crops harvested for dry grain; within this scenario, wheat represents one of the most important ones, with over 700 million tonnes being harvested worldwide annually.

All wheat species derive from the genus *Triticum*, and the most important are *Triticum Aestivum* (also known as common wheat and usually employed to make bread or other baked goods) and *Triticum Turgidum* (also known as durum wheat and usually employed to make pasta) and the latter will be the species studied in this chapter.

Durum wheat is largely produced in Italy, while other major producing countries are France, Greece, Spain and USA (Morin, Vermeulen, & Maestri, 2018).

Product Identity

Current standards of identity or related legislation (Codex, EU, US, ISO, Trade associations…)

Reg (EEC) 742/2010 - a regulation for the establishment of the eligibility criteria to be met by cereals for public interventions and the methods that should be used for the assessment of this eligibility

Council Directive 2009/74/EC – it establishes rules on the production, packaging, sampling, sealing and marking of cereals.

In addition. there are different regulations, directives and recommendations at national level: for example, in the Italian regulation, the presidential decree N°187 (D.P.R. n 187, 9 February 2001) set the maximum content of common wheat in durum wheat products.

Moreover, the Codex Alimentarius published a document that resumes all the texts adopted by its commission up to 2007 (FAO/WHO, 2007).

Authenticity issues

The **most risky factors** that can have impact on the durum wheat authenticity are:

- **Species substitution**

  It is the deliberate substitution with cheaper species or varieties, as for example the addition of common wheat to durum wheat

- **Geographical provenance**

  The geographical origin is sometimes presented as an added value of the product but there are no standard methods that are able to assess this parameter; for this reason, fraudulent origin declarations could affect the supply chain.

- **Certification of organic production**

  A wheat product can be labelled as "organic" only if the producers followed specific international rules that lead to an higher cost of the production and subsequently of the final product.

  A fraud happens when cheaper non-organic products are sold as organic

Aim of the work

In the following chapter, a non-targeted metabolomic method will be presented as a new tool able to assess the geographical origin of durum wheat.

This chapter is going to be submitted to "*Food Chemistry"*. For additional details see section "Author".

Bibliographic references

D.P.R. n 187, 9 February 2001-Regolamento per la revisione della normativa sulla produzione e commercializzazione di sfarinati e paste alimentari, a norma dell'articolo 50 della legge 22 febbraio 1994, n. 146. (2001). Official Italian Journal, 117, p. 6-12.

Joint FAO/WHO Codex Alimentarius Commission, World Health Organization, Food and Agriculture Organization of the United Nations & Joint FAO/WHO Food Standards Programme, eds. (2007). – *Codex alimentarius: cereals, pulses, legumes and vegetable proteins*. 1st ed, World Health Organization : Food and Agriculture Organization of the United

Morin, J., Vermeulen, P., & Maestri, E. (2018). Cereals and cereal-based products. In J. Morin, & M. Lees, *FoodIntegrity Handbook. A guide to food authenticity issues and analytical solutions* (p. 101-114). Nantes: Eurofins Analytics France.

A non-targeted high-resolution mass spectrometry approach for the assessment of the geographical origin of durum wheat

Daniele Cavanna [1,2], Cecilia Loffi [1,2], Chiara Dall'Asta [2], Michele Suman [1]*

[1]Advanced Laboratory Research, Barilla G. e R. Fratelli S.p.A., Via Mantova 166-43122 Parma (Italy)
[2]Department of Food and Drug, University of Parma, Parco Area delle Scienze 95/A-43124 Parma,(Italy)

Introduction

Durum wheat (triticum turgidum subsp. durum) is the preferred raw material for the production of pasta due to its technological and nutritional properties (Pauly, Pareyt, Fierens, & Delcour, 2013).

The declaration of its geographical origin, sometimes claimed on pasta labels, represents an added value to the commodity, especially if it is made only with wheat coming from a specific nation or region (i.e. "100% Italiano").

The chemical composition of wheat, indeed, is partially related to cultivation soil and to climate conditions, that are different among different geographical areas and thus can affect the quality and subsequently the commercial value of this raw material (Zhao, et al., 2011).

For these reasons, the false origin declaration of this commodity or the creation of fraudulent mixtures of wheat coming from different countries could represent an illegal economical gain for wheat sellers and that is why, in the last decade, pasta producers made a special effort to develop methods able to assess the geographical origin of wheat (Morin, Vermeulen, & Maestri, 2018).

Several papers in literature suggest the use of light stable isotopes (i.e. $\delta^{15}N$, $\delta^{13}C$, $\delta^{18}O$) (Rashmi, Shree, & Singh, 2017) (Wadood , Boli, & Yimin, Geographical traceability of wheat and its products using multielement light stable isotopes coupled with chemometrics, 2019), the Sr isotopic ratios (Liu , et al., 2017) or a combination of them (Liu, et al., Combination of the 87Sr/86Sr ratio and light stable isotopic values (d13C, d15N and dD) for identifying the geographical origin of winter wheat in China, 2016) as interesting tools to face this problem; however, according to the scientific papers studied, the feasibility is proven only taking into account a small amount of geographies.

Recently, also a near-infrared and chemometric approach was presented as a possible tool for the discrimination of wheat kernels coming from different China regions (Wadood , Guo, Zhang, & Wei, 2019), but again the geographical regions explored are limited.

Another interesting tool is the use of non-targeted high resolution mass spectrometry approaches: as largely reported in the literature, these studies are able to record a fingerprint of the raw material and,

after data elaboration, some markers responsible of specific groups can be selected and identified (Castro-Puyana, Perez-Míguez, Montero, & Herrero, 2017).

Metabolomic studies on wheat are presented in literature for different purposes, like, for example , the detection of fraudolent addition of common wheat to durum wheat (Righetti, et al., A novel approach based on untargeted lipidomics reveals differences in the lipid pattern among durum and common wheat, 2018) or for the discrimination of ancient grains (Righetti, et al., Characterization and Discrimination of Ancient Grains: A Metabolomics Approach, 2016) but, as far as we know, this approach was not used for wheat geographical discrimination.

Metabolic fingeprint, however, is close to the plants phenotype and is very sensitive to exogenous factors, like growing location, weather or soil composition (Oldiges, et al., 2007) (Creydt & Fischer, Omics approaches for food authentication, 2018) and, for these reasons, could represent a promising tool for the assessment of the geographical origin of specific raw materials (Cajka, Showalter, Riddellova, & Fiehn, 2016) (Saia, Fragasso, De Vita, & Beleggia, 2019).

Literature, indeed, presents several examples of non-targeted analyses applied for this purpose, as reported, for example, for Extra Virgin Olive Oils (Gil-Solsona, et al., 2016) (Ghisoni, et al., 2019), Asparagus (Creydt, Hudzik, Rurik, Kohlbacher, & Fischer, 2018) or Parmigiano Reggiano (Popping, De Dominicis, Dante, & Nocetti, 2017).

Despite the technique used, a method for the assessment of geographical origin can be considered robust only if the inter-year variability is considered during the experimental design set up: the samples used for the model creation and for the model validation should come from different harvesting years; otherwise, the results obtained should not be valid if applied on samples coming from a new production campaign. This is due to the fact that many critical factors related to the soil and to the climate conditions may change during the years, maybe affecting the non volatile profile of the wheat. For these reasons, a model can be considered robust only if it works also including samples subjected to these natural changes (Camin, et al., 2017).

On the basis of what described above, this work presents a non-targeted Liquid Chromatography – High Resolution Mass Spectrometry study for the detection of new chemical markers responsible of the geographical origin of wheat and able to discriminate if the raw material comes from Italy, Europe or Extra Europe.

Authentic samples coming from different Italian, European and Extra European regions were collected during two years of harvesting, 2016 and 2018: the 2016 campaign was used for model creation and for markers selection, while the 2018 campaign was used for model and markers validation.

Materials and Methods

Chemicals

Methanol and Ammonium Formate, were purchased from VWR International, Ltd (Poole, United Kingdom).

Formic Acid and the analytical standard Chloramphenicol were purchased from Sigma Aldrich (St. Louis, MO).

Water was purified using a Milli-Q system (Millipore, Bedford, MA).

Wheat Samples

Wheat from two harvesting years (2016 and 2018) were selected with the help of a third part company authorized to perfom official sampling; therefore, samples were stored in the laboratory with an official certification of their geographical origin.

91 samples were collected during the 2016 campaign (56 from Italy, 21 from other European countries and 14 from Extra European countries) and 49 samples were collected during the 2018 campaign (11 from Italy, 21 from other European countries and 17 from Extra European countries).

The complete list of samples is presented in supplementary materials (table S1).

Samples were stored at 4°C and were cleaned from foreign materials using specific sieves; subsequently, after a successive visual inspection, only the intact grains were selected.

Ultimately, samples were grinded with a Knife Mill GRINDOMIX GM 200 (Retsch, Germany) before starting the extraction procedure.

Sample Extraction

1 gram of grinded wheat was transferred in a 50 ml falcon tube, together with 250 µl of a 20 µg/ml Chloramphenicol solution in methanol (used as internal standard) and 10 ml of a mixture of Methanol:water in the ratio 50:50 (v/v).

Sample was then shaken for 30 min at 240 RPM using an inclined rotating plane agitator (InterContinental, Italy) and then centrifuged at 5000 RPM for 5 min with a Rotina 380R (Hettic lab technology, Germany).

Finally, the extract was filtered into an HPLC vial with a 0.22 µm PTFE syringe filter (Phenomenex, Torrance, CA) and stored at -20°C until analysis with LC-HRMS instrument.

For the evaluation of method reproducibility, 20% of the samples were double prepared, as detailed in supplementary materials (table S1).

According to the common metabolomic approach presented in literature, extraction blanks were collected and quality control (QC) samples were created mixing 10 µl of each sample extract (Godzien, Alonso-Herranz, Barbas, & Armitage, 2015).

LC-HRMS analysis

Chromatographic separation was performed with a Dionex UltiMate 3000 UHPLC (Thermo Fisher Scientific, Inc., Waltham, MA) equipped with a BEH C18 150 x 2.1 mm, 1.7 µm particle size analytical column (Waters, Milford, MA, USA) maintained at 40°C. Formic Acid (FA) and Ammonium Formate (AF) were used as mobile phase modifiers during the gradient elution, while the flow rate of the run was kept at 0.3 ml/min.

Gradient conditions were set as follows: after 1 minute with 95 % of mobile phase A (0.1% FA and 5mM AF in Water) and 5% of mobile phase B (0.1% FA and 5mM AF in Methanol), the percentage of solvent B increased to 100% in 8 minutes and then was maintained at this percentage for 5 minutes before column re-equilibration (7 minutes).

Autosampler was maintained at 18 °C, while the injection volume was 4 µl.

Mass spectrometry detection was performed with a benchtop Q Exactive™ Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific, Waltham, MA) equipped with a heated (H) electrospray ionization (ESI) interface (Thermo Fisher Scientific,Waltam, MA). Two analytical sequences (one with positive and one with negative ionization mode) were executed.

The heather temperature of the ion source was set up at 290 °C while the capillary was kept at 270 °C. All the other HESI parameters are the same described in a previous work (Cavanna, Catellani, Dall'Asta, & Suman, 2018).

Both the sequences were collected performing a "full scan" experiment, analyzing a QC sample every ten injections and with all the other samples randomly injected. For both the sequences, the inspected mass range was between 75 and 1000 Da, with a resolution of 140000 FWHM (m/z 200), automatic gain control (AGC) target at 1e6 and maximum injection time (IT) 200 ms.

At the end of the sequence, QC samples were also analyzed with a "full scan-data dependent" approach in order to collect the MS/MS spectra and try to perform features identification. For this experiment, the full-scan resolution was 70000 FWHM (m/z 200), while 17500 FWHM (m/z 200) was the resolution used for the MS/MS experiment.

The intensity threshold set for the execution of a fragmentation acquisition was 6e4. The quadrupole isolation window was kept at 2.0 m/z with a TopN value of 5. The scan range was from 50 to the

fragmented mass m/z (m/z +25), AGC target 2e5, maximum IT 50 ms, normalized collision energy (NCE) 30% with ±50% step.

Data elaboration

Raw data were acquired using Xcalibur software (version 3.0, Thermo Fisher Scientific, Waltham, Massachusetts).

Peaks alignment, "extraction blanks" subtraction and features extraction were performed using Compound Discoverer software (version 2.1 SP1, Thermo Fisher Scientific, Waltham, Massachusetts); the mass range inspected was between 75 m/z and 1000 m/z from 1 to 14 minutes of the chromatographic runs; all the other critical parameters are the same presented in a recent work (Cavanna, Catellani, Dall'Asta, & Suman, 2018) with the exception of the "precursor ion deviation" parameter, that was set to 5 ppm for both positive and negative ionization modes.

The two data matrixes (for positive and negative acquisitions) were exported and processed with SIMCA software (version 14.1 Umetrics, Umea, Sweden) for chemometric data elaboration.

Data were log transformed and pareto scaled, then preliminary Principal Component Analysis (PCA) models were executed for the evaluation of the positioning of QCs and replicates in the scores plot.

Subsequently, replicate extracts were averaged, features were selected and only the ones that had a CV% lower than 40% in the QC samples were included for further studies.

Samples set was divided in two groups: the samples coming from the 2016 campaign (training set) were used for the models creation and for markers selection, while the 2018 samples (test set) were used for models and markers validation.

Two predictive orthogonal partial least square discriminant analysis (OPLS-DA) models were created classyfing the samples as "Italian", "European" and "Not European". The goodness of the chemometric models was firstly evaluated with an internal leave 1/7 out cross-validation and then with the 2018 samples, used as external set.

For markers selection, different two classes OPLS-DA models were built comparing the "Italian" samples against the "European" samples and against the "Not European" samples; subsequently, using the S-plots, statistically significant markers responsible of the clusterization were selected (Jandric, Islam, Singh, & Cannavan, 2017).

In addition, their VIP (variable influence on projection) values were evaluated to assess their relevance for the chemometric model (VIP values had to be >1.4) (Akarachantachote, Chadcham, & Saithanu, 2014).

Features that survived this process were studied, and a tentative of compounds identification was performed.

The identification process was the same described in other works (Cavanna, Catellani, Dall'Asta, & Suman, 2018).

Markers validation

The compounds selected with the workflow detailed above were searched in the samples coming from the 2018 harvest campaign that included, especially for the european areas, an higher variability in the geographical origins.

The reliability of these molecules as "geographical" markers implies they have to be confirmed not only with their presence or absence in the new samples but also highlighting the same intensity trend through the geographical areas (Italy, Europe and Extra Europe).

Results and Discussion

Evaluation of the analytical process

Before starting the non targeted study, an evaluation of the analytical process was executed studying the behavior of Chloramphenicol, the internal standard that was added at the beginning of the extraction procedures (Want, et al., 2013). The instrumental response was acceptable during the whole sequences (approximately 200 injections each) for both the ionization modes.

The details of the obtained results are presented in the Supplementary Materials – Section 2.

Multivariate data elaboration

For data elaboration, the whole analytical sequences (including wheat coming from two years of harvesting, extraction replicates, blanks and QCs) were processed.

After peaks alignment, blank subtraction and isotope merging, 4224 features were detected in positive mode and 1538 in negative mode.

The first explorative PCA highlighted a good clustering of the QCs approximately in the center of the scores plot, as presented in figure 1 for negative acquisition and in figure S1 (Supplementary Materials) for positive acquisition.

The replicate extracts approximately overlaid in the graph and this is a further certification of the repeatability of the extraction procedure; furthermore, a partial separation of the "Italian" samples from the others seems visible.

After data filtering (obtained by selecting the features with a CV lower than 40% in QC samples), two predictive three classes OPLS-DA models (obtained with the samples from the 2016 campaign) were created using 2996 features in positive and 1344 features in negative.

The separation dramatically increased with a clear clustering of the three groups, as presented in Figure 2 for positive acquisition mode and in figure S2 (Supplementary Materials) for negative acquisition mode.

The variance of the x and y variables explained by the model (R^2X (cum) and R^2Y (cum)), and the cumulative predicted variation in the Y matrix (Q^2 (cum)) are reported in the plots as well.

Subsequently, models were tested with the samples used for their creation and with the 2018 ones, treated as test set.

The good predictive capabilities, especially for those obtained with data acquired in positive ionization mode, are certified by the fact that all 2016 samples were correctly classified, together with approximately 88% of the samples from 2018 harvest, that were treated as unknown samples. The details are presented in Table 1.

Despite these encouraging results, non-targeted mass spectrometric approaches should lead to the selection of chemical markers responsible of the clusterization, so as to make the obtained results easily transferable into other laboratories (Cavanna, Righetti, Elliott, & Suman, 2018).

In order to reach this goal, different two classes OPLS-DA models were created comparing "Italian" samples against "European" samples and against "Not European" ones, using related S-plot and VIP values tools for features selection (Black, Haughey, Chevalier, Galvin-King, & Elliott, 2016).

All these models showed a clear separation of the groups: as example, figure 3 presents the OPLS-DA created between "Italian" and "European" samples acquired in positive ionization mode, while table 2 summarizes the values of R^2X (cum), R^2Y (cum) and Q^2 (cum) obtained for all the models.

Markers identification

Features responsible of clusterizations were selected, and for a group of them an attempt of compounds identification was executed.

Markers were classified as "Italian" or "Not Italian" depending on whether they were present only or mostly in Italian or in not Italian samples; in addition, two molecules were selected as discriminant markers for the "Extra European" geographical origin of wheat.

Tables 3, 4 and 5 summarize the compounds selected, ranked according to the level of identification suggested by the Standard Initiative in Metabolomics (Sumner, et al., 2007) and subsequent improvements (Schymanski, et al., 2014).

Due to lack of full chemical identification, arbitrary names were assigned to the majority of the compounds (Wielogorska, et al., 2018); on the contrary, the compounds identified as Sucrose 3-lauroyl-3'-isovaleric acid and 3-Methylenenorleucine seems plausible, taking into account that their precursors are commonly found in wheat (Quílez, Ruiz, & Romero, 2006), as well as the compounds that are part of the metabolic pathways of Glucose and Mannose.

Markers validation

As also suggested by US Pharmacopoeia (U.S. Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection, 2016) and as presented in literature (Jandrić, et al., 2014) (Andersen, et al., 2014), markers should be validated with an external set of samples not used for the model creation.

For this reason, all the selected features were searched in the 2018 wheat samples and their presence or absence, together with their trends across the different geographies, were evaluated and compared with the results obtained during the model creation, as presented in Tables 6, 7 and 8.

All the target molecules were detected with almost the same trend through the different geographies, certifying that the compounds identified are robust markers for the assessment of the geographical origin of wheat.

Conclusions

The current study presented for the first time the possibility to discriminate the geographical origin of wheat using a metabolomic approach, assessing its origin as 100% Italian, European or Extra European.

Different regions worldwide were included in the study, covering different continents and assuring in this way that the results obtained could be considered general and not related to specific nations; moreover, the evaluation of the inter year variability was included in this study.

A predictive model was created and subsequently target compounds were selected as markers of a specific origin.

Both model and markers were validated with an external set of samples not used for the model creation and related to a different year of harvesting, certifying the robustness of the results obtained.

Further improvements of this study should lead to increase the identification level of the markers and to the development of target methods focused on the detection of these specific molecules, that could be easily applied in quality control laboratories.

Tables

	ITALY	EUROPE	EXTRA EUROPE
% Correctly classified in the training set	100%	100%	100%
% Correctly classified in the test set	91%	90%	82%

Table 1. Predicted group affinity of training and test set with the ESI+ OPLS-DA model.

MODEL	IONIZATION MODE	LATENT COMPONENT	ORTHOGONAL COMPONENT	R^2X (cum)	R^2Y (cum)	Q^2(cum)
"Italy" vs "Europe"	ESI +	1	3	0.381	0.985	0.923
	ESI -	1	1	0.415	0.730	0.673
"Italy" vs "Not Europe"	ESI +	1	3	0.373	0.989	0.939
	ESI -	1	5	0.576	0.980	0.854

Table 2. Global value of R^2X (cum), R^2Y (cum), and Q^2 (cum) parameters for the two classes OPLS-DA models

Name	Pseudomolecular ion	Detected m/z	RT (min)	Predicted Formula	Mass error (ppm)	CV IN QCs (%)	VIP Value	ID type
Compound 1	$[M-H]^+$	434.2016	6.30	C19 H31 N O10	0.92	15	2.21[a]	4
Compound 2	$[M-H]^+$	342.1762	5.41	C13 H27 N O9	0.88	6	3.03[b]	4
Compound 3	$[M-H]^+$	624.3594	8.55	N.A.	N.A.	10	3.33[b]	5
Compound 4	$[M+H]^-$	462.1625	5.70	C20 H25 N5 O8	1.08	7	2.62[a]	4
Compound 5	$[M+H]^-$	794.3702	6.14	N.A.	N.A.	25	2.22[a]	5
Compound 6	$[M-H]^-$	583.3344	7.53	C27 H52 O13	1,54	15	2.57[b]	4
Compound 7	$[M-H]^-$	625.3446	7.95	C29 H54 O14	0.96	8	2.73[b]	4
Compound 8	$[M-H]^-$	383.1142	8.00	C21 H20 O7	1.57	13	2.81 [a]	4
Sucrose 3-lauroyl-3'-isovaleric acid	$[M-H]^-$	607.3336	8.15	C29 H52 O13	0.16	10	2.90[b]	3
Compound 9	$[M+H]^-$	449.1016	8.02	C22 H23 Cl O8	1.78	13	2.15[a]	4
Compound 10	$[M-H]^-$	476.1201	8.02	C23 H19 N5 O7	2.31	16	2.74[a]	4
Compound 11	$[M-H]^-$	527.1176	8.02	C22 H20 N6 O10	3.60	33	2.51[a]	4
β-D-Fructofuranosyl 2-O-dodecanoyl-3-O-(3-methyl-2-butenoyl)-α-D-glucopyranoside	$[M-H]^-$	605.3187	8.14	C29 H50 O13	1.49	11	2.99[a]	3
Compound 12	$[M-H]^-$	589.3232	8.39	C29 H50 O12	0.34	16	2.67[a]	4
Compound 13	$[M+H]^-$	563.3084	8.64	C27 H48 O12	1.95	20	2.46[a]	4
(2R,3S)-2,3,4-Trihydroxybutyl 4-O-acetyl-2-O-propionyl-3-O-tetradecanoyl-β-D-mannopyranoside	$[M+H]^-$	591.3399	9.05	C29 H52 O12	2.20	20	2.88[a]	3

Abbreviations: QC, quality control; RT, retention time; VIP, variable influence on projection; N.A., not available

Table 3. "Not Italian" Markers. [a]= VIP value obtained from the OPLS-DA model "Italy" vs "Non Europe" ; b= VIP value obtained from the OPLS-DA model "Italy" vs "Europe"

Name	Pseudomolecular ion	Detected m/z	RT (min)	Predicted Formula	Mass error (ppm)	CV IN QCs (%)	VIP Value	ID type
Compound 14	$[M-H]^-$	187.0615	4.33	C8 H12 O5	1.60	6	1.70	4
Compound 15	$[M-H]^-$	669.3949	7.58	C28 H59 N6 O10 P	1.34	38	2.33	4

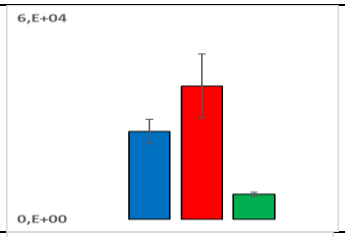
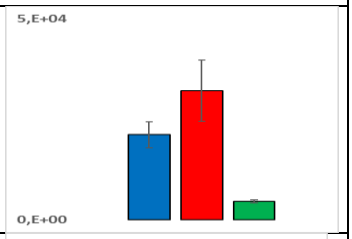
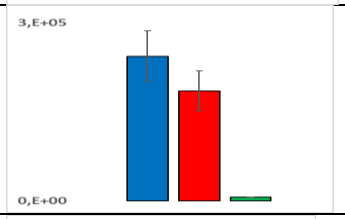
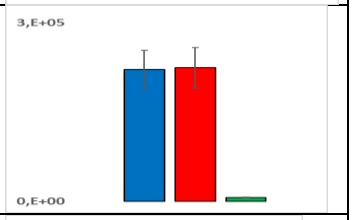
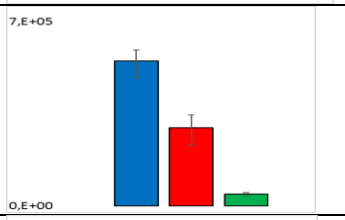
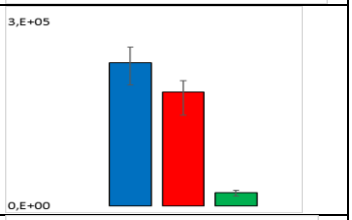
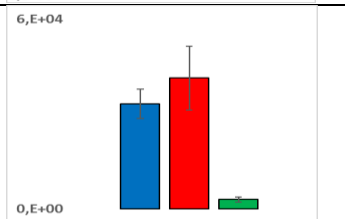
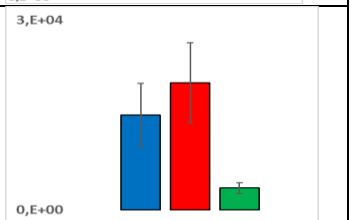
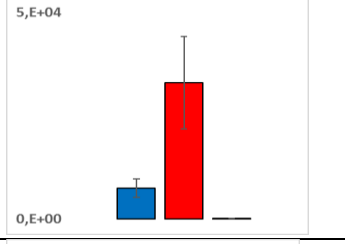
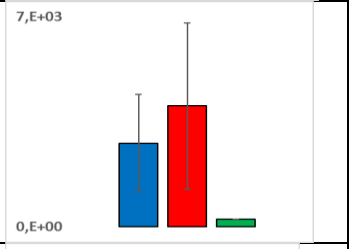
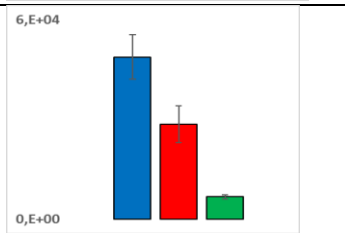
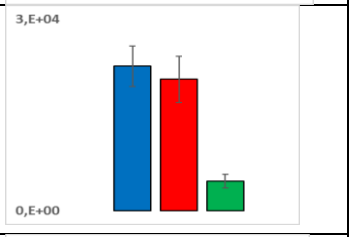
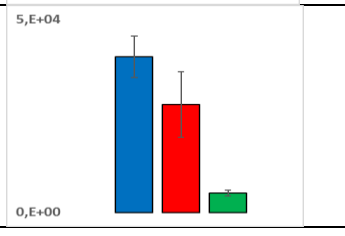
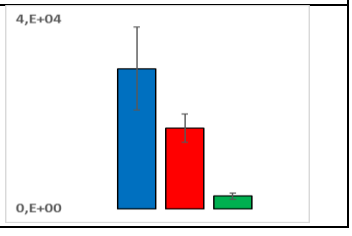
Abbreviations: QC, quality control; RT, retention time; VIP, variable influence on projection

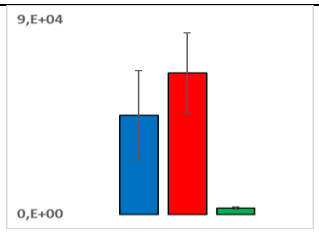
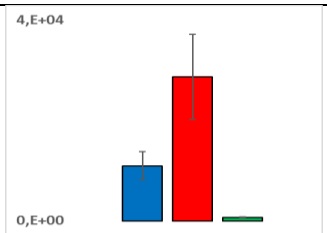
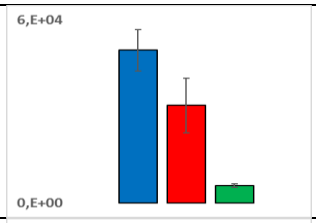
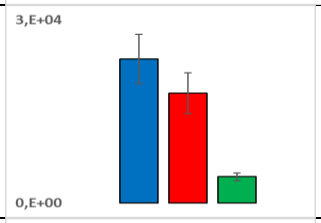
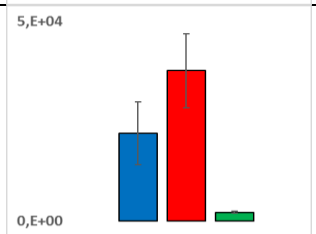
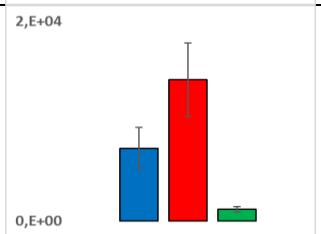
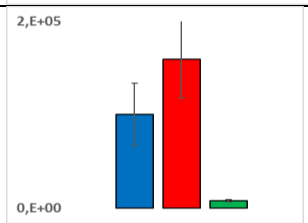
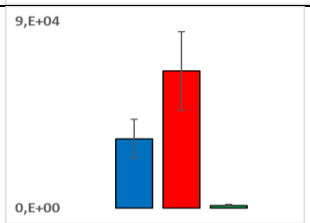
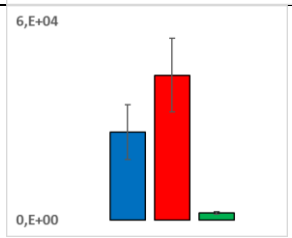
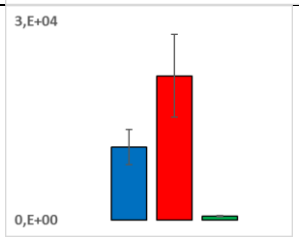
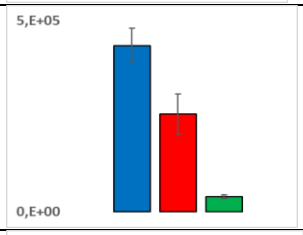
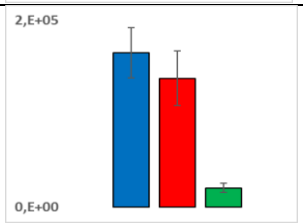
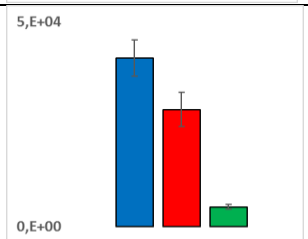
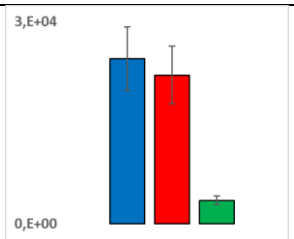
Table 4. "Extra Europe" Markers. VIP values obtained from the OPLS-DA model "Italy" vs "Non Europe"

Name	Pseudomolecular ion	Detected m/z	RT (min)	Predicted Formula	Mass error (ppm)	CV IN QCs (%)	VIP Value	ID type
3-Methylenenorleucine	$[M-H]^+$	144.1018	1.95	C7 H13 N O2	0.69	22	2.05	3
Compound 16	$[M-H]^+$	657.4442	10.26	C34 H56 N8 O5	0.61	11	2.81	4
Compound 17	$[M-H]^-$	291.1468	5.36	C14 H20 N4 O3	2.06	6	2.18	4
Compound 18	$[M+H]^-$	230.1034	5.38	C10 H17 N O5	0.44	9	2.67	4
Vanillylactic acid	$[M+H]^-$	225.0770	6.50	C11 H14 O5	0.89	10	1.68	3
Compound 19	$[M-H]^-$	355.2488	8.14	C20 H36 O5	0.28	11	2.47	4
Compound 20	$[M-H]^-$	353.2333	8.34	C20 H34 O5	0.28	7	2.75	4
Compound 21	$[M-H]^-$	383.2449	8.42	C22 H32 N4 O2	0.78	17	1.78	4
Compound 22	$[M-H]^-$	237.1259	9.36	C10 H23 O4 P	0.84	33	1.69	4

Abbreviations: QC, quality control; RT, retention time; VIP, variable influence on projection

Table 5. "Italian" Markers. VIP values obtained from the OPLS-DA model "Italy" vs "Europe

NAME	2016 CAMPAIGN	2018 CAMPAIGN
Compound 1		
Compound 2		
Compound 3		
Compound 4		
Compound 5		
Compound 6		
Compound 7		

NAME	2016 CAMPAIGN	2018 CAMPAIGN
Compound 8		
Sucrose 3-lauroyl-3'-isovaleric acid		
Compound 9		
Compound 10		
Compound 11		
β-D-Fructofuranosyl 2-O-dodecanoyl-3-O-(3-methyl-2-butenoyl)-α-D-glucopyranoside		
Compound 12		

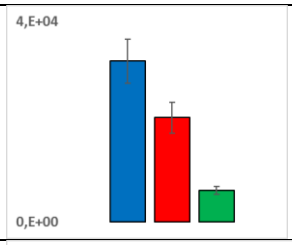
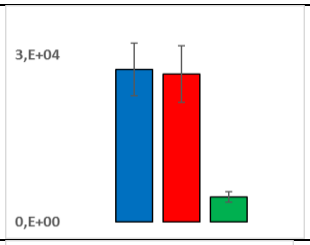
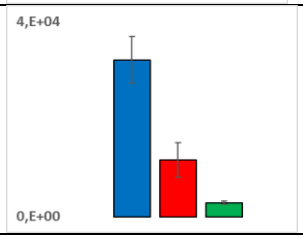
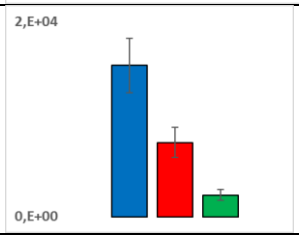
NAME	2016 CAMPAIGN	2018 CAMPAIGN
Compound 13		
(2R,3S)-2,3,4-Trihydroxybutyl 4-O-acetyl-2-O-propionyl-3-O-tetradecanoyl-β-D-mannopyranoside		

Table 6. Comparison of the mean area values (+/− standard errors) of the three geographical areas between the 2016 and the 2018 campaigns for the “Not Italian” markers. Blue bar: European wheat; red bar: Extra-European wheat; green bar: Italian wheat

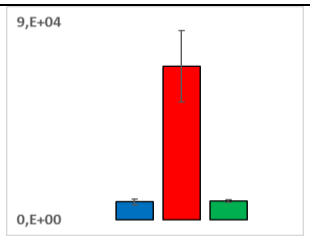
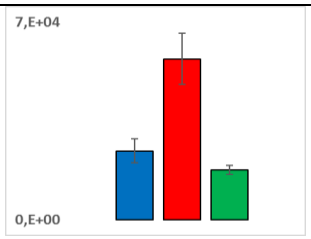
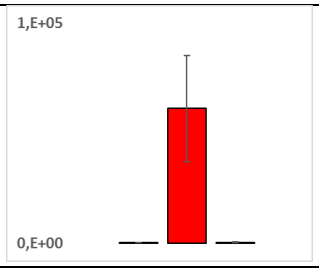
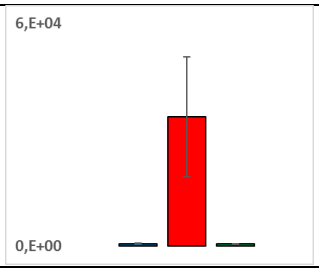
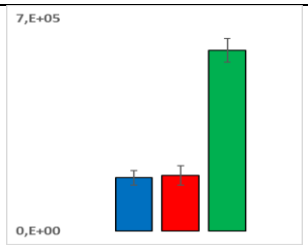
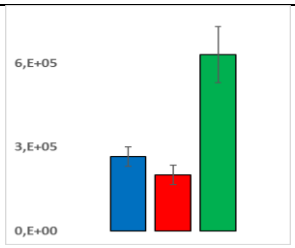
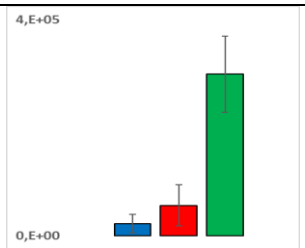
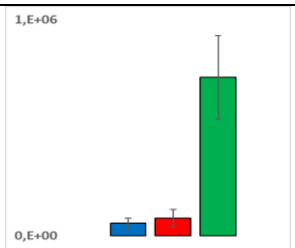
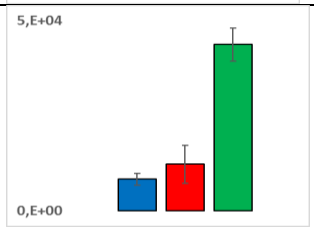
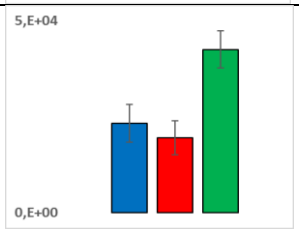
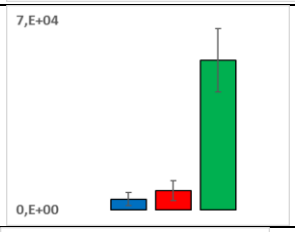
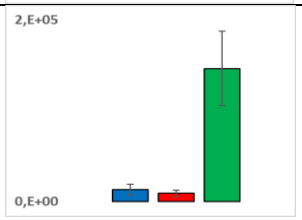
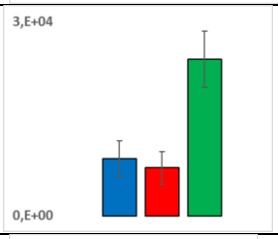
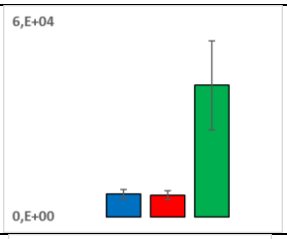
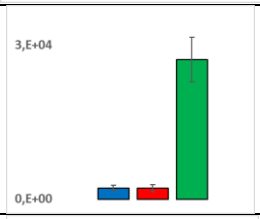
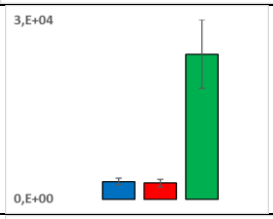
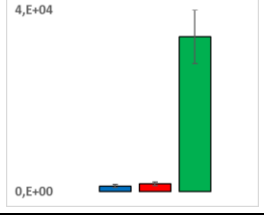
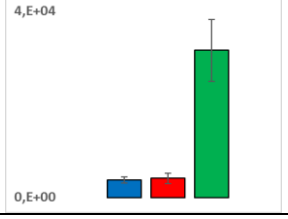
NAME	2016 CAMPAIGN	2018 CAMPAIGN
Compound 14		
Compound 15		

Table 7. Comparison of the mean area values (+/− standard errors) of the three geographical areas between the 2016 and the 2018 campaigns for the “Extra Europe” markers. Blue bar: European wheat; red bar: Extra-European wheat; green bar: Italian wheat

NAME	2016 CAMPAIGN	2018 CAMPAIGN
3-Methylenenorleucine		
Compound 16		
Compound 17		
Compound 18		
Vanillylactic acid		
Compound 19		
Compound 20		

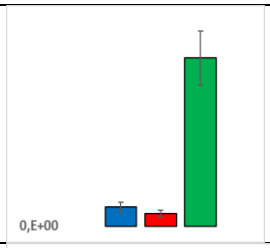
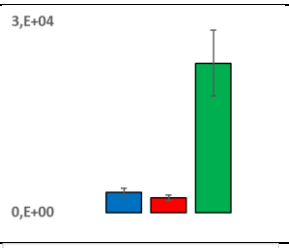
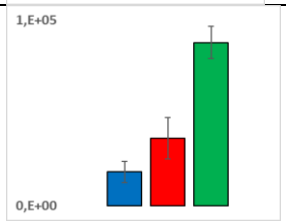
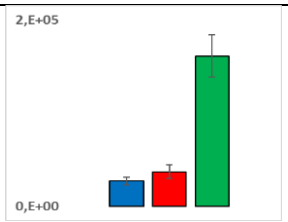
NAME	2016 CAMPAIGN	2018 CAMPAIGN
Compound 21		
Compound 22		

Table 8. Comparison of the mean area values (+/− standard errors) of the three geographical areas between the 2016 and the 2018 campaigns for the "Italian" markers. Blue bar: European wheat; red bar: Extra-European wheat; green bar: Italian wheat

Figures

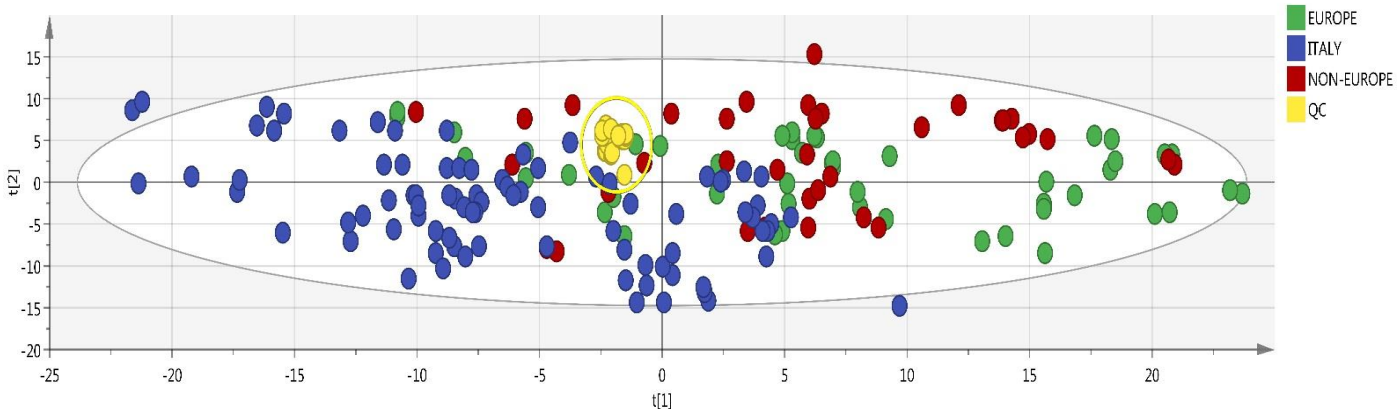


Figure 1. ESI - PCA scores plot of wheat samples. Green dots: European samples; blue dots: Italian samples; red dots: Extra European samples; yellow dots: QC. R^2X (cum): 0.724 Q^2(cum): 0.566. Explained variance of the first two PCs: 31.0%.

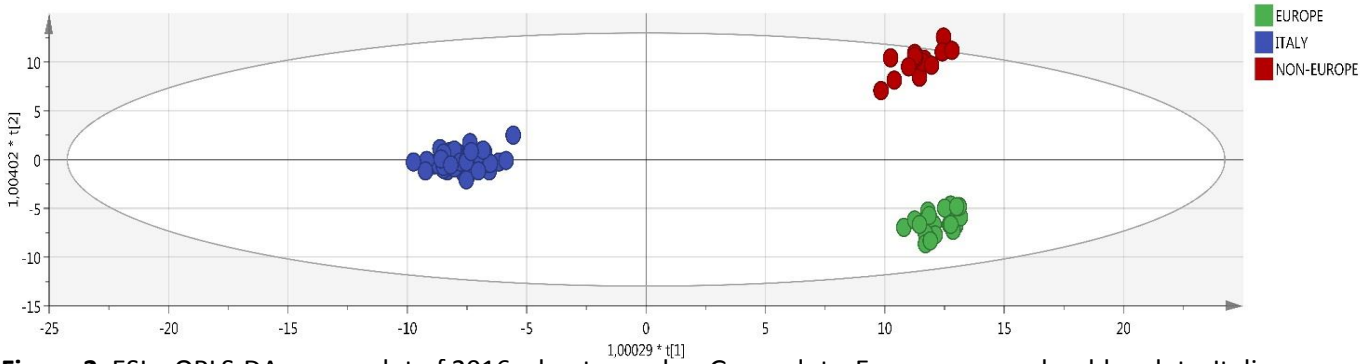


Figure 2. ESI + OPLS-DA scores plot of 2016 wheat samples. Green dots: European samples; blue dots: Italian samples; red dots: Extra European samples;. R^2X (cum): 0.532 R^2Y (cum): 0.981 Q^2(cum): 0.900.

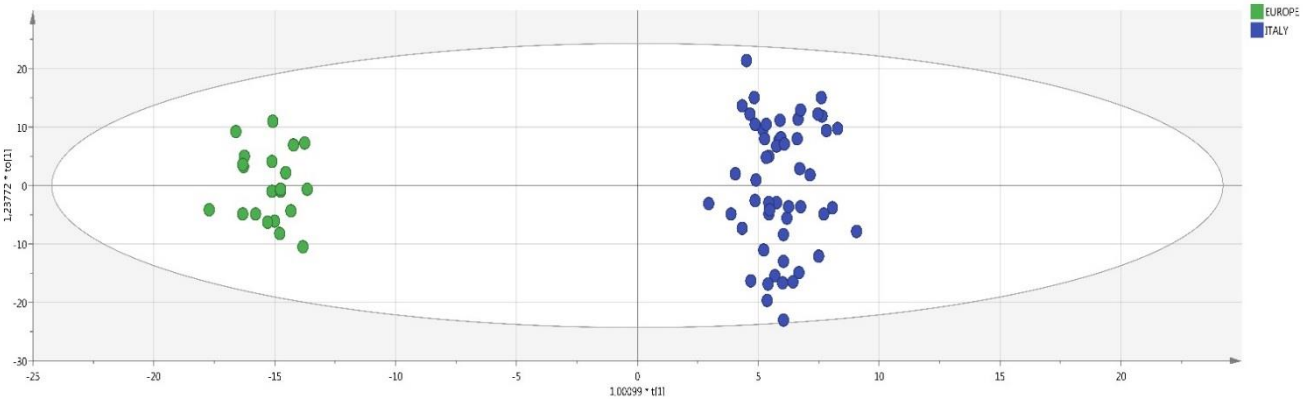


Figure 3. ESI + OPLS-DA scores plot of 2016 "Italian" against "European" samples. Green dots: European samples; blue dots: Italian samples. R^2X (cum): 0.381 R^2Y (cum): 0.985 Q^2(cum): 0.923

Supplementary Material

Section 1 – Wheat samples

The global list of the samples used in this study is detailed in table S1.

SAMPLE NAME	GROUP	NATION	YEAR
EU_1	EUROPE	Greece	2016
EU_2	EUROPE	Greece	2016
EU_3	EUROPE	Greece	2016
EU_4	EUROPE	Greece	2016
EU_5*	EUROPE	Greece	2016
EU_6	EUROPE	Greece	2016
EU_7	EUROPE	Greece	2016
EU_8	EUROPE	Greece	2016
EU_9	EUROPE	Greece	2016
EU_10	EUROPE	Greece	2016
EU_11*	EUROPE	Lithuania	2016
EU_12	EUROPE	Turkey	2016
EU_13	EUROPE	Turkey	2016
EU_14	EUROPE	Turkey	2016
EU_15*	EUROPE	Turkey	2016
EU_16	EUROPE	Turkey	2016
EU_17*	EUROPE	Turkey	2016
EU_18	EUROPE	Turkey	2016
EU_19	EUROPE	Turkey	2016
EU_20	EUROPE	Turkey	2016
EU_21	EUROPE	Turkey	2016
ITA_1	ITALY	Italy_Center	2016
ITA_2*	ITALY	Italy_Center	2016
ITA_3	ITALY	Italy_Center	2016
ITA_4*	ITALY	Italy_Center	2016
ITA_5	ITALY	Italy_Center	2016
ITA_6	ITALY	Italy_Center	2016
ITA_7	ITALY	Italy_Center	2016

ITA_8*	ITALY	Italy_Center	2016
ITA_9*	ITALY	Italy_Center	2016
ITA_10	ITALY	Italy_Center	2016
ITA_11	ITALY	Italy_Center	2016
ITA_12	ITALY	Italy_Center	2016
ITA_13	ITALY	Italy_Center	2016
ITA_14	ITALY	Italy_Center	2016
ITA_15	ITALY	Italy_Center	2016
ITA_16	ITALY	Italy_Center	2016
ITA_17	ITALY	Italy_Center	2016
ITA_18	ITALY	Italy_Center	2016
ITA_19*	ITALY	Italy_North	2016
ITA_20	ITALY	Italy_North	2016
ITA_21	ITALY	Italy_North	2016
ITA_22	ITALY	Italy_North	2016
ITA_23*	ITALY	Italy_North	2016
ITA_24	ITALY	Italy_North	2016
ITA_25	ITALY	Italy_North	2016
ITA_26*	ITALY	Italy_North	2016
ITA_27*	ITALY	Italy_North	2016
ITA_28*	ITALY	Italy_North	2016
ITA_29	ITALY	Italy_North	2016
ITA_30	ITALY	Italy_North	2016
ITA_31	ITALY	Italy_North	2016
ITA_32	ITALY	Italy_South	2016
ITA_33	ITALY	Italy_South	2016
ITA_34*	ITALY	Italy_South	2016
ITA_35*	ITALY	Italy_South	2016
ITA_36	ITALY	Italy_South	2016
ITA_37	ITALY	Italy_South	2016
ITA_38	ITALY	Italy_South	2016
ITA_39	ITALY	Italy_South	2016
ITA_40	ITALY	Italy_South	2016

ITA_41	ITALY	Italy_South	2016
ITA_42	ITALY	Italy_South	2016
ITA_43	ITALY	Italy_South	2016
ITA_44	ITALY	Italy_South	2016
ITA_45	ITALY	Italy_South	2016
ITA_46	ITALY	Italy_South	2016
ITA_47	ITALY	Italy_South	2016
ITA_48	ITALY	Italy_South	2016
ITA_49	ITALY	Italy_South	2016
ITA_50	ITALY	Italy_South	2016
ITA_51	ITALY	Italy_South	2016
ITA_52	ITALY	Italy_South	2016
ITA_53*	ITALY	Italy_South	2016
ITA_54	ITALY	Italy_South	2016
ITA_55	ITALY	Italy_South	2016
ITA_56	ITALY	Italy_South	2016
EXTRA_EU_1*	NON-EUROPE	Canada	2016
EXTRA_EU_2	NON-EUROPE	Kazakhstan	2016
EXTRA_EU_3	NON-EUROPE	Mexico	2016
EXTRA_EU_4*	NON-EUROPE	Mexico	2016
EXTRA_EU_5	NON-EUROPE	Mexico	2016
EXTRA_EU_6	NON-EUROPE	Mexico	2016
EXTRA_EU_7	NON-EUROPE	Mexico	2016
EXTRA_EU_8	NON-EUROPE	Russia	2016
EXTRA_EU_9*	NON-EUROPE	Russia	2016
EXTRA_EU_10*	NON-EUROPE	Russia	2016
EXTRA_EU_11	NON-EUROPE	Russia	2016
EXTRA_EU_12	NON-EUROPE	Russia	2016
EXTRA_EU_13	NON-EUROPE	USA	2016
EXTRA_EU_14	NON-EUROPE	USA	2016
VAL_EU_1	EUROPE	Austria	2018
VAL_EU_2	EUROPE	Austria	2018
VAL_EU_3	EUROPE	France	2018

VAL_EU_4	EUROPE	France	2018
VAL_EU_5*	EUROPE	France	2018
VAL_EU_6	EUROPE	France	2018
VAL_EU_7*	EUROPE	France	2018
VAL_EU_8	EUROPE	France	2018
VAL_EU_9	EUROPE	Greece	2018
VAL_EU_10	EUROPE	Greece	2018
VAL_EU_11	EUROPE	Greece	2018
VAL_EU_12	EUROPE	Greece	2018
VAL_EU_13	EUROPE	Slovakia	2018
VAL_EU_14	EUROPE	Spain	2018
VAL_EU_15	EUROPE	Spain	2018
VAL_EU_16*	EUROPE	Spain	2018
VAL_EU_17*	EUROPE	Spain	2018
VAL_EU_18	EUROPE	Turkey	2018
VAL_EU_19	EUROPE	Turkey	2018
VAL_EU_20	EUROPE	Turkey	2018
VAL_EU_21	EUROPE	Turkey	2018
VAL_ITA_1*	ITALY	Italy_Center	2018
VAL_ITA_2	ITALY	Italy_Center	2018
VAL_ITA_3	ITALY	Italy_Center	2018
VAL_ITA_4	ITALY	Italy_Center	2018
VAL_ITA_5*	ITALY	Italy_North	2018
VAL_ITA_6	ITALY	Italy_North	2018
VAL_ITA_7*	ITALY	Italy_North	2018
VAL_ITA_8	ITALY	Italy_South	2018
VAL_ITA_9	ITALY	Italy_South	2018
VAL_ITA_10	ITALY	Italy_South	2018
VAL_ITA_11*	ITALY	Italy_South	2018
VAL_EXTRA_EU_1	NON-EUROPE	Australia	2018
VAL_EXTRA_EU_2	NON-EUROPE	Australia	2018
VAL_EXTRA_EU_3	NON-EUROPE	Australia	2018
VAL_EXTRA_EU_4	NON-EUROPE	Canada	2018

VAL_EXTRA_EU_5	NON-EUROPE	Canada	2018
VAL_EXTRA_EU_6	NON-EUROPE	Kazakhstan	2018
VAL_EXTRA_EU_7*	NON-EUROPE	Kazakhstan	2018
VAL_EXTRA_EU_8	NON-EUROPE	Russia	2018
VAL_EXTRA_EU_9	NON-EUROPE	Russia	2018
VAL_EXTRA_EU_10	NON-EUROPE	Russia	2018
VAL_EXTRA_EU_11	NON-EUROPE	Russia	2018
VAL_EXTRA_EU_12	NON-EUROPE	Russia	2018
VAL_EXTRA_EU_13	NON-EUROPE	Russia	2018
VAL_EXTRA_EU_14	NON-EUROPE	USA	2018
VAL_EXTRA_EU_15	NON-EUROPE	USA	2018
VAL_EXTRA_EU_16	NON-EUROPE	USA	2018
VAL_EXTRA_EU_17	NON-EUROPE	USA	2018

Table S1: list of wheat samples. *= double extraction

Section 2 – Evaluation of the analytical process

The area values and the retention times shift of Chloramphenicol were important for the assessment of the goodness of the analytical sequences, providing at the same time information concerning the reproducibility of the extraction procedure.

For the sequence analyzed in positive ionization mode, the areas CV was 11.7%, while the CV in the retention time was 2.1%; for the negative acquisition, the values were 11.0% (areas) and 0.3% (retention time) respectively.

According to the results obtained, we can deduce that no issues occurred during sample extraction procedures; moreover, the slight shift in the retention times is a further proof of the goodness of the analytical sequences.

Finally, the results obtained highlight that the response of the instrument was acceptable during the whole sequences (approximately 200 injections each) for both the ionization modes.

Section 3 – Multivariate data elaboration

The PCA (positive mode) and OPLS-DA scores plot (negative mode) of the results obtained are presented in figures S1 and S2

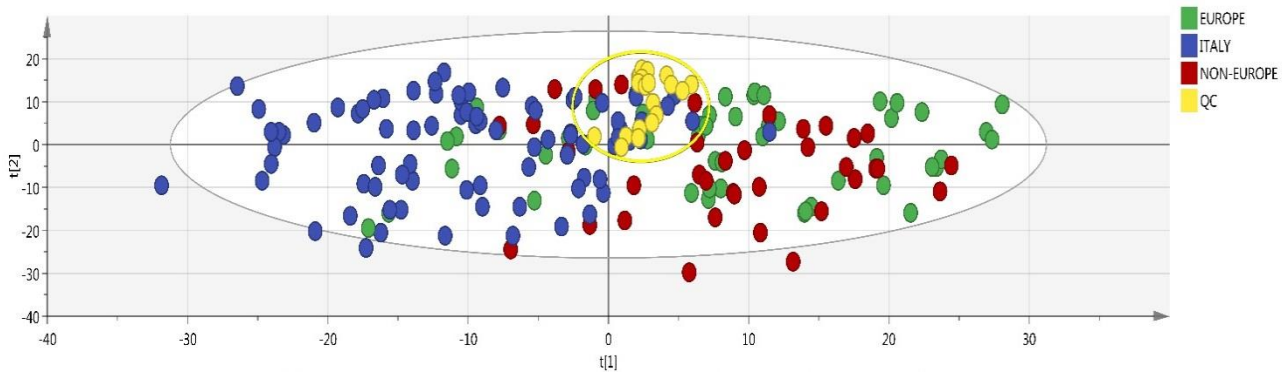


Figure S1. ESI + PCA scores plot of wheat samples. Green dots: European samples; blue dots: Italian samples; red dots: Extra European samples; yellow dots: QC. R^2X (cum): 0.724 Q^2(cum): 0.566. Explained variance of the first two PCs: 23.0%

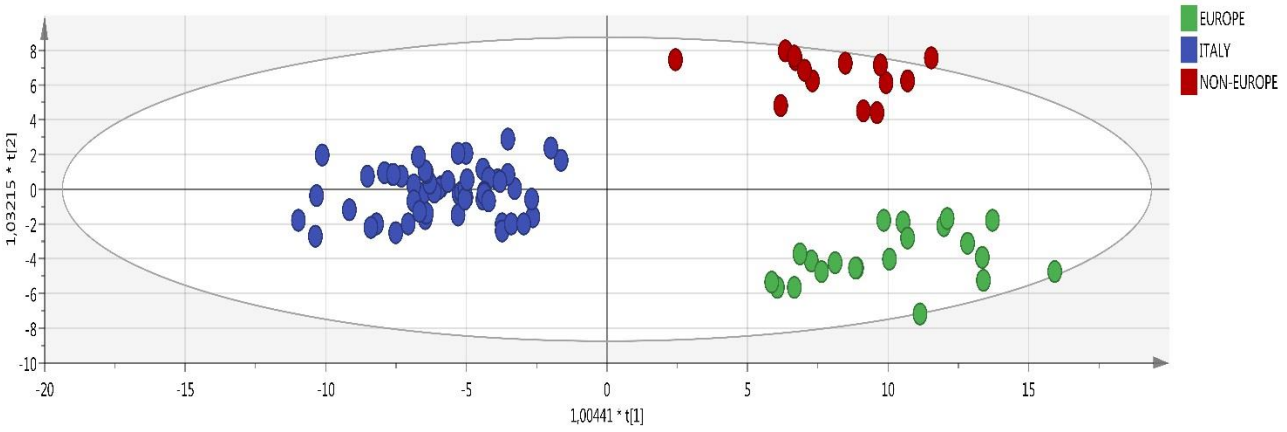


Figure S2. ESI - OPLS-DA scores plot of 2016 wheat samples. Green dots: European samples; blue dots: Italian samples; red dots: Extra European samples;. R^2X (cum): 0.555 R^2Y (cum): 0.883 Q^2(cum): 0.642.

Bibliographic references

Akarachantachote, N., Chadcham, S., & Saithanu, K. (2014). Cutoff Threshold of Variable Importance In Projection For Variable Selection. *International Journal of Pure and Applied Mathematics*, *94*(3), 307-322.

Andersen, M., Kristensen, M., Manach , C., Pujos-Guillot, E., Poulsen, S., Larsen , T., . . . Dragsted, L. (2014). Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics. *Analytical and Bioanalytical Chemistry, 406*(7), 1829-1844.

Black, C., Haughey, S., Chevalier, O., Galvin-King, P., & Elliott, C. (2016). A comprehensive strategy to detect the fraudulent adulteration of herbs: The oregano approach. *Food Chemistry , 210*, 551-557.

Cajka, T., Showalter, M., Riddellova, K., & Fiehn, O. (2016). Advances in Mass Spectrometry for Food Authenticity Testing: An Omics Perspective. In G. Downey, *Advances in Food Authenticity Testing* (pp. 171-200). Oxford: Woodhead Publishing.

Camin, F., Boner, M., Bontempo, L., Fauhl-Hassek, C., Kelly, S., Riedl, J., & Rossmann, A. (2017). Stable isotope techniques for verifying the declared geographical origin of food in legal cases. *Trend in Food Science & Technology, 61*, 176-187.

Castro-Puyana, M., Perez-Míguez, R., Montero, L., & Herrero, M. (2017). Application of mass spectrometry-based metabolomics approaches for food safety, quality and traceability. *Trends in Analytical Chemistry, 93*, 102-118.

Cavanna, D., Catellani, D., Dall'Asta, C., & Suman, M. (2018). Egg product freshness evaluation: a metabolomic approach. *Journal Of Mass Spectrometry, 53(9)*, 849-861.

Cavanna, D., Righetti, L., Elliott, C., & Suman, M. (2018). The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach. *Trends in Food Science & Technology, 80*, 223-241.

Creydt, M., & Fischer, M. (2018). Omics approaches for food authentication. *Electrophoresis, 39*(13), 1569-1581.

Creydt, M., Hudzik, D., Rurik, M., Kohlbacher, O., & Fischer, M. (2018). Food Authentication: Small-Molecule Profiling as a Tool for the Geographic Discrimination of German White Asparagus. *Journal Of Agricultural and Food Chemistry, 66*, 13328-13339.

Ghisoni, S., Lucini, L., Angilletta, F., Rocchetti, G., Farinelli, D., Tombesi, S., & Trevisan, M. (2019). Discrimination of extra-virgin-olive oils from different cultivars and geographical origins by untargeted metabolomics. *Food Research International, 121*, 746-753.

Gil-Solsona, R., Raro, M., Sales, C., Lacalle, L., Díaz, R., Ibáñez, M., . . . Hernández, F. (2016). Metabolomic approach for Extra virgin olive oil origin discrimination making use of ultra-high performance liquid chromatography – Quadrupole time-of-flight mass spectrometry. *Food Control, 70*, 350-359.

Godzien, J., Alonso-Herranz, V., Barbas, C., & Armitage, E. (2015). Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample. *Metabolomics, 11*(3), 518-528.

Jandric, Z., Islam, M., Singh, D., & Cannavan, A. (2017). Authentication of Indian citrus fruit/fruit juices by untargeted and targeted metabolomics. *Food Control, 72*, 181-188.

Jandrić, Z., Roberts, D., Rathor, M., Abrahim, A., Islam, M., & Cannavan, A. (2014). Assessment of fruit juice authenticity using UPLC-QToF MS: a metabolomics approach. *Food Chemistry, 148*, 7-17.

Liu , H., Wei, Y., Lu, H., Wei, S., Jiang, T., Zhang, Y., . . . Guo, B. (2017). The determination and application of (87) Sr/(86) Sr ratio in verifying geographical origin of wheat. *Journal Of Mass Spectrometry, 52*(4), 248-253.

Liu, H., Wei, Y., Lu, H., Wei, S., Jiang, T., Zhang, Y., & Guo, B. (2016). Combination of the 87Sr/86Sr ratio and light stable isotopic values (d13C, d15N and dD) for identifying the geographical origin of winter wheat in China. *Food Chemistry, 212*, 367-373.

Morin, J., Vermeulen, P., & Maestri, E. (2018). Cereals and cereal-based products. In J. Morin, & M. Lees, *FoodIntegrity Handbook. A guide to food authenticity issues and analytical solutions* (pp. 101-114). Nantes: Eurofins Analytics France.

Oldiges, M., Lütz, S., Pflug, S., Schroer, K., Stein, N., & Wiendahl, C. (2007). Metabolomics: current state and evolving methodologies and tools. *Applied Microbiology and Biotechnology, 76*(3), 495-511.

Pauly, A., Pareyt, B., Fierens, E., & Delcour, J. (2013). Wheat (Triticum aestivum L. and T. turgidum L. ssp. durum) Kernel Hardness: II. Implications for End- Product Quality and Role of Puroindolines Therein. *Comprehensive Reviews in Food Science and Food Safety, 12*(4), 427-438.

Popping, B., De Dominicis, E., Dante, M., & Nocetti, M. (2017). Identification of the Geographic Origin of Parmigiano Reggiano (P.D.O.) Cheeses Deploying Non-Targeted Mass Spectrometry and Chemometrics. *Foods, 6*(13).

Quílez, J., Ruiz, J., & Romero, M. (2006). Relationships Between Sensory Flavor Evaluation and Volatile and Nonvolatile Compounds in Commercial Wheat Bread Type Baguette. *Journal of Food Science, 71*(6), S423-S427.

Rashmi, D., Shree, P., & Singh, D. (2017). Stable isotope ratio analysis in determining the geographical traceability of Indian wheat. *Food Control, 79*, 169-176.

Righetti, L., Rubert, J., Galaverna, G., Folloni, S., Ranieri, R., Stranska-Zachariasova, M., . . . Dall'Asta, C. (2016). Characterization and Discrimination of Ancient Grains: A Metabolomics Approach. *International Journal of Molecular Sciences, 17*(1217), 1-14.

Righetti, L., Rubert, J., Galaverna, G., Hurkova, K., Dall'Asta, C., Hajslova, J., & Stranska-Zachariasova, M. (2018). A novel approach based on untargeted lipidomics reveals differences in the lipid pattern among durum and common wheat. *Food Chemistry, 240*, 775-783.

Saia, S., Fragasso, M., De Vita, P., & Beleggia, R. (2019). Metabolomics provide valuable insight for the study of durum wheat: a review. *Journal of Agricultural and Food Chemistry, 67*(11), 3069-3085.

Schymanski, E., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology, 48*, 2097-2098.

Sumner, L., Amberg, A., Barrett, D., Beale, M., Beger, R., Daykin, C., . . . Viant, M. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics, 3*, 211–221. doi:10.1007/s11306-007-0082-2

(2016). *U.S. Pharmacopeial Convention. Guidance on developing and validating nontargeted methods for adulteration detection.*

Wadood , S., Boli, G., & Yimin, W. (2019). Geographical traceability of wheat and its products using multielement light stable isotopes coupled with chemometrics. *Journal of Mass Spectrometry, 54*(2), 178-188.

Wadood , S., Guo, B., Zhang, X., & Wei, Y. (2019). Geographical origin discrimination of wheat kernel and white flour using near- infrared reflectance spectroscopy fingerprinting coupled with chemometrics. *International Journal of Food Science & Technology*.

Want, E., Masson, P., Michopoulos, F., Wilson, I., Theodoridis, G., Plumb, R., . . . Nicholson, J. (2013). Global metabolic profiling of animal and human tissues via UPLC-MS. *Nature Protocols, 8*(1), 17-32.

Wielogorska, E., Chevallier, O., Black, C., Galvin-King, P., Delêtre, M., Kelleher, C., . . . Elliott, C. (2018). Development of a comprehensive analytical platform for the detection and quantitation of food fraud using a biomarker approach. The oregano adulteration case study. *Food Chemistry, 238*, 32-39.

Zhao, H., Guo, B., Wei, Y., Zhang, B., Sun, S., Zhang, L., & Yan, J. (2011). Determining the Geographic Origin of Wheat Using Multielement Analysis and Multivariate Statistics. *Journal Of Agricultural and Food Chemistry, 59*, 4397-4402.

Durum Wheat: industrial benefits and future perspectives

At the moment, the assessment of wheat geographical origin is executed with paper trail documentation and performing internal and external audits directly on the fields.

On the contrary, an objective method able to evaluate this attribute is not available in the company.

For this reason, the research presented in this chapter represents the first stage of a process that will lead Barilla G.& R Fratelli S.p.A. to introduce an objective analytical criterion able to assess the geographical origin of durum wheat.

Further studies will be executed for markers identification and, subsequently, for the development of a quantitative targeted method able to set the concentration limits for each molecule.

Subsequently, binary or ternary mixtures of wheat samples coming from different geographies will be analyzed, in order to check the adulteration limits that can be detected.

After the completion of this wheat study, the same method will be applied also for the analysis of semolina and of pasta products, in order to check the reliability of these molecules as geographical markers along the production chain.

CHAPTER 4

SENSOMICS

Sensomics

General overview

Although aided by visual inspection, the final recognition and quality evaluation of food relies on chemoreceptive events in our nose and oral space, mediated by volatile odor-active and non-volatile taste-active molecules.

The possibility to find a correlation between consumers' judgments and chemical compounds should help in the characterization and formulation of food products, with the aim of identifying criteria to objectively characterize them.

The science that studies these kind of tasks is named "Sensomics" and was firstly defined by Hofmann's lab at Technische Universität München, Germany (http://www.molekulare-sensorik.de/index.php?id=58&L=1).

This approach combines analytical results with human evaluations and judgments, using bioinformatics tools to merge the data; the analytical techniques and the statistical approaches that are used are similar to the ones that were presented in the previous chapters.

For a food company, the identification of active molecules responsible of positive or negative evaluations provided by the consumers represents a crucial tool that could help in the identification of issues during the production chain or, in general, in finding objective criteria that could improve the values of the final products.

Aim of the work

On the basis of what presented above, in the last chapter of this thesis, a characterization of the molecules responsible of consumer overall liking in pasta is presented. This chapter was accepted for publication in "*Journal of AOAC international"*. For additional details see section "Author".

LC-HRMS for characterizing durum wheat pasta production variability and consumers overall liking

Francesca Lambertini[1], Daniele Cavanna[1,2], Dante Catellani[1], Mario Li Vigni[3], Caterina Durante[3], Alessandro D'Alessandro[1], Michele Suman[1]

[1]Barilla Advanced Laboratory Research, Via Mantova 166 - 43122 Parma, Italy
[2]Department of Food and Drug, University of Parma, Parco Area delle Scienze 95/A- 43124 Parma, Italy
[3]ChemStamp srl, via Campi 103- 41125 Modena, Italy

Introduction

Pasta is one of the most common staple food products in the world and is usually produced using durum wheat. Italy is the leader in pasta production; however, in recent decades, its diffusion all over the world has begun to grow tremendously. There are several pasta producers on the market, and one of the most important features that can increase the market share of a specific brand relative to others is clearly the quality of the final product.

The flavor of a food product is obviously a crucial part of its quality; usually, consumer choices and loyalty to a particular product over another is primarily driven by the subjective perception of consumers. Therefore, the production of high quality food commodities requires that the final products always maintain the same quality and taste expected by consumers.

This issue is of particular importance at the industrial level, where production is often managed over different countries and different plants using different raw materials.

Historically, the analytical approach for assessing the flavor of a food product is a panel test; however, even if the opinions provided by the panelists are an important parameter for product quality assessment, the subjectivity of the test always remains an issue. In addition, the high cost of the analysis and the continuous need for trained panelists cannot be ignored.

Promising methods using a sensomics perspective have recently been proposed to overcome these drawbacks, with the aim of identifying criteria to objectively characterize the food products.

The so-called sensomics methodology was defined by Hofmann's lab at Technische Universität München, Germany, and was developed by combining state of the art natural product analytics, human psychophysical techniques, and bioinformatics tools (http://www.molekulare-sensorik.de/index.php?id=2&L=1). In our context, this approach requires the untargeted analysis of

food extracts as a starting point for the identification of specific marker compounds, which are related to the target aspect of the product being assessed.

In recent years, many studies have been conducted to correlate the sensory properties of food with different analytical measurements, and various statistical methods have been used to obtain the most reliable results (Seisonen, Vene, & Koppel, 2016).

Several studies have been published in which some flavor marker compounds have been identified using gas chromatographic techniques coupled with mass spectrometry detection. For example, this technique is successfully exploited for the characterization of key volatile aroma compounds in rum (Franitza, Granvogl, & Schieberle, 2016), native cold-pressed rapeseed oil (Matheis & Granvogl, 2016), Chinese Vidal ice-wine (Ma, Tang, Xu, & Li, 2017) and hazelnuts (Kiefl , Pollner, & Schieberle, 2013).

This approach, however, is useful for the detection of volatile and thermally stable compounds. In contrast, a more versatile analytical approach that can be applied to broader categories of compounds is liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS). In recent years, LC-HRMS has opened new doors for the feasibility of finding relevant chemical information linked to sensory evaluations. With a liquid chromatography approach, easier sample preparation can be carried out, no derivatizations are required, a high number of compounds can be potentially detected simultaneously, and the analyses are usually fast, with great peak shape and resolution.

Moreover, the use of HRMS, which allows the detection of the accurate masses, provides high sensitivity and high mass spectra resolution in compound detection (Rubert, Zachariasova, & Hajslova, 2015), providing a robust method for identifying molecules. Different applications of LC-MS techniques for the characterization of flavor molecules have been reported in the literature; for example, gel permeation chromatography combined with a LC-MS/MS approach is able to identify the molecules responsible for the kokumi sensation in matured Gouda cheese (Toelstede, Dunkel, & Hofmann, 2009). In a more recent work, cation-exchange chromatography, followed by MS/MS quantitation, has been used to identify the flavor profile compounds in the cooked meat of king prawns (Meyer, Dunkel, & Hofmann, 2016).

Furthermore, an UV-Vis, LC-TOF-MS approach in combination with NMR has been presented in the literature as a robust way to identify the bitter compounds in the hard resin of hops used to produce Pilsner-type beers (Dresel, Dunkel, & Hofmann, 2015). Finally, chromatographic fingerprinting seems to be a very promising and innovative approach for food characterization (Cuadros-Rodr´ıguez, Ruiz-Sambl´as, Valverde-Som, Perez-Casta~no, & Gonzalez-Casado, 2016). Frequently, multivariate statistical analyses are required for data management since the amount of raw data is

massive and the information originates from different sources (Seisonen, Vene, & Koppel, 2016). The use of data fusion strategies - in particular, mid-level strategies - to integrate data from different platforms has been demonstrated to be a useful tool for the correct classification of different food samples (Biancolillo, Bucci, Magrì, Magrì, & Marini, 2014).

In the present work, an LC-HRMS method coupled with chemometric data elaboration is proposed as a useful tool for high quality pasta fingerprinting, demonstrating its potential in the identification of marker compounds directly or inversely related to overall consumer liking.

The use of data fusion strategies - in particular, mid-level strategies - to integrate the data from the different platforms allowed the correct classification of all the experimental and validation samples.

Materials and Methods

Chemicals

All solvents used were of HPLC-grade and were purchased from VWR International, Ltd. (Poole, England). Deionized water was prepared using a Milli-Q system (Millipore, Bedford, MA, USA), and analytical standards of sodium bicarbonate, dibasic potassium phosphate, potassium dihydrogen phosphate and sodium chloride were obtained from Sigma-Aldrich (St. Louis, MO, USA). α-Amylase from *Bacillus* sp. type II-A, lyophilized powder was obtained from Sigma-Aldrich.

Sample description

Durum wheat pasta samples in the shape of short wide tubes (penne) were appropriately industrially produced with different recipes in different plants as described in *Table 1*.

Sample preparation

Before the instrumental analysis, durum wheat pasta was cooked and partially digested to reproduce, as much as possible, the stage at which the consumer can evaluate his overall liking of the taste.

The process of cooking pasta at home was reproduced in the laboratory, taking in to account the standard ratio of pasta:water:sodium chloride of 80 g:1 L:7 g (AACC International, 1999). Water was brought to a boil, and then, sodium chloride and pasta were added. The cooking time reported on the label, 11 min, was respected. After that time, the cooking process was stopped by the addition of 200 ml of cold water. Cooked pasta was then poured through a strainer, and the excess water was shaken out. The cooked pasta was left to rest on filter paper for 5 min to allow the evaporation of residual water.

A simplified version of artificial saliva was prepared primarily following the procedure reported by Deibler et al (Deibler, Lavin, Linforth, Taylor, & Acree, 2001) and was composed of sodium bicarbonate (20 mmol/L), dibasic potassium phosphate (2.75 mmol/L), potassium dihydrogen phosphate (12.2 mmol/L), sodium chloride (15 mmol/L), and α-amylase from *Bacillus* sp. (266 U/ml). A 5 gram sample of cooked pasta was minced, wet with 3.5 ml of artificial saliva (without enzymes), vigorously shaken by vortex for 1 min, and then kept in a thermostatic bath at 37 °C. A 0.5 ml sample of artificial saliva was then added, and the mixture was incubated for 20 seconds. The enzymatic reaction was stopped by the addition of 1 ml of acetonitrile (to maintain the 80:20 ratio between the aqueous and organic phases). The solution was centrifuged at 10,000 *rpm* at 4 °C for 10 min, and the supernatant was withdrawn and transferred into a vial for the LC-HRMS analysis.

Technological/chemical evaluations

Chemical evaluations of water sorption and ashes on pasta were executed following the respective AACC International methods (AACCI Method 54-50.01 and AACCI Method 08-12.01).

Pasta colour was evaluated using the CIELab color space measured by a Minolta CR-410 colorimeter (Konica Minolta, Inc.) on milled dry pasta.

Technological texture (Z3) was determined on cooked pasta using an internal method based on the measure of the extrusion energy in an Ottawa cell.

Consumer Testing/Sensory Analysis

Hundreds of consumers living part in North and part in South Italy were recruited for a blind product test in Central location, assessing the pasta product samples, dressed with olive oil, in randomized order. Respondents completed a questionnaire for each sample, consisting of hedonic and diagnostic questions (e.g. (i) how much they like the appearance, flavour or texture of a product, or like specific sensory characteristics, such as its colour, bitterness, …; (ii) intensity rating questions ask respondents to rate the strength of a sensory attribute, for example, its bitterness, on a scale; and (iii) just about-right questions ask respondents to rate whether the level of a sensory attribute is ‘‘too high’’, ‘‘just right’’ or ‘‘too low’’, etc…).

Liquid chromatography - high resolution mass spectrometry analysis

LC separation was performed using a 100 × 2.1 mm inner diameter, 2.6 μm particle size, Kinetex C18 biphenyl analytical column (Phenomenex, Torrance, CA, USA) thermostated at 30 °C using a Dionex UltiMate® 3000 Standard LC system (Thermo Fisher Scientific Inc., Waltham, MA, USA). Gradient elution separation was carried out using formic acid (FA) as the mobile phase modifier with

a flow rate of 500 µl min^{-1} (solvent A: 0.1% FA in water; solvent B: 0.1% FA in acetonitrile) under the following conditions: solvent B was initially set at 0% for 2 min, increased linearly to 50% in 12 min, held constant for 2 min, increased to 100% in 0.5 min, and then maintained at 100% for 3.5 min before column re-equilibration (3.5 min). The injection volume was 10 µl.

The mass spectrometer was a Q Exactive benchtop hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific Inc.) equipped with a heated electrospray ionization interface (H-ESI; Thermo Fisher Scientific Inc.). A total of 4 different MS experiments were carried out: positive and negative ion modes at low and high *m/z*. The sheath gas (nitrogen, 99.9% purity), the auxiliary gas (nitrogen, 99.9% purity), and the sweep gas (nitrogen, 99.9% purity) were delivered at flow rates of 40, 20 and 0 arbitrary units, respectively.

Source conditions, in positive ion mode, were as follows: spray voltage, 3.2 kV; capillary temperature, 220 °C; S-lens RF level, 55 arbitrary units; and aux gas heater temperature, 250 °C.

Full-scan accurate low mass spectra from 50 to 1000 Da and full-scan accurate high mass spectra from 1000 to 2000 Da were obtained at a mass resolution of 70,000 FWHM (*m/z* 200).

Source conditions in negative ion mode were as follows: ESI voltage, 3.0 kV; capillary temperature, 220 °C; S-lens RF level, 55 arbitrary units; and aux gas heater temperature, 250 °C. Full-scan accurate low mass spectra from 50 to 1000 Da and full-scan accurate high mass spectra from 1000 to 2000 Da were obtained at a mass resolution of 70,000 FWHM (*m/z* 200).

Instrumental data were acquired using Xcalibur 3.0 software (Thermo Fisher Scientific Inc.), while processing was performed by Xcalibur 3.0 (qualitative analysis) in combination with SIEVE 2.0 for the untargeted/clustering parts.

The reproducibility of sample analysis was evaluated with a preliminary PCA including all the HRMS variables. The scores plot obtained showed a good overlapping of the replicate samples.

Multivariate data analysis

Multivariate data analysis to extract relevant chemometric information was performed by SIMCA 14 software (Umetrics, MKS Instruments AB, Malmo, Sweden; iPLS Toolbox 8.0 for Matlab).

Besides the main tools for Multivariate Analysis (Li Vigni, Durante, & Cocchi, 2013), i.e. Principal Component Analysis (PCA) for exploratory data evaluation and Partial Least Squares Regression (PLS-R) for the elaboration of quantitative relationships between the X block of analytical and sensory data and the Y variable related to Consumers' Overall Liking , the following tools and approaches have been implemented in this work:

- Interval Partial Least Squares Regression (iPLS-R) (Nørgaard, et al., 2000). iPLS-R is based on the use of PLS-R for the computation of quantitative regression models. Starting from the model with all the descriptors, variables are removed and model performance is compared to evaluate if a statistically significant improvement is obtained. The process is then repeated by adding back variables and evaluating the significance of their contribution to improve the model. This procedure is iterated until model improvement is not found significant anymore on the basis of F-test evaluation. Model validity and robustness is evaluated according to cross-validation: the limited number of samples makes it necessary to use Leave One Out Cross-Validation. Since this CV method is sometimes prone to overfitting, additional parameters were used to evaluate the significance of the regression; in particular, a permutation test was performed. This test consists in the repetition over a number of iteration of the random reordering of the y-block: the model is rebuilt by considering each sample as assigned to a therefore "incorrect" y-value, so that model statistics are an index, when compared to the true model, of how high the "chance correlation" is. Statistical comparison tests compare the results for original model and permuted ones under the null hypothesis that there is no significant difference between the "real" correlation and the "chance" correlation, i.e. the model obtained is as good as predicting random values. The two tests used in this work are the paired t-test, which assumes that data come from normal distributions, and the non-parametric analogous, the Pairwise Wilcoxon signed rank test, which does not make this assumption.

- Data Fusion strategy. In a Data Fusion strategy (Biancolillo, Bucci, Magrì, Magrì, & Marini, 2014), data acquired by means of different techniques are aggregated so that a better understanding and/or prediction ability is obtained by considering the characterization of the system under investigation from different "points of view". When the most intuitive fusion is performed (Low-Level), data are augmented variable-wise: in this case, a proper scaling of the concatenated blocks is necessary in order to avoid that only those with higher variance contribute to the modelling. Mid-Level Data Fusion consists in the independent extraction/compression of each of the data blocks so that, when fused, only relevant and/or compressed variables are used for model evaluation. In this work, a low/mid-level Data Fusion approach was used: data arising from LC-HRMS were first analyzed in order to extract only the features which correlate the most with Consumers' Overall Liking. In this way, the high variance of this data block, which is made both of predictive (informative) and non-predictive

(sample variability not related to consumers' liking) contributions, is filtered and only the most relevant features enter in the final modelling phase. This block of data was directly concatenated to the other set of data (sensory attributes and other chemical, physical and recipe variables) due to the similar number of descriptors in the two blocks after selection. A global scaling to unit variance (autoscaling) showed similar numeric results when compared to a block scaling to unit variance, so the first preprocessing was used for simplicity.

Data set was created using as variables the intensity values belonging to the mean of the two replicate analyses.

Results and Discussion

Sample production was set up with the objective of covering the natural variability in the usual production of semolina pasta. Since one of the largest sources of variability in final short shape pasta is associated with the operating conditions of the drying cycles (Guler, Koksel, & Ng, 2002) (Petitot, et al., 2009) (Piwinska, Wyrwisz, Kurek, & Wierzbicka, 2015), this study used samples produced by different industrial lines. In this context, other sources of variability may be the protein content, length and thickness.

The set of products listed in Table 1 were submitted to consumer testing/sensory analysis and technological/chemical analysis (chemical-physical testing, volatile compound analysis and untargeted LC-HRMS analysis).

Based on consumer opinion, the main reasons for liking semolina pasta are the pleasant flavor, the right texture and the fact that it remains firm after cooking.

Initially, work was carried out to optimize sample treatment conditions, followed by optimization of both the chromatographic separation and mass spectrometry detection. The aim was to ensure that the largest number of compounds could be efficiently detected.

To perform instrumental analysis on a pretreated sample that meets consumer perceptions as much as possible, pasta samples were first cooked in a method that simulates at-home preparation and then submitted to an oral digestion protocol as reported in the experimental section.

To avoid possible systematic variability affecting the experiment, all samples were prepared twice and doubly injected to account for random handling order.

Untargeted detection of a large number of compounds, which all have different chemical properties, by LC-HRMS requires finding a reasonable balance between sample preparation, chromatographic separation, Heated Electrospray Ionization (HESI) efficiency and quality of the mass spectra recorded.

For this reason, the LC-HRMS conditions used for this experiment were deliberately set as general as possible to meaningfully detect the largest number of compounds.

From the chromatographic side, a linear elution gradient from 0 to 50% of organic phases was set such that the largest number of compounds could potentially be separated.

Formic acid was the only mobile phase modifier used without the addition of a buffer, which served two purposes: first, the use of other compounds could increase the intensity of some compounds but could potentially decrease the intensity of others; second, the use of formic acid alone minimized the in-source adduct formation, facilitating compound identification. At the same time, mild operating conditions for the HESI source with respect to the gas flow rates and capillary voltages, among others, were chosen based on our previous studies. Moreover, different detection experiments, such as positive and negative ionization modes and *m/z* acquisition over a wide range (*m/z* from 50 to 2000 Da, recorded in two ranges: from *m/z* 50 to 1000 Da and from *m/z* 1000 to 2000 Da) with or without fragmentation, have been performed. Although most of the compounds can be detected in HESI in the positive ionization mode, parallel analysis in HESI in the negative ionization mode was conducted to provide additional information for the identification step of the compounds that can be detected in both modes.

The raw data obtained in the positive ionization mode were processed with Sieve software. The experimental design was based on small molecule domain, chromatographic signal detection and non-differential single class analysis.

A framing algorithm was used with previous base-peak alignment considering 2.5 min of scan time, a *m/z* tolerance of 5 ppm, a background threshold of $3 \cdot 10^6$ units and a maximum scan number of 200.

The final check of the integrated peaks was carried out manually and peak areas were used for statistical analysis.

The investigated dataset was composed of nine samples of penne, produced on the industrial lines, characterized by LC-HRMS for a total of 52 *m/z* values. A preliminary exploratory analysis was performed by means of principal component analysis (PCA) in order to evaluate the presence of clusters of samples according to different levels of liking by consumers (not included in the dataset but used for data visualization only – see the color scale on the right side of Figure 1a), and to detect the degree of correlation among *m/z* values and its influence on sample differentiation, if present.

Figure 1a and 1b shows score plots (after autoscaling the entire LC-HRMS dataset) and biplot for the first two principal components, respectively. A clear separation of samples according to consumer overall liking is present along with PC2: samples F32_11_D and P06_13_B are located at negative

values of PC2 and correspond to the samples with the lowest overall liking (< 5.7). The highest overall liking values correspond to F32_11_A and C04_11_A, which have positive values of PC2. As far as the other samples are concerned, PC1 shows a differentiation possibly related to the production context: C04 and F32 products have negative PC1 values, whereas P06, P08 and F21 have positive PC1 values. Considering the loadings, which are the single largest *m/z* contributors to PCA, a high degree of correlation is present, and it is possible to identify some *m/z* values that are closer to the samples with the lowest PC2 scores; hence, their signals are more intense. Meanwhile, the *m/z* values that do not appear correlated to PC2 (positive contribution) are less present in the two samples.

To determine if a quantitative relationship between ion profile and overall liking could be established, thus enhancing the possibility of understanding the connection between chemical composition of the product and consumer preferences, a partial least squares regression model was computed using the 9 samples x 52 molecular ions dataset and the overall liking as the response. Due to the number of samples, a leave-one-out cross-validation scheme was used to evaluate model performance. A 2 Latent Variables (LV's) second-level model was obtained, with an acceptable performance in terms of fit (84% explained variance of Y) but poor cross-validation results (25% explained variance of Y and 0.3 RMSECV). Therefore, a variable selection strategy was implemented using interval-PLS regression (iPLS). This gave a significant improvement of model performance both in fit (99% explained variance of Y) and in cross-validation (94% explained variance of Y and 0.1 RMSECV); moreover, it allowed narrowing the number of m/z values to be further investigated to 11 out of 52. Among the selected variables, the two following groups of ions were selected:

a) with a positive correlation with overall liking: 151.0086, 146.1652, 123.0139;

b) with a negative correlation with overall liking: 197.0809, 203.0525, 182.0812, 121.1225, 213.1485, 174.1490, 229.1547, 360.1499.

Ions in each group here above described are ordered in terms of decreasing variable importance in projection (VIP); the two compounds with masses of *m/z* 197.0809 and 151.0086 are the ones with highest VIP score of all samples.

iPLS analysis showed a quantitative relationship between the LC-HRMS fragmentation fingerprint of different products with the overall liking of the same, which, for this technique, is well summarized by considering the eleven *m/z* values that were selected as the most significant for the regression.

This approach makes it possible to obtain a data set from analytical technique that is both significantly less cumbersome than the original (11 variables compared to the original 52 for 9 samples) and purged of the less relevant, or noisy entries which reduces variability. The dataset was therefore combined

with other available information (that was important for predicting consumer liking in a previous work), in a data fusion approach. The information from other techniques included the following:

a) technological/chemical evaluations: texture (Z3), color (L*, b*), ashes, and water sorption;
b) sensory panel characterization: consistency in the mouth, flavor balance, flavor fullness, and hardness;
c) recipe as % of regrind semolina added to the mixer;
d) concentration of chemical compounds quantified by GC-MS: pentanal, 3-methyl-butanal, 2-pentyl-furan, and benzaldehyde

A PLS model was obtained predicting the consumers overall liking from the dataset by mid-level fusing of the aforementioned variables, for a total of 25 descriptors.

The performance of a second-level model was excellent, with a 99% of Y variance explained in the fit and a 92% of Y variance explained in cross-validation, with a RMSECV of 0.1 (leave-one-out method).

Permutation tests were carried out over 250 iteration and statistical tests results showed that the model is significant at the 95% confidence level (p-value for Pairwise Wilcoxon signed rank test for calibration: 0.02, for cross-validation: 0.01; p-value for Randomization t-test for calibration: 0.04, for cross-validation: 0.02).

Figure 2 shows the correlation between the variables considered in the model by presenting the PLS weights and scores in a biplot. The overall liking of the samples increases going from negative to positive values of LV1; therefore, it can be seen that all sensorial attributes (e.g., consistency in mouth, hardness, flavor balance and fullness) increase with a higher overall liking, together with the high abundance of *m/z* values such as 151.0086 and 146.1652, and the low abundance of others such as 197.0809, 182.0812 and 203.0525, which is consistent with the previous conclusions.

Other chemical compounds (from GC-MS analysis) appear to influence the overall liking, such as pentanal (with a positive contribution) and 3-methyl-butanal and 2-pentyl-furan (with negative contributions).

As far as the other descriptors are concerned, the presence of regrind semolina seems to lower the overall liking, while color properties do not appear as important as the other parameters. It is also interesting to notice the degree of direct correlation of a technological parameter such as Z3 with sensorial attributes that are connected to pasta consistency and hardness since Z3 is a mechanical measurement of consistency and hardness due to the structure of the proteins involved.

For the most interesting ions identified by statistical analysis, a systematic workflow for their identification was followed:

1: XIC generation of the target ion and deep study of the compound's mass spectrum (with background subtraction at the beginning and at the end of the peak)

2: corresponding molecular weight assignment after the evaluation of

charge status

isotopic pattern distribution

potential adducts (i.e., Na or K adducts)

3: study of the mass spectra recorded under the target chromatographic peak and evaluation of other *m/z* values that could correspond to neutral losses (i.e., water loss) or to other fragments (i.e., loss of a carboxylic group)

4a: XIC overlay the hypothetical fragments under the target chromatographic peak to confirm that they belong to the target compound after source fragmentation

4b: if data depending scan spectra of target compound is available, the fragmentation pattern is studied

5: generation of the most likely chemical formulas related to the target *m/z* value using Xcalibur software

6: reliability study of the suggested formulas: determine if the number of atoms proposed by the software is plausible and compare of the corresponding theoretical isotopic pattern with the experimental pattern

7: submit the plausible chemical formula and evaluation of the proposed compounds to the Chemspider on-line database

Based on the procedure described above, the compounds corresponding to the ions observed at *m/z* 182.0812 and 203.0525 were identified as tyrosine and glucose, respectively. Their identities were further confirmed by comparison to analytical standards.

According to the previously described statistical approach, both tyrosine and glucose are negatively correlated with the overall liking.

Tyrosine typically has a bitterness taste; therefore, this negative relation should be associated with a taste perception. The scientific literature contains other works in which this amino acid is connected to a bitter taste, for example, in a sensory study of fermented pastes of soybeans and soybean-maze blends (Ng'ong'ola-Manan, Mwangwela, Schuller, Østlie, & Østlie, 2014).

In contrast to this study, from a taste point of view, glucose usually has a positive correlation with overall liking, as seen in different works on tomatoes (Piombino, et al., 2013) and lettuce (Chadwick, Gawthrop, Michelmore, Wagstaff, & Methven, 2016).

A possible explanation is that the negative correlation with glucose obtained here for pasta samples is not related to a taste perception but to unpleasant organoleptic sensations that may be perceived in the mouth during chewing, potentially related to starch content (which is essentially a glucose polymer).

In particular, the smoother the texture, the higher the potential for starch release and therefore the higher the level of glucose.

Conclusions

The production process of pasta can be affected by many different factors on an industrial scale, such as ingredients, recipe, conditions, operation parameters and performance of the production line used. On the other side, a typical target of an industrial food company is to guarantee all consumers around the world the same quality and the same taste/sensorial characteristics of its products.

For this reason, a consumer test study was conducted that was devoted to identifying some potential factors that can be considered responsible for different consumer perceptions of pasta attributes.

Liquid chromatography coupled with high resolution mass spectrometry was exploited to execute a non-targeted analytical exercise on pasta samples that were produced and subjected to consumer test; in this way, the final goal was to potentially identify some molecular markers responsible for consumer perceptions and relate those markers to some influencing factors.

An effective cluster analysis using post-acquisition processing techniques of various groups of samples, characterized by the different combinations of parameter production, was ultimately devised.

Acknowledgement

The authors thank Gilda Re and all the sensory team of Barilla for the support of consumer tests and organoleptic evaluations/data elaboration and thank Daniele Ballestrieri and Elena Bergamini (Barilla Advanced Laboratory Research) for their analytical support and for active collaboration on the project in terms of suggestions and fruitful discussions.

Tables

SAMPLE CODE (DIFFERENT PLANT-LINE COMBINATION)	RECIPE	COOKING TIME
F32_11_A	A	11

F32_13_A	A	13
F21_11_A	A	11
C04_11_A	A	11
P08_11_B	B	11
P06_11_B	B	11
P06_13_B	B	13
P06_11_C	C	11
F32_11_D	D	11

Table 3: Sample description

Figures

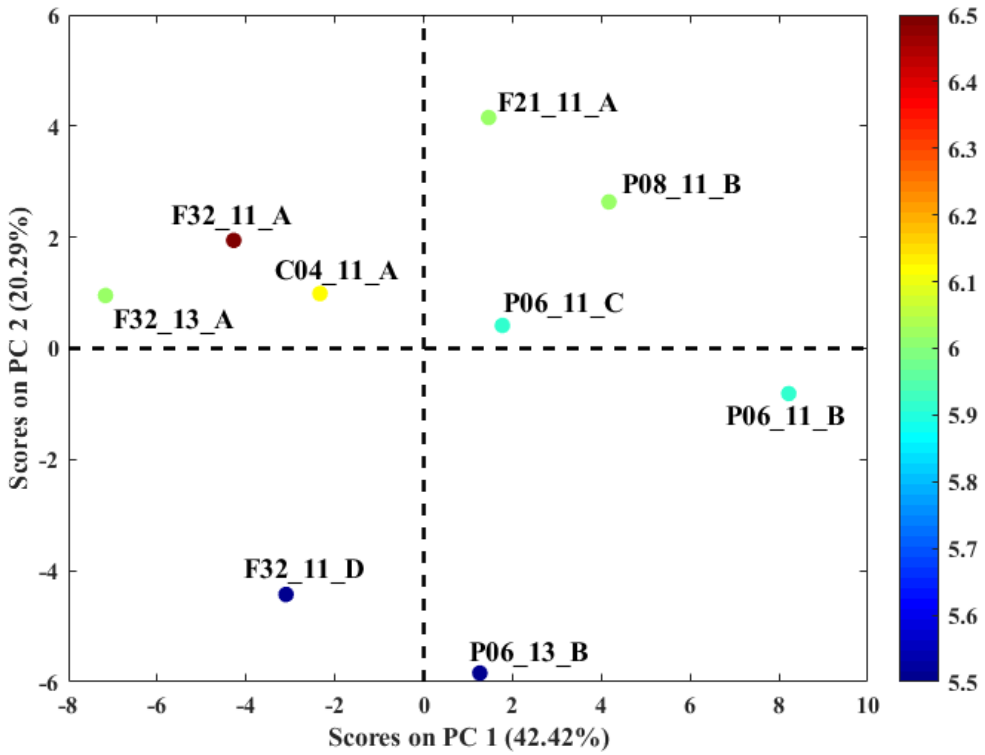


Figure 1a. PCA score plot (PC1 vs PC2) obtained using all 52 m/z values as variables. Samples are colored according to the overall liking values.

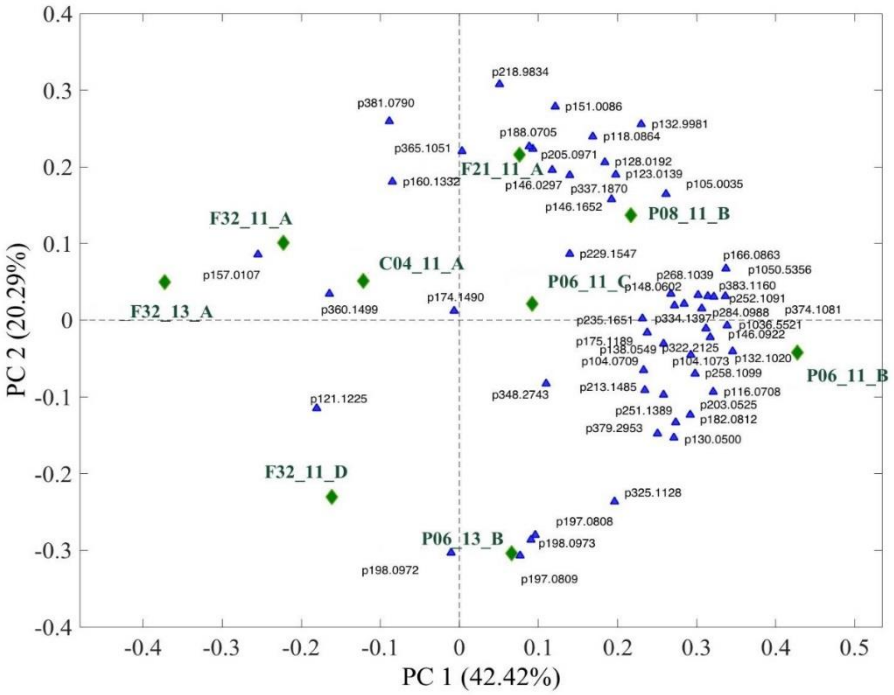


Figure 1b. PCA biplot (PC1 vs PC2): pasta samples are represented by green rhombuses, and the 52 m/z variables are represented by blue triangles.

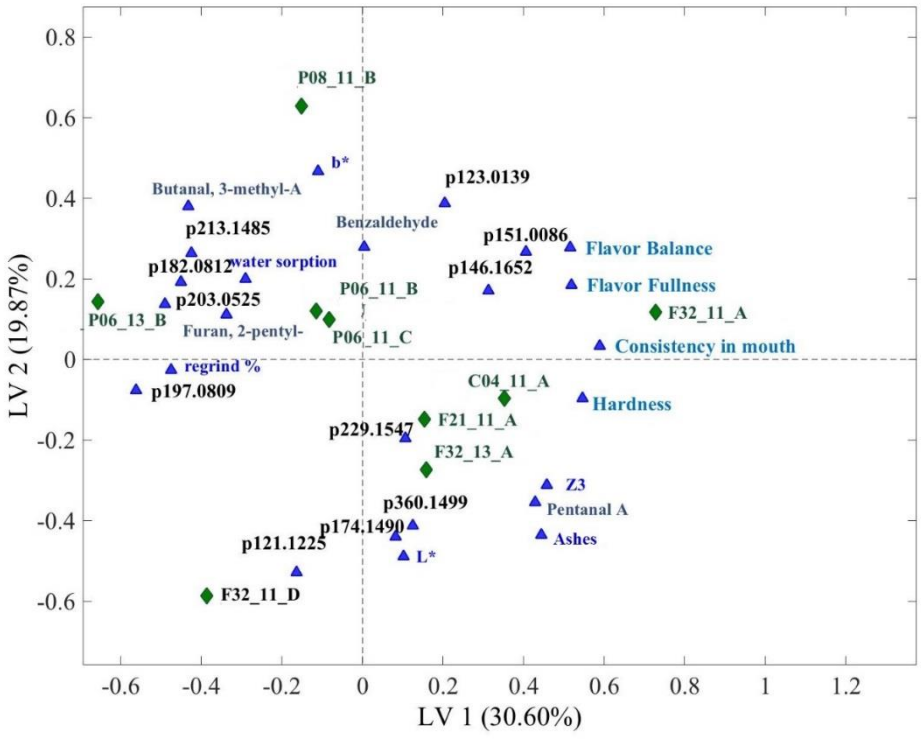


Figure 2. Supervised PLS biplot (LV1 vs LV2) of the samples after the data fusion. Scores are represented by green rhombuses, and variables (the 25 extracted descriptors obtained with the mid-level data fusion approach) are represented by blue triangles. The variables belonging to the same analytical technique are shown in the same color.

Bibliographic references

Biancolillo, A., Bucci, R., Magrì, A., Magrì, A., & Marini, F. (2014). Data-fusion for multiplatform characterization of an italian craft beer aimed at its authentication. *Analytica Chimica Acta, 820*, 23-31.

Chadwick, M., Gawthrop, F., Michelmore, R., Wagstaff, C., & Methven, L. (2016). Perception of bitterness, sweetness and liking of different genotypes of lettuce. *Food Chemsitry, 197*, 66-74.

Cuadros-Rodr´ıguez, L., Ruiz-Sambl´as, C., Valverde-Som, L., Perez-Casta~no, E., & Gonzalez-Casado, A. (2016). Chromatographic fingerprinting: An innovative approach for food 'identitation' and food authentication - A tutorial. *Analytica Chimica Acta, 909*, 9-23.

Deibler, K., Lavin, E., Linforth, R., Taylor, A., & Acree, T. (2001). Verification of a Mouth Simulator by in Vivo Measurements. *Journal of Agricultural and Food Chemistry, 49*(3), 1388-1393.

Dresel, M., Dunkel, A., & Hofmann, T. (2015). Sensomics Analysis of Key Bitter Compounds in the Hard Resin of Hops (Humulus lupulus L.) and Their Contribution to the Bitter Profile of Pilsner-Type Beer. *Journal of Agricultural and Food Chemistry, 63*(13), 3402-3418.

Franitza, L., Granvogl, M., & Schieberle, P. (2016). Characterization of the Key Aroma Compounds in Two Commercial Rums by Means of the Sensomics Approach. *Journal of Agricultural and Food Chemistry, 64*, 637-645.

Guler, S., Koksel, H., & Ng, P. (2002). Effects of industrial pasta drying temperatures on starch properties and pasta quality. *Food Research International, 35*(5), 421-427.

http://www.molekulare-sensorik.de/index.php?id=2&L=1 .

International, A. (1999). AACCI Method 66-50.01–Pasta and Noodle Cooking Quality – Firmness. In *Approved Methods of Analysis,11th Ed.* St. Paul, MN: AACC International.

Kiefl , J., Pollner, G., & Schieberle, P. (2013). Sensomics analysis of key hazelnut odorants (Corylus avellana L. 'Tonda Gentile') using comprehensive two-dimensional gas chromatography in combination with time-of-flight mass spectrometry (GC×GC-TOF-MS). *Journal of Agricultural and Food Chemistry, 61*(22), 5226-5235.

Li Vigni, M., Durante, C., & Cocchi, M. (2013). Exploratory data analysis. In F. Marini, *Data Handling in Science and Technology, Ch. 3, Vol. 28: Chemometrics in Food Chemistry* (p. 55-126). Amsterdam, The Netherlands: Elsevier.

Ma, Y., Tang, K., Xu, Y., & Li, J.-M. (2017). Characterization of the Key Aroma Compounds in Chinese Vidal Icewine by Gas Chromatography–Olfactometry, Quantitative Measurements, Aroma Recombination, and Omission Tests. *Journal of Agricultural and Food Chemistry, 65*, 394-401.

Matheis, K., & Granvogl, M. (2016). Characterisation of the key aroma compounds in commercial native cold-pressed rapeseed oil by means of the Sensomics approach. *European Food Research and Technology, 242*(9), 1565-1575.

Meyer, S., Dunkel, A., & Hofmann, T. (2016). Sensomics-Assisted Elucidation of the Tastant Code of Cooked Crustaceans and Taste Reconstruction Experiments. *Journal of Agricultural and Food Chemistry, 64*(5), 1164-1175.

Ng'ong'ola- Manan, T., Mwangwela, A., Schuller, R., Østlie, H., & Østlie, T. (2014). Sensory evaluation and consumer acceptance of naturally and lactic acid bacteria- fermented pastes of soybeans and soybean–maize blends. *Food Science and Nutritions, 2*(2), 114-131.

Nørgaard, L., Saudland, A., Wagner, A., Nielsen, J., Munck, L., & Engelsen, S. (2000). Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. *Applied Spectroscopy, 54*(3), 413-419.

Petitot, M., Brossard, C., Barron, C., Larre, C., Morel, M., & Micard, V. (2009). Modification of pasta structure induced by high drying temperatures. Effects on the in vitro digestibility of protein and starch fractions and the potential allergenicity of protein hydrolysates. *Food Chemistry, 116*(2), 401-412.

Piombino, P., Sinesio, F., Moneta, E., Cammareri, M., Genovese, A., Lisanti, M., . . . Grandillo, S. (2013). Investigating physicochemical, volatile and sensory parameters playing a positive or a negative role on tomato liking. *Food Research International, 50*(1), 409-419.

Piwinska, M., Wyrwisz, J., Kurek, M., & Wierzbicka, A. (2015). Effect of oat β-glucan fiber powder and vacuum-drying on cooking quality and physical properties of pasta. *CYTA Journal of Food, 14*, 101-108.

Rubert, J., Zachariasova, M., & Hajslova, J. (2015). Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review. *Food Additives & Contaminants: Part A, 32*(10), 1685-1708.

Seisonen, S., Vene, K., & Koppel, K. (2016). The current practice in the application of chemometrics for correlation of sensory and gas chromatographic data. *Food Chemistry, 210*, 530-540.

Toelstede, S., Dunkel, A., & Hofmann, T. (2009). A series of kokumi peptides impart the long-lasting mouthfulness of matured Gouda cheese. *Journal of Agricultural and Food Chemistry, 57*(4), 1440-1448.

Author Contributions

Conceptualization: F.L., A.D., M.S.; study execution: F.L., D.Cat., D.Cav.; data elaboration: M.L.V., C. D.., A.D.; writing-original draft preparation: D.Cav.; F.L.; writing -review and editing: all authors; supervision. A.D., M.

Conclusions and future perspectives

In the present PhD thesis, three specific frauds related to different raw materials have been investigated, exploring rapid and confirmatory approaches.

In the first section, a predictive model able to assess the egg products freshness was created using a GC-IMS technique. This method was developed and tested with samples coming from the production plants with encouraging results; in the future it will be probably used for routine analyses in Barilla's quality control laboratories.

Simultaneously, for confirmatory purposes, new chemical molecules were identified as freshness markers using LC-HRMS techniques: in the future, target methods will be developed and specific concentration limits for each compound will be set up.

In the second section, the same approaches were applied for the detection of soft refined oils addition to extra virgin olive oils.

Two different fast GC instruments were able to similarly detect mixtures until 10% of soft refined oils; in addition, an inter-laboratory non-targeted LC-HRMS study was executed for the selection and identification of specific markers responsible of this kind of fraud. Despite some differences between the two laboratories, the results obtained are comparable and brought to the selection and to a partial identification of new chemical markers that are related to low quality olives and that will be better identified in the future.

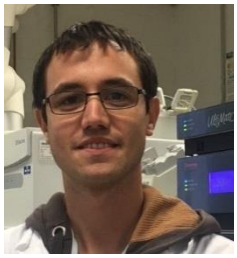
In the third part of this work, the geographical origin of durum wheat was studied with a Liquid Chromatography non-targeted high resolution mass spectrometry approach.

The results obtained brought to the selection and to a partial identification of new molecules characteristic of the Italian and of the not Italian regions; in addition, their values as discriminant markers were confirmed analysing samples coming from a different year of harvesting.

In the last section, the attention was focused on the sensomic field and some molecules responsible of consumers overall liking were identified, showing that the same approaches used for frauds detection could also be applied, with specific adjustments, in the sensory field.

The possibility to objectively predict the consumer's judgments with the measurement of specific parameters is crucial for the food companies, that could improve the quality of the final products more easily.

Author



Daniele Cavanna

Born in Moncalieri (Torino, Italy), on August 18th, 1987.

e-mail: daniele.cavanna@barilla.com

Studies

Daniele graduated summa cum laude in Chemistry in 2011 at "University of Torino".

After a one year period as research assistant at Politecnico of Torino, in 2013 he became a laboratory analyst in "Marini Group". Together with the routine analyses, he was involved in the development of analytical methods for the detection of drugs and illicit additives in foods.

In 2015 he joined "Actavis Pharmaceutical Company" and he worked as laboratory analyst in the "Research and development group".

In 2016 he obtained a second level post graduate master in "Chemical and Toxicological Analyses".

In November 2016, under the supervision of Dr. Michele Suman and Prof.ssa Chiara Dall'Asta, he started his PhD project financed by Barilla and developed together with the University of Parma.

The topic of the research project was the development of innovative "targeted" and "non-targeted" analytical methods that, merged with chemometric data analysis, were able to help in the detection of food frauds. The same approach was also applied in the sensomics field, trying to identify active sensory molecules within different products.

During the PhD project he was visiting student at the "University of Chemistry and Technology" in Prague, at the "Institut des Sciences Analytiques et de PhysicoChimie pour l'Environnement et les Matériaux (IPREM)" in Pau and at the Mathematic department of the "University of York" in York.

Scientific Activities

Papers published

- A. Hydalgo, D. Galbiati, D. Cavanna, M. Suman - **Evaluation of chemical indices for the identification of incubator-reject eggs in egg products** – Food Control, Volume 107 (2020) DOI: 10.1016/j.foodcont.2019.106767
- D. Cavanna, S. Zanardi, C. Dall'Asta, M. Suman - **Ion mobility spectrometry coupled to gas chromatography: A rapid tool to assess eggs freshness**– Food Chemistry, Volume 271 (2019) page 691-696. DOI: 10.1016/j.foodchem.2018.07.204
- D. Cavanna, L. Righetti, C. Elliott, M. Suman - **The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach**– Trends In Food Science &Technology, Volume 80 (2018) page 223-241. DOI: 10.1016/j.tifs.2018.08.007
- D. Cavanna, D. Catellani, C. Dall'Asta, M. Suman – **Egg product freshness evaluation: a**

metabolomic approach – Journal of Mass Spectrometry, Volume 53, Issue 9 (2018), page 849-861. DOI: 10.1002/jms.4256

- L. Righetti, L. Dellafiora, D. Cavanna, E. Rolli, G. Galaverna, R.Bruni, M. Suman, C. Dall'Asta **- Identification of acetylated derivatives of zearalenone as novel plant metabolites by high-resolution mass spectrometry -** Analytical and Bioanalytical Chemistry, Volume 410, Issue 22 (2018), page 5583-5592. DOI: 10.1007/s00216-018-1066-y
- F. Lambertini, D. Cavanna, D. Catellani, A. D'Alessandro, M. Vigni ,C. Durante, M. Suman – **LC-HRMS for characterizing Durum Wheat pasta production variability and consumer overall liking** – Journal of AOAC International Volume 101 Issue 2 (2018) , page 360-366 – DOI:10.5740/jaoacint.17-0209

Book chapters

- M.Suman, D.Cavanna, M.Zerbini, D.Ricchetti, D.Sanfelici ,E.Cavandoli, L.Mirone – **Egg and products** – Food Integrity Handbook – A guide to food authenticity issues and analytical solution, edited by J.F. Morin and M.Lees , page 28-41, Eurofins Analytics France (2018) – ISBN: 978-2-9566303-0-2- DOI:10.32741/fihb.2.eggs

Papers submitted

- D. Cavanna, K. Hurkova, Z. Džuman , A. Serani, M. Serani, C. Dall'Asta, M. Tomaniova, J. Hajslova, M. Suman– **A novel multi-platform High Resolution Mass Spectrometry non-targeted approach facing extra virgin olive oil adulteration with soft refined oils** – Journal of Agricultural and Food Chemistry *(submitted)*
- T. Damiani, D. Cavanna, A. Serani, C. Dall'Asta, M. Suman– **Head Space-based profiling techniques for screening purposes to assess authenticity issues of the EVOO sector** – Food Analytical Methods *(submitted)*
- C. Loffi, D. Cavanna, F. Brigante, A. D'Alessandro, M. Suman - **Non-Targeted LC-HRMS Analysis of Basil and Pesto Sauce -** Journal of the Science of Food and Agriculture (*submitted*)

Manuscript in preparation

- D. Cavanna, C. Loffi, C. Dall'Asta, M. Suman– **A non-targeted high-resolution mass spectrometry approach for the assessment of the geographical origin of durum wheat**– Food Chemistry *(manuscript in preparation)*

Oral communications provided at scientific meetings

- D. Cavanna, C. Dall'Asta, M. Suman - **Validation of non targeted mass spectrometric methods for the detection of food frauds: from theory to real applications–** 2nd Food Chemistry Conference , Seville, 17th-19th September 2019
- D. Cavanna – **Innovative analytical methods merged with chemometrics for food integrity assurance –** XXIV Workshop on the developments in the Italian PhD research on food science, technology and biotechnology , Firenze, 11th-13th September 2019

- D. Cavanna – **Innovative analytical methods merged with chemometrics for food integrity and sensomics –** Fera science seminars (**invited speaker**), York, 6th June 2019
- D.Cavanna, D.Catellani, C.Dall'Asta, L. Righetti, M.Suman – **Non-targeted and metabolomic approaches for food frauds detection**– La Spettrometria di massa per la ricerca scientifica (**invited speaker**), Milano, 28th February 2019
- D. Cavanna, K. Hurkova, A. Serani, C. Dall'asta, J. Hajslova, M. Suman - **A novel multi-platform high resolution mass spectrometry non-targeted approach facing extra virgin olive oil adulteration –** Final Food Integrity Conference, Nantes, 14th-15th November 2018
- D.Cavanna, D.Catellani, C.Dall'Asta, L. Righetti, M.Suman – **Non-targeted and metabolomic approaches for food frauds detection**– Orbitrap MS revolution Day (**invited speaker**), Milano, 13th November 2018
- D.Cavanna, D. Catellani, C. Dall'Asta, M. Suman - **"Non-targeted" analytical methods to detect food frauds: new markers for egg products freshness evaluation –** Asset 2018, Belfast, 28th-31th May 2018
- D.Cavanna, D. Catellani, C. Dall'Asta, M. Suman - **Egg product freshness evaluation: a metabolomic approach -** 5th MS Food Day, Bologna, 11th-13th October 2017

Posters presented at scientific meetings

- D. Cavanna, K. Hurkova, Z. Džuman , A. Serani, C. Dall'Asta, J. Hajslova, M. Suman - **"A Novel Multi-Platform High Resolution Mass Spectrometry Non-Targeted Approach Facing Extra Virgin Olive Oil Adulteration"**, International Mass Spectrometry Conference 2018, Firenze, 26th-31st August 2018
- D.Cavanna, D.Catellani, C.Dall'Asta, M.Suman - **"New Chemical Markers For the Assessment of Egg Products Freshness"**, at Asset 2018, Belfast, 28th-31th May 2018
- D.Cavanna, S.Zanardi, C.Dall'Asta, M.Suman - **"Ion mobility spectrometry: a rapid tool to assess egg products freshness",** 4th Food Integrity Conference, Parma, 10th-11th May 2017