Food Chemistry 373 (2022) 131491

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

FIA–HRMS fingerprinting subjected to chemometrics as a valuable tool to address food classification and authentication: Application to red wine, paprika, and vegetable oil samples

Guillem Campmajó [a,b,*], Javier Saurina [a,b], Oscar Núñez [a,b]

[a] Department of Chemical Engineering and Analytical Chemistry, University of Barcelona, Martí i Franquès 1-11, E08028 Barcelona, Spain
[b] Research Institute in Food Nutrition and Food Safety, University of Barcelona, Av. Prat de la Riba 171, Edifici Recerca (Gaudí), E08921 Santa Coloma de Gramenet, Spain

ARTICLE INFO

Keywords:
FIA–HRMS
Chemometrics
Data fusion
Red wine
Paprika
Olive oil

ABSTRACT

The rise of food fraud practices, affecting a wide variety of goods and their specific characteristics (e.g., quality or geographical origin), demands rapid high-throughput analytical approaches to ensure consumers protection. In this context, this study assesses flow injection analysis coupled to high-resolution mass spectrometry (FIA–HRMS), using a fingerprinting approach and combined with chemometrics, to address four food authentication issues: (i) the geographical origin of three Spanish red wines, (ii) the geographical origin of three European paprikas, (iii) the distinction of olive oil from other vegetable oils and (iv) the assessment of its quality category. In each case, negative and positive ionisation FIA–HRMS fingerprints, and two different data fusion strategies, were evaluated. After external validation, excellent classification accuracies were reached. Moreover, high-resolution mass spectrometry (HRMS) allowed sample matrices characterisation by the putative identification of the most common ions.

1. Introduction

Globalisation has notoriously expanded international trade, increasing the number of participants between the production and consumption in the food chain. In this context, food fraud, which encompasses a sort of intentional manipulation practices in food products (i.e., adulteration, mislabelling, grey market, and counterfeit) aiming an economic gain (Morin & Lees, 2018), has become of great concern among consumers, food businesses, the scientific community, and government administrations. Because of the economic purpose behind food fraud, its likelihood is generally estimated using supply-and-demand and financial indicators influenced by macroeconomic trends and directly affected by unexpected situations such as the Suez Canal blockage or the COVID-19 pandemic (Points, Manning, & Group, 2020). Moreover, goods well-valued for specific labelled particularities (e.g., geographical origin or production system), which enhance their reputation and increase their price, are likely to be affected by fraudulent practices since the difference between authentic and non-authentic products is difficult to measure.

Chromatographic and related techniques —such as capillary electrophoresis (CE), gas chromatography (GC), and liquid chromatography (LC)— with spectroscopic detection or coupled to mass spectrometry (MS), and combined with chemometrics, have proven excellent capacity to address complex food authentication issues through fingerprinting strategies (Cuadros-Rodríguez, Ruiz-Samblás, Valverde-Som, Pérez-Castaño, & González-Casado, 2016; Medina, Perestrelo, Silva, Pereira, & Câmara, 2019). However, the need for more rapid high-throughput analytical approaches, minimising sample analysis time and even costs, has focused the attention on direct MS techniques (Ibáñez, Simó, García-Cañas, Acunha, & Cifuentes, 2015). In this line, both ambient mass spectrometry (AMS) and flow injection analysis coupled to mass spectrometry (FIA–MS) seem to be potential alternatives to non-targeted chromatographic methods.

AMS comprises several techniques, mainly spray- or plasma-based —such as desorption electrospray ionisation (DESI) (Takáts, Wiseman, Gologan, & Cooks, 2004) and direct analysis in real-time (DART) (Cody, Laramée, & Durst, 2005), respectively—, that provide direct desorption/ionisation of analytes from the native sample or with minimal sample

* Corresponding author at: Department of Chemical Engineering and Analytical Chemistry, University of Barcelona, Martí i Franquès 1-11, E08028 Barcelona, Spain.
E-mail address: gcampmajo@ub.edu (G. Campmajó).

https://doi.org/10.1016/j.foodchem.2021.131491
Received 21 June 2021; Received in revised form 26 October 2021; Accepted 26 October 2021
Available online 28 October 2021

treatment. Despite the significant number of advances that have been made in this field in the last decade and its advantages (such as real-time, in situ, and in vivo analysis), AMS techniques still lack three crucial aspects in fingerprinting approaches: good reproducibility, sensitivity, and wide molecular coverage (Kuo, Dutkiewicz, Pei, & Hsu, 2020). Instead, FIA–MS, which is based on injecting a small volume of a liquid sample or a sample extract into an organic phase continuous stream that carries the sample bolus up to the mass spectrometer ion source, provides satisfactory reproducibility because of modern automatic autosamplers and injectors precision and repeatability (Gachumi, Purves, Hopf, & El-Aneed, 2020) and allows better ionisation efficiency through any atmospheric pressure ionisation (API) source —mostly electrospray ionisation (ESI)— than AMS. Therefore, FIA–MS offers a balance between chromatographic-MS and AMS techniques regarding analytical capability and sample analysis throughput (Nanita & Kaldon, 2016).

Several analytical methods aiming to solve food authentication issues by FIA–MS fingerprinting, combined with chemometrics, have been developed in the last years. In this line, nominal mass fingerprints acquired by flow injection analysis coupled to low-resolution mass spectrometry (FIA–LRMS), mainly based on ion trap technology MS instruments —i.e., ion trap (IT) and linear ion trap (LIT)— due to their higher sensitivity in full-scan MS mode than quadrupole mass analysers, have proved their potential in particular applications, such as the organic and conventional sage sample differentiation (Gao, Lu, Sheng, Chen, & Yu, 2013) or cinnamon species classification (Chen, Sun, & Ford, 2014). In contrast, exact mass fingerprints are obtained by flow injection analysis coupled to high-resolution mass spectrometry (FIA–HRMS), using time-of-flight- (TOF) or Orbitrap-based mass analysers, which present a maximum resolving power up to 50,000 and 500,000 full-width at half maximum (FWHM), respectively (Rubert, Zachariasova, & Hajslova, 2015). Thus, FIA–HRMS leads to richer fingerprints, where near-isobaric compound signals are well-resolved, enhancing selectivity and providing better molecular coverage than FIA–LRMS. For instance, it has been already successfully applied to address oregano (Gao et al., 2014) and lettuce (Sun et al., 2018) production system or milk adulteration detection (Du et al., 2018). Moreover, taking advantage of its accurate mass measurements and isotopic abundance ratios, molecular *formulae* of specific ions can be determined and compared to publicly accessible databases for putative compound identification.

This study aimed to demonstrate FIA–HRMS suitability to address certain food authenticity issues through a fingerprinting approach and its combination to principal component analysis (PCA), partial least squares regression-discriminant analysis (PLS-DA), and soft independent modelling of class analogy (SIMCA). Thus, the geographical origin of three Spanish red wines (*Catalunya*, *La Rioja*, and *Castilla y León*) and three European paprikas (*La Vera*, *Murcia*, and the Czech Republic), as well as the distinction of olive oil from other vegetable oils and the assessment of its quality category, were evaluated.

2. Materials and methods

2.1. Reagents and solutions

For the sample treatment, purified water was obtained using an Elix® 3 coupled to a Milli-Q® system (Millipore Corporation, Bedford, MA, USA) and filtered through a 0.22-μm nylon membrane; hexane and formic acid (96%) were provided from Merck (Darmstadt, Germany); UHPLC-supergradient acetonitrile was from Panreac (Castellar del Vallès, Spain); and ethanol was purchased from VWR International Eurolab S. L. (Barcelona, Spain). Instead, for the FIA–HRMS, LC–MS grade water and acetonitrile were from Merck.

2.2. Instrumentation

FIA was performed using an ultra-high-performance liquid chromatography (UHPLC) system equipped with an Accela 1250 quaternary pump and an Accela autosampler (Thermo Fisher Scientific, San Jose, CA, USA). The sample injection volume was 10 μL. The carrier consisted of a 50:50 (*v/v*) mix, composed of water acidified with 0.1% formic acid (*v/v*) and acetonitrile, and was pumped isocratically at 150 $\mu L \cdot min^{-1}$ for 1.5 min.

The UHPLC system was coupled to a hybrid quadrupole-Orbitrap (Q-Orbitrap) mass spectrometer (Q-Exactive Orbitrap, Thermo Fisher Scientific) equipped with a heated electrospray ionisation source (H-ESI II) operating in both negative and positive ionisation modes. The H-ESI source was set in an off-axis position to prevent and minimise mass spectrometer contamination. Nitrogen with a purity of 99.98%, purchased from Linde (Barcelona, Spain), was used for the ESI sheath, sweep, and auxiliary gas at flow rates of 40, 0, and 12 a.u. (arbitrary units), respectively. Moreover, the vaporiser temperature was set at 250 °C, the capillary temperature at 350 °C, the spray voltage at ± 3.0 kV (depending on the ionisation mode), and the S-lens RF level at 50 V. The Q-Orbitrap mass analyser worked in full-scan MS mode, with an *m/z* range from 100 to 1500, at a mass resolution of 70,000 FWHM at *m/z* 200. Besides, an automatic gain control (AGC) target of 3.0×10^6, which is the number of ions to fill the C-Trap, and a maximum injection time of 100 ms, were established. Simultaneously to the full-scan MS mode, data-dependent scan mode ($ddMS^2$) was also performed with an intensity threshold of 1.0×10^5, a fixed first *m/z* of 50 for the registered product ion scan range, a quadrupole isolation window of 0.5 *m/z*, and applying stepped normalised collision energies (NCE) of 17.5, 35.0, and 52.5 eV for ion fragmentation. Besides, in this event acquisition mode, a mass resolution of 17,500 FWHM at *m/z* 200, an AGC target value of 5.0×10^5, and a maximum injection time of 100 ms were also set.

The Q-Orbitrap system was tuned and calibrated every three days, using commercially available calibration solutions for both negative and positive ion modes (Thermo Fisher Scientific). Moreover, the Xcalibur software *v* 4.1 (Thermo Fisher Scientific) was used to control the LC–MS system and acquire and process data.

2.3. Samples and sample treatment

In this study, three different sample sets (red wine, paprika, and olive and other vegetable oils), detailed in the present Section, were under evaluation by the proposed FIA–HRMS method. In all their corresponding sample sequences, in order to ensure the quality of the results avoiding and controlling systematic errors and cross-contamination, a quality control (QC) sample —constructed by pooling equal aliquots of each sample of the set— and an extracting solvent blank were injected at the beginning and after every ten sample injections. Besides, samples were also randomly injected to minimise the effect of instrumental drifts on the chemometric models.

2.3.1. Red wine

A set of 94 red wine samples from three Spanish areas —50 from *Catalunya*, 25 from *La Rioja*, and 19 from *Castilla y León*— encompassing 15 different Protected Designation of Origin (PDO) labels (*Bierzo*, *Catalunya*, *Conca de Barberà*, *Costers del Segre*, *Empordà*, *Montsant*, *Penedès*, *Pla de Bages*, *Priorat*, *Ribera del Duero*, *Rioja*, *Tarragona*, *Terra Alta*, *Tierra de Castilla*, and *Toro*) and ten production years (1996, 2002, 2006, 2007, 2009, 2010, 2011, 2012, 2013, and 2014), and made from various grape varieties, were analysed. Prior to FIA–HRMS analysis, samples were filtered with a 0.22-μm nylon filter (Scharlab, Sentmenat, Spain).

2.3.2. Paprika

One hundred eleven paprika samples, including different geographical origins —72 with *La Vera* PDO (Spain), 24 with *Murcia* PDO (Spain), and 15 from the Czech Republic (their specific region was

not labelled)— and types (hot, bittersweet, and sweet), were directly purchased from their production companies or bought in Spanish or Czech commercial supermarkets.

Regarding the sample treatment, a previously developed procedure was followed (Cetó et al., 2018). Briefly, 0.3 g of the sample were subjected to solid–liquid extraction (SLE) with 3 mL of water:acetonitrile (20:80, *v/v*) mix. After stirring in a Vortex (Stuart, Stone, United Kingdom) for 1 min, sonicating (5510 Branson ultrasonic bath, Hampton, NH, USA) for 15 min, and centrifuging (ROTANTA 460 RS Centrifuge, Hettich, Germany) for 30 min at 4500 rpm, the resulting supernatant extract was filtered, using a 0.22-μm nylon filter and kept at 4 °C in a glass injection vial until its analysis.

2.3.3. Olive and other vegetable oil

In this study, a total of 85 vegetable oil samples —46 olive oils, 15 sunflower oils, 6 corn oils, 6 soy oils, and 12 oils produced from mixtures of seeds (6 sunflower/corn oils and 6 sunflower/soy oils)— from various trademarks and purchased from Barcelona markets were analysed. Moreover, among the olive oil sample set, 12 were refined olive oils (OO), 4 virgin olive oils (VOO), and 30 extra-virgin olive oils (EVOO).

The employed sample treatment was based on a previously described method (Gosetti, Bolfi, Manfredi, Calabrese, & Marengo, 2015) with slight modifications. First, liquid–liquid extraction with low-temperature partition (LLE-LTP), using ethanol:water (70:30, *v/v*) as the extracting solvent, was carried out. Thus, in a 15 mL-polytetrafluoroethylene (PTFE) tube (Serviquimia, Barcelona, Spain), 2.00 g of oil sample were extracted by stirring for 2 min in a Vortex in 2 mL of the extracting solvent. After centrifugation for 5 min at 3500 rpm, the mixture was frozen for 24 h at −18 °C. Then, the resulting supernatant extract was transferred into another PTFE tube for a defatting step with 2 mL of hexane, also by stirring in a Vortex followed by centrifugation for 5 min at 3500 rpm. Finally, the aqueous ethanolic sample extract was filtered with a 0.22-μm nylon filter and stored at −18 °C in a 2-mL glass injection vial until FIA–HRMS analysis.

2.4. Data analysis

2.4.1. Data matrix construction

Raw data results were submitted to exact mass detection, chromatogram builder, isotopic peak grouper, and join aligner, using the mzMine 2.53 software (Pluskal, Castillo, Villar-Briones, & Orešič, 2010). First, the exact mass detection step generated mass lists for each scan acquired in a sample, considering a noise level of 1.0×10^5. Then, the chromatogram builder allowed the joining of exact mass signals found in contiguous scans in a sample, establishing a peak time range of 0.05 – 0.40 min, an *m/z* tolerance of 5 ppm, and an intensity threshold of 1.0×10^5. After this, isotopes were removed, considering that the most representative isotope was the most intense and setting an *m/z* tolerance of 5 ppm. Finally, the join aligner allowed matching of exact masses detected across samples, establishing a mass tolerance of 5 ppm, a peak time tolerance of 0.35 min (the whole time range under evaluation), 95% of weight for *m/z*, and a 0% of weight for time. At the end of this workflow, a data matrix was constructed containing FIA–HRMS fingerprints of the studied samples: samples × variables, where variables consisted of ion signal intensity values as a function of *m/z*. Moreover, to reduce the matrix dimensions, molecular features were filtered and only were selected those with a relative standard deviation (RSD, %) lower than 20% in the signals of the QC samples injected during the sample sequence.

2.4.2. Chemometric analysis

The obtained FIA–HRMS fingerprints were then subjected to PCA, PLS-DA, and SIMCA, which were performed using Solo 8.6 chemometrics software from Eigenvector Research (Manson, WA, USA). Details of the theoretical background of these chemometric methods are addressed elsewhere (Massart et al., 1997; Wold, 1976).

PCA relies on the concentration of the dataset's relevant information, originally arranged in the X-matrix containing sample FIA–HRMS fingerprints, into a reduced number of principal components (PCs). In this study, it allowed an exploratory chemometric analysis to evaluate QC sample behaviour (i.e., QC samples forming a compact group in the PCA scores plot indicates the absence of systematic errors during the sample injection and validates the chemometric results) and sample trends and groups.

Instead, PLS-DA, which uses the same X-matrix as PCA, assigns each given sample into a numerically encoded class in the Y-matrix, depending on predefined sample characteristics (e.g., geographical or botanical origin). In this case, a reduced number of latent variables (LVs) contain the most relevant information that links both matrices. The most appropriate number of LVs to build the PLS-DA models was established at the first significant minimum point of the Venetian blinds cross-validation (CV) error. Besides, considering the complexity of the studied issues, where various sample classes were assessed, the hierarchical model builder (HMB) was used, segregating the complete classification in a consecutive combination of two-input class PLS-DA models (classification decision tree). To evaluate and validate the predictive ability of the whole classificatory chemometric model, 60% of samples were randomly stratified as the calibration set and the remaining 40% as the external validation set. In this line, the performance of the developed classificatory method was checked through each class sensitivity (capability to detect true positives, i.e., samples belonging to a given class that have been correctly assigned) and specificity (capability to detect true negatives, i.e., samples that do not belong to a given class correctly assigned as negative), as well as the overall accuracy of the model (well-classified and misclassified sample ratio).

Finally, SIMCA is based on the definition of a target class by a PCs subspace. In this study, one-class SIMCA was applied to sample authentication, and therefore, each SIMCA model was composed of a unique PCA submodel corresponding to a specific sample class. Then, since it consists of a distance-based method of class modelling, when a new sample is projected into the model, its class membership is assessed according to its distance from the PCA submodel —calculated from the reduced Q residuals and Hotelling T^2 values (normalised to 95% confidence limit) and combined using $d_i = \sqrt{\left(T^2_{red,i}\right)^2 + Q^2_{red,i}}$, being i the index of each given unknown sample to be classified— and a previously established decision threshold. The latter was optimised in each case, maximising the performance of SIMCA in the calibration step by reaching the minimum error. Moreover, same external validation as in the PLS-DA study was carried out.

It should be pointed out that for each sample set, four different X-matrices were used as chemical descriptors in the PLS-DA and SIMCA models: FIA–HRMS fingerprints obtained with negative ionisation, with positive ionisation, and using a low-level (LLDF) or a mid-level data fusion (MLDF) strategy (Borràs et al., 2015). The LLDF X-matrix concatenated both negative and positive ionisation FIA–HRMS data. Instead, the MLDF only contained ten variables per each PLS-DA model involved in the classification decision tree. These variables corresponded to those presenting the highest selectivity ratio among the 50 with the highest variable importance in projection (VIP) values obtained in the LLDF loadings of each PLS-DA model involved in the classification. In all cases, data was autoscaled to provide the same weight to each variable by suppressing differences in their magnitude and amplitude scales.

3. Results and discussion

3.1. FIA–HRMS fingerprint characterisation

Before the chemometric analysis, the obtained FIA–HRMS fingerprints were visually inspected, and some of the most intense ions were

putatively identified to assess sample matrix characterisation, taking advantage of high-resolution mass spectrometry (HRMS) capabilities. In this line, Table S1 summarises the putative identification of some of the most characteristic ions found in the food matrices under study, following Schymanski et al. HRMS identification levels (Schymanski et al., 2014). Public databases, containing MS^2 data —mzCloud (HighChem LLC, Bratislava, Slovakia), Metlin (Smith et al., 2005), and The Human Metabolome Database (Wishart et al., 2018)— and polyphenolic content in food, such as Phenol-Explorer (Rothwell et al., 2013), were consulted. Criteria followed in this process were established as follows: 5 ppm of exact mass tolerance, >90% of isotopic pattern fit, and MS^2 data agreement.

The negative and positive ionisation FIA–HRMS spectra of a *Catalunya*, *La Rioja*, and *Castilla y León* red wine sample are shown in Fig. S1. A priori, no noticeable interregional differences could be highlighted since the most intense ions were commonly found in all samples without following a characteristic pattern due to geographical origin. Among the putatively identified compounds, negative ionisation FIA–HRMS fingerprints contained certain molecules known to be found in wine, such as several organic acids (being tartaric acid the base peak) (Ivanova-Petropulos et al., 2018), hydroxybenzoic (e.g., gallic acid) and hydroxycinnamic acids (e.g., caffeic and caffeoyl tartaric acid) (Gutiérrez-Escobar, Aliaño-González, & Cantos-Villar, 2021), and monosaccharide and sugar-related compounds. Instead, amino acids and choline and furan compounds were found in positive ionisation spectra. Several anthocyanins, which influence the wine colour (Garrido & Borges, 2013), were also found. Moreover, some coumarins, released from wood into the wine during the maturation stage (Hroboňová & Sádecká, 2020), were detected in their $[M + Na]^+$ form.

Regarding paprika samples, as an example, Fig. 1 depicts typical negative and positive ionisation FIA–HRMS fingerprints for hot *La Vera*, *Murcia*, and the Czech Republic samples. At first glance, *La Vera* samples

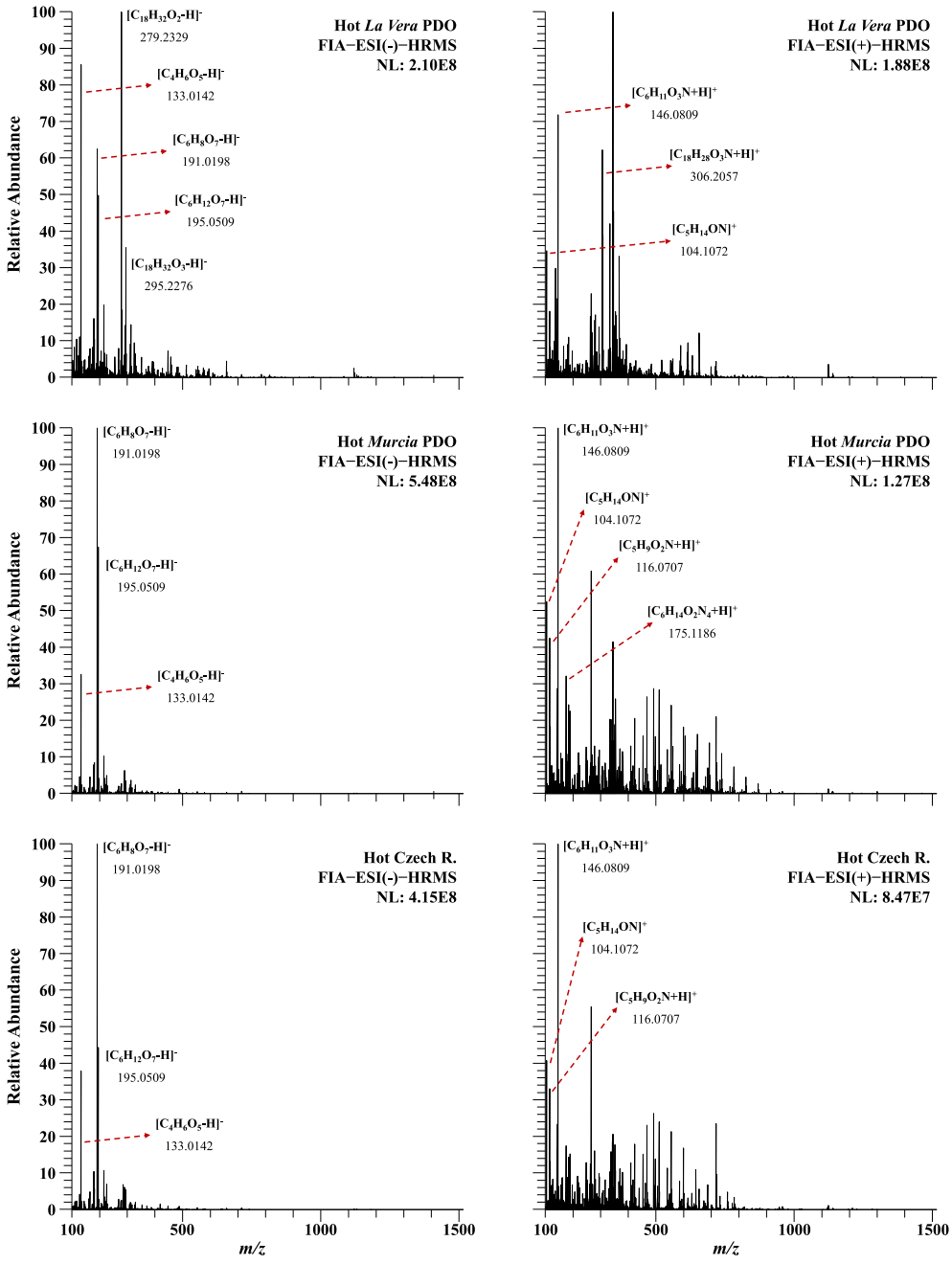

Fig. 1. Negative and positive ionisation FIA–HRMS fingerprints obtained for a *La Vera*, *Murcia*, and the Czech Republic paprika sample.

presented distinctive fingerprints in the negative mode comparing to the remaining samples. For instance, the *m/z* 279.2329, corresponding to the deprotonated molecule of linoleic acid, one of the major fatty acids found in *Jaranda* and *Jariza Capsicum annuum* L. varieties (Pérez-Gálvez, Garrido-Fernández, Mínguez-Mosquera, Lozano-Ruiz, & Montero-de-Espinosa, 1999) that are used in *La Vera* paprika production, was particularly intense in their negative ionisation spectra. In addition, these paprika fingerprints, reproducible among samples belonging to the same geographical origin, contained common compounds such as organic acids or certain mono- and polyunsaturated fatty acids. *La Vera* samples particularity was also highlighted in positive ionisation FIA–HRMS fingerprints that included many signals in the *m/z* range from 100 to 800. However, in this case, spectra were slightly altered in *Murcia* and the Czech Republic samples according to the paprika type, while significantly modified in *La Vera* ones. In this line, compounds such as capsaicinoids may be related to these differences (Arrizabalaga-Larrañaga et al., 2021). Moreover, several amino acids and other compounds, including choline, 6-(hydroxymethyl)pyridin-3-ol, tropine, and 4-hydroxy-1-methyl-2-pyrrolidine carboxylic acid, were also detected in the positive ionisation FIA–HRMS spectra as reported in Table S1.

Finally, both negative and positive ionisation olive oil FIA–HRMS fingerprints were noticeably different from the remaining vegetable oil ones. Instead, as shown in Fig. S2, more similarities were found in olive oil spectra depending on their quality grade. Besides, while VOO and EVOO showed comparable fingerprints among samples belonging to the same group, more variability was observed in OO samples, which may be due to the different percentages of VOO added to them for taste improvement. As shown in Table S1, among other compounds, several polyphenolic compounds well-known to be found in olive oil were identified in the analysed ethanolic sample extract. In this line, tyrosols predominated the negative ionisation spectra, although other polyphenols such as luteolin or dihydro-*p*-coumaric acid were also detected (Farré, Picó, & Barceló, 2019). In the positive ionisation mode, tyrosols were detected, forming an adduct with Na.

3.2. Red wine geographical origin classification and authentication

In this study, FIA–HRMS fingerprints were proposed as chemical markers to address the geographical origin classification of three Spanish red wines: *Catalunya*, *La Rioja*, and *Castilla y León*. Thus, in a first attempt to evaluate their discriminating ability, an exploratory PCA was performed to both negative and positive ionisation data —104 × 440 and 104 × 972 (samples × variables) dimension data matrices, respectively—, aiming to observe QC sample behaviour as well as

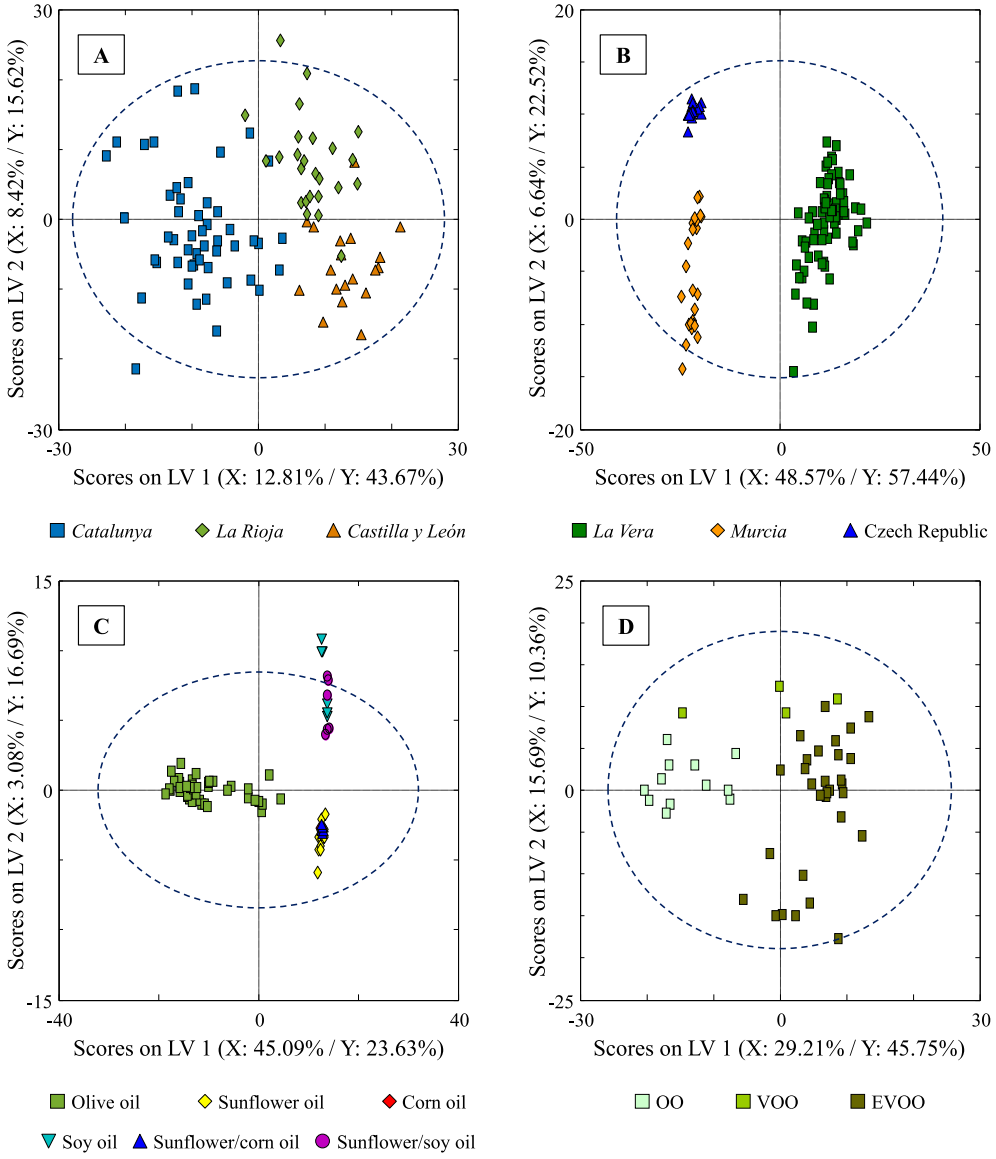

Fig. 2. PLS-DA scores plot of LV1 *vs.* LV2 obtained for: (A) the red wine samples analysed according to their geographical origin, using positive ionisation FIA–HRMS fingerprints; (B) the paprika samples analysed according to their geographical origin, using negative ionisation FIA–HRMS fingerprints; (C) the olive and vegetable oil samples analysed according to their botanical origin, using negative ionisation FIA–HRMS fingerprints; and (D) the olive oil samples analysed according to their quality grade, using negative ionisation FIA–HRMS fingerprints. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sample groups and trends. Similar results were found in both cases. For instance, Fig. S3 shows the PCA scatter plot of scores on the PC2-PC1 (explaining 28.49% of the variance) obtained using positive ionisation data. While QC samples were jointly located in the centre of the plot, indicating the lack of systematic errors, *Catalunya* samples were clearly distinguished at the bottom of the plot displaying negative PC2 values. Instead, *La Rioja* and *Castilla y León* samples shared more similarities since there was no evident discrimination between them (Fig. S3A). In this line, the representation of samples in the PCA scores plot according to non-*Tempranillo* and *Tempranillo*-based red wines (all *La Rioja* and most *Castilla y León* samples were mainly produced from *Tempranillo* grapes, while most *Catalunya* samples did not) allowed to prove its influence on sample distribution. A trend in the PC2 from non-*Tempranillo*, displaying negative values, to *Tempranillo*-based red wines, displaying positive values, was observed (Fig. S3B).

After the exploratory chemometric analysis, QC samples were excluded from X-data matrices (resulting in 94 × 440 and 94 × 972 data matrices for negative and positive ionisation, respectively), which were then subjected to PLS-DA using the corresponding Y-data matrices, indicating sample geographical origin. As expected, the obtained PLS-DA scores plots improved non-supervised chemometric results for both negative and positive ionisation data. For instance, Fig. 2A depicts the scores plot of LV1 *vs.* LV2 obtained when using positive ionisation FIA–HRMS data. In this case, six LVs explaining the 87.98% Y-variance were required to build the PLS-DA model, allowing a good sample distribution according to their geographical origin.

In view of these results, a classification decision tree consisting of two consecutive rule nodes —1) *Catalunya vs.* Others and 2) *La Rioja vs. Castilla y León*— was proposed to address red wine geographical origin classification. As previously mentioned in Section 2.4.2, negative ionisation, positive ionisation, LLDF, and MLDF FIA–HRMS data were tested. In this line, the data matrix dimensions and the number of LVs used in each PLS-DA calibration model involved, as well as the resulting external validation classification parameters (class sensitivity, class specificity, and global accuracy), can be found in Table 1. In this context, LLDF FIA–HRMS fingerprints provided the best external validation classificatory results with 86.8% accuracy. Contrarily, MLDF data, which contained much less sample information, only reached a 60.5% classification rate.

Instead, considering the suitability of class-modelling chemometric methods in the authentication field (Rodionova, Titova, & Pomerantsev, 2016), SIMCA was proposed to test the capacity of FIA–HRMS to generate a characteristic fingerprint for each red wine class. Table 2 summarises the data matrix dimensions, the number of PCs established in each SIMCA model, and the assignation performance after the external validation. Satisfactory overall accuracy results were obtained for the four data matrices used (above 75.4%), although these values were generally slightly below those obtained in the classificatory study with PLS-DA. In fact, only in MLDF, SIMCA provided a better accuracy result than the obtained with the PLS-DA classification decision tree, mainly because of a substantial increase in *La Rioja* sensitivity and *Castilla y León* specificity.

Table 1
Calibration model parameters —data matrix dimensions (samples × variables) and number of LVs— for each of the PLS-DA models built in the classificatory studies and corresponding obtained external validation classification results —class sensitivity (%), class specificity (%), and global accuracy (%)—.

RED WINE GEOGRAPHICAL ORIGIN											
	Calibration: model parameters				**External validation: classification performance**						
	***Catalunya vs.* Others**		***La Rioja vs. Castilla y León***		***Catalunya***		***La Rioja***		***Castilla y León***		**Accuracy**
	Data matrix	**LVs**	**Data matrix**	**LVs**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	
FIA−HRMS (-)	56 × 440	2	26 × 440	4	80.0	94.4	70.0	89.3	87.5	90.0	78.9
FIA−HRMS (+)	56 × 972	5	26 × 972	1	94.4	100.0	90.0	92.9	62.5	100.0	84.2
LLDF	56 × 1412	3	26 × 1412	1	90.0	100.0	100.0	89.3	62.5	100.0	86.8
MLDF	56 × 20	2	26 × 20	3	75.0	100.0	20.0	82.1	75.0	76.7	60.5
PAPRIKA GEOGRAPHICAL ORIGIN											
	Calibration: model parameters				**External validation: classification performance**						
	***La Vera vs.* Others**		***Murcia vs.* The Czech Republic**		***La Vera***		***Murcia***		**The Czech Republic**		**Accuracy**
	Data matrix	**LVs**	**Data matrix**	**LVs**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	
FIA−HRMS (-)	66 × 553	1	23 × 553	2	100.0	100.0	100.0	100.0	100.0	100.0	100.0
FIA−HRMS (+)	66 × 601	2	23 × 601	2	96.6	100.0	100.0	100.0	83.3	100.0	95.6
LLDF	66 × 1154	1	23 × 1154	2	100.0	100.0	100.0	100.0	100.0	100.0	100.0
MLDF	66 × 20	1	23 × 20	1	100.0	100.0	100.0	100.0	100.0	100.0	100.0
OLIVE OIL											
Botanical origin											
	Calibration: model parameters				**External validation: classification performance**						
	Olive oil *vs.* Others				**Olive Oil**						**Accuracy**
	Data matrix	**LVs**			**Sensitivity**		**Specificity**				
FIA−HRMS (-)	53 × 368	1			100.0		100.0				100.0
FIA−HRMS (+)	53 × 739	1			100.0		100.0				100.0
LLDF	53 × 1107	1			100.0		100.0				100.0
MLDF	53 × 10	1			94.4		100.0				96.9
Quality											
	Calibration: model parameters				**External validation: classification performance**						
	EVOO and VOO *vs.* OO				**EVOO and VOO**			**OO**			**Accuracy**
	Data matrix	**LVs**			**Sensitivity**	**Specificity**		**Sensitivity**	**Specificity**		
FIA−HRMS (-)	27 × 368	1			92.9	100.0		100.0	100.0		94.7
FIA−HRMS (+)	27 × 739	3			85.7	100.0		100.0	92.9		89.5
LLDF	27 × 1107	1			92.9	100.0		100.0	92.9		94.7
MLDF	27 × 10	2			92.9	100.0		80.0	92.9		89.5

Table 2
Calibration model parameters —data matrix dimensions (samples × variables) and number of PCs— for each of the SIMCA models built in the class assignation studies and corresponding obtained external validation assignation results —class sensitivity (%), class specificity (%), and global accuracy (%)—.

RED WINE GEOGRAPHICAL ORIGIN											
	Calibration: model parameters				**External validation: assignation performance**						
	Data matrix	**PCs**			***Catalunya***		***La Rioja***		***Castilla y León***		**Accuracy**
		Catalunya	***La Rioja***	***Castilla y León***	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	
FIA–HRMS (-)	56 × 440	6	3	4	65.0	72.2	70.0	89.3	25.0	86.7	75.4
FIA–HRMS (+)	56 × 972	4	3	3	75.0	61.1	80.0	92.9	50.0	80.0	77.2
LLDF	56 × 1412	6	3	4	55.0	72.2	70.0	96.4	25.0	90.0	76.3
MLDF	56 × 20	2	4	2	80.0	88.9	50.0	89.3	75.0	93.3	84.2
PAPRIKA GEOGRAPHICAL ORIGIN											
	Calibration: model parameters				**External validation: assignation performance**						
	Data matrix	**PCs**			***La Vera***		***Murcia***		**The Czech Republic**		**Accuracy**
		La Vera	***Murcia***	**The Czech Republic**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**	
FIA–HRMS (-)	66 × 553	1	2	2	96.6	100.0	100.0	100.0	100.0	100.0	99.3
FIA–HRMS (+)	66 × 601	1	2	1	93.1	100.0	70.0	100.0	66.7	100.0	94.8
LLDF	66 × 1154	1	2	1	96.6	100.0	80.0	100.0	83.3	100.0	97.0
MLDF	66 × 20	3	3	2	89.7	100.0	80.0	100.0	50.0	100.0	94.1
OLIVE OIL											
Botanical origin											
	Calibration: model parameters				**External validation: assignation performance**						
	Data matrix	**PCs**			**Olive Oil**						**Accuracy**
		Olive Oil			**Sensitivity**		**Specificity**				
FIA–HRMS (-)	53 × 368	2			83.3		100.0				90.6
FIA–HRMS (+)	53 × 739	4			77.8		100.0				87.5
LLDF	53 × 1107	1			94.4		100.0				96.9
MLDF	53 × 10	1			94.4		100.0				96.9
Quality											
	Calibration: model parameters				**External validation: assignation performance**						
	Data matrix	**PCs**			**EVOO and VOO**						**Accuracy**
		EVOO and VOO			**Sensitivity**		**Specificity**				
FIA–HRMS (-)	27 × 368	3			85.7		100.0				89.5
FIA–HRMS (+)	27 × 739	3			78.6		100.0				84.2
LLDF	27 × 1107	3			78.6		100.0				84.2
MLDF	27 × 10	1			85.7		100.0				89.5

3.3. Paprika geographical origin classification and authentication

The geographical origin authentication of paprika was also assessed through FIA–HRMS fingerprinting. In this line, three different European paprika samples —*La Vera*, *Murcia*, and the Czech Republic— were analysed in negative and positive ESI modes, providing 123 × 553 and 123 × 601 (samples × variables) dimension data matrices, respectively. The obtained plots of scores of PC1 *vs*. PC2 are presented in Fig. S4. At first glance, QC samples appeared compactly in the centre of the plot, guaranteeing the validity of the obtained chemometric results. As expected after visual inspection of paprika FIA–HRMS fingerprints in Section 3.1, *La Vera* samples were manifestly differentiated from the other samples, standing on the right side of both PCA plots. Moreover, in the negative ionisation data PCA scores plot (Fig. S4A), *Murcia* samples were separated according to their type (hot and sweet), being hot samples nearly located to the Czech Republic ones. Instead, *Murcia* and Czech samples slightly overlapped in the positive ionisation data PCA scores plot (Fig. S4B), independently of their type.

After QC exclusion, PLS-DA was applied to both matrices. Excellent geographical origin sample classification was achieved either in negative or positive ionisation modes. For example, the negative ionisation FIA–HRMS data scores plot of LV1 *vs*. LV2 (three LVs, describing 90.32% of Y-variance, were used to build the PLS-DA model) is depicted in Fig. 2B. To test and validate the classification ability of the acquired FIA–HRMS fingerprints, the following nodes were proposed to build a classification decision tree: 1) *La Vera vs*. Others and 2) *Murcia vs*. the Czech Republic. In this line, the predictive capability of the built PLS-DA models was excellent, as shown in Table 1. The lowest classification rate was 95.6%, obtained with the positive ionisation data matrix, while negative ionisation and data fusion matrices allowed the complete correct classification of the test set samples. Moreover, as shown in Table 2, similar results were obtained when subjecting FIA–HRMS fingerprints to SIMCA, proving the ability of the proposed model to authenticate the studied samples. Besides, in both PLS-DA and SIMCA studies, MLDF

provided excellent results as LLDF, indicating that a profiling approach focusing on the adequate specific markers could achieve similar results to a fingerprinting approach, which is in agreement with previous studies (Barbosa, Campmajó, Saurina, Puignou, & Núñez, 2020).

3.4. Olive oil botanical origin and quality classification and authentication

Finally, several vegetable oils described in Section 2.3.3, including olive oil, were analysed through the FIA–HRMS fingerprinting method, aiming at both botanical origin and quality olive oil authenticity. Firstly, 94 × 368 and 94 × 739 (samples × variables) matrices corresponding to negative and positive ionisation data, respectively, were subjected to PCA. After checking QC sample correct behaviour, excellent discrimination of olive oil in front of the other vegetable oil samples was observed in both cases. Besides, a trend among olive oil samples according to their quality category was also found. For instance, Fig. S5 depicts the PCA scores plot of PC1 *vs*. PC2 (describing 57.24% of the variance) obtained with the negative ionisation FIA–HRMS data matrix. In this case, olive oil samples were isolated from the other samples on the right side of the plot, displaying positive PC1 values. Moreover, PC1 also allowed a visual separation of EVOO and OO samples, while most of the VOO samples were jointly located to EVOO.

After the exploratory analysis, supervised classificatory PLS-DA was performed to address olive oil authentication. On the one hand, focusing on the botanical origin authentication issue, negative and positive ionisation data matrices were subjected to PLS-DA, providing similar results. For instance, Fig. 2C contains the PLS-DA scatter plot of scores on the LV2-LV1 —corresponding to the three LVs built model— obtained with the negative ionisation FIA–HRMS data matrix, where olive oil samples were discriminated on the left side of the plot, displaying negative LV1 values. Besides, among the other vegetable oil samples, located on the right side of the plot, soy oil and its mix with sunflower oil samples displayed positive LV2 values, whereas the remaining samples presented negative LV2 values.

Since this study aimed to classify and authenticate olive oil in front of other vegetable oils, independently of their botanical origin, a two-input class PLS-DA single step —Olive oil *vs*. Others— was proposed. Table 1 shows the most relevant parameters of the PLS-DA calibration models and the classification performance after the external validation. As a result, except for the MLDF fingerprints that allowed a 96.9% classification accuracy, the constructed chemometric models provided the correct classification of all the test samples. Moreover, as shown in Table 2, excellent results were obtained when using SIMCA, particularly for LLDF and MLDF FIA–HRMS fingerprints, which allowed an overall accuracy result of 96.9%.

Therefore, considering the great discrimination ability achieved with the MLDF FIA–HRMS fingerprints, which suggested the suitability of a profiling strategy to address this food authentication issue, the *m/z* signals that formed the matrix were studied, and some of them were putatively identified. In this line, the *m/z* 123.0451, 137.0243, and 137.0606, found in negative ionisation FIA–HRMS spectra, were assigned as the deprotonated molecule of 4-methylcatechol, hydroxybenzoic acid, and tyrosol, respectively; while in the positive ionisation, the *m/z* 415.1360 was identified as the $[M + Na]^+$ form of methyl oleuropein aglycone. Instead, the remaining discriminating ions (*m/z* 263.0534, 277.0329, 281.0644, 309.0595, and 735.4107 in the negative ionisation fingerprints, and 805.5800 in the positive ionisation ones) could not be identified. The fact that the identified compounds corresponded to substances well-known for their presence in olive oil proved the correct variable selection strategy, detailed in Section 2.4.2, through VIP and selectivity ratio values (see Fig. S6). Moreover, when comparing the corresponding PCA scores plot with its loadings plot (Fig. S7), the selected ten variables were found on the right side of the plot, showing a direct correlation with olive oil samples.

On the other hand, olive oil quality authentication was also evaluated by subjecting the acquired FIA–HRMS data —46 × 368 and 46 × 739 (samples × variables), negative and positive ionisation data matrices— to PLS-DA. In this context, Fig. 2D represents the scatter plot of scores of LV1 *vs*. LV2, describing 56.11% of Y-variance, attained using negative ionisation fingerprints. Similarly to the previous exploratory analysis results, good discrimination along the LV1 between EVOO and OO samples was observed. Concerning VOO samples, they seemed to be nearly positioned to EVOO ones. Thus, considering the EVOO and VOO similarities found in both exploratory PCA and supervised PLS-DA and the scarcity of VOO samples in the sample set, they were conjointly considered in the following classification and authentication study. Again, as performed in the botanical origin classification, a two-input class PLS-DA model, consisting of EVOO and VOO *vs*. OO, was proposed. As shown in Table 1, while positive ionisation and MLDF fingerprints reached an 89.5% classification accuracy, the corresponding negative ionisation and LLDF ones achieved a 94.7%. Regarding the SIMCA study, as shown in Table 2, negative ionisation and LLDF FIA–HRMS fingerprints provided again the best results, reaching an accuracy of 89.5%.

4. Conclusions

The three representative cases under study (red wine, paprika, and olive oil) differed in the complexity of the samples as follows: (i) red wines presented similar FIA–HRMS fingerprints without following a clear characteristic pattern due to geographical origin and lacking specific markers for the different classes. Besides, the interregional diversity due to varietal, climatic and geographical features made sample classification a complex issue. (ii) Instead, for paprika samples, distinctive FIA–HRMS fingerprints were observed according to sample geographical origin. These differences could be related to the manufacturing processes and peculiarities of each origin. (iii) Finally, in the case of oils, a similar situation was faced since compositional FIA–HRMS fingerprints of olive oils differed considerably from those of other vegetable sources (specific biomarkers could be encountered).

Therefore, FIA–HRMS fingerprinting, combined with chemometrics, has proved to be a suitable high-throughput analytical approach to address the food classification and authentication issues under study since remarkable classification accuracies were obtained after external validation. Moreover, HRMS conferred great selectivity and molecular coverage, leading to rich fingerprints, resulting in satisfactory results when using either negative ionisation, positive ionisation, or LLDF data. Furthermore, the successful application of the MLDF strategy to some of the studied food authentication cases also suggested the eligibility of targeted profiling approaches, focusing on specific compounds, to assess them.

Funding

The authors are grateful for the financial support received from the Spanish Ministry of Science, Innovation and Universities, through the grant PGC2018-095013-B-I00, and from the Agency for Administration of University and Research Grants (Generalitat de Catalunya), through the grants 2017SGR-171 and 2017SGR-310.

CRediT authorship contribution statement

Guillem Campmajó: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Supervision. **Javier Saurina:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **Oscar Núñez:** Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

G. Campmajó thanks the University of Barcelona for the PhD APIF fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2021.131491.

References

Arrizabalaga-Larrañaga, A., Campmajó, G., Saurina, J., Núñez, O., Santos, F. J., & Moyano, E. (2021). Determination of capsaicinoids and carotenoids for the characterization and geographical origin authentication of paprika by UHPLC–APCI–HRMS. *Lwt, 139*, 110533. https://doi.org/10.1016/j.lwt.2020.110533

Barbosa, S., Campmajó, G., Saurina, J., Puignou, L., & Núñez, O. (2020). Determination of Phenolic Compounds in Paprika by Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectrometry: Application to Product Designation of Origin Authentication by Chemometrics. *Journal of Agricultural and Food Chemistry, 68*(2), 591–602. https://doi.org/10.1021/acs.jafc.9b0605410.1021/acs.jafc.9b06054.s00110.1021/acs.jafc.9b06054.s002

Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., & Busto, O. (2015). Data fusion methodologies for food and beverage authentication and quality assessment - A review. *Analytica Chimica Acta, 891*, 1–14. https://doi.org/10.1016/j.aca.2015.04.042

Cetó, X., Serrano, N., Aragó, M., Gámez, A., Esteban, M., Díaz-Cruz, J. M., & Núñez, O. (2018). Determination of HPLC-UV fingerprints of spanish paprika (Capsicum annuum L.) for its classification by linear discriminant analysis. *Sensors, 18*(12), 4479. https://doi.org/10.3390/s18124479

Chen, P., Sun, J., & Ford, P. (2014). Differentiation of the four major species of cinnamons (C. burmannii, C. verum, C. cassia, and C. loureiroi) using a flow injection mass spectrometric (FIMS) fingerprinting method. *Journal of Agricultural and Food Chemistry*, *62*(12), 2516–2521. https://doi.org/10.1021/jf405580c.

Cody, R. B., Laramée, J. A., & Durst, H. D. (2005). Versatile new ion source for the analysis of materials in open air under ambient conditions. *Analytical Chemistry, 77*(8), 2297–2302. https://doi.org/10.1021/ac050162j10.1021/ac050162j.s00110.1021/ac050162j.s00210.1021/ac050162j.s003

Cuadros-Rodríguez, L., Ruiz-Samblás, C., Valverde-Som, L., Pérez-Castaño, E., & González-Casado, A. (2016). Chromatographic fingerprinting: An innovative approach for food "identitation" and food authentication - A tutorial. *Analytica Chimica Acta, 909*, 9–23. https://doi.org/10.1016/j.aca.2015.12.042

Du, L., Lu, W., Cai, Z. (Julia), Bao, L., Hartmann, C., Gao, B., & Yu, L. (Lucy). (2018). Rapid detection of milk adulteration using intact protein flow injection mass spectrometric fingerprints combined with chemometrics. *Food Chemistry*, *240*, 573–578. https://doi.org/10.1016/j.foodchem.2017.07.107.

Farré, M., Picó, Y., & Barceló, D. (2019). Direct analysis in real-time high-resolution mass spectrometry as a valuable tool for polyphenols profiling in olive oil. *Analytical Methods, 11*(4), 472–482. https://doi.org/10.1039/c8ay01865k

Gachumi, G., Purves, R. W., Hopf, C., & El-Aneed, A. (2020). Fast Quantification without Conventional Chromatography, the Growing Power of Mass Spectrometry. *Analytical Chemistry, 92*(13), 8628–8637. https://doi.org/10.1021/acs.analchem.0c00877

Gao, B., Lu, Y., Sheng, Y., Chen, P., & Yu, L. (2013). Differentiating organic and conventional sage by chromatographic and mass spectrometry flow injection fingerprints combined with principal component analysis. *Journal of Agricultural and Food Chemistry, 61*(12), 2957–2963. https://doi.org/10.1021/jf304994z

Gao, B., Qin, F., Ding, T., Chen, Y., Lu, W., & Yu, L. (2014). Differentiating organically and conventionally grown oregano using ultraperformance liquid chromatography mass spectrometry (UPLC-MS), headspace gas chromatography with flame ionization detection (Headspace-GC-FID), and flow injection mass spectrum (FIMS). *Journal of Agricultural and Food Chemistry, 62*(32), 8075–8084. https://doi.org/10.1021/jf502419y

Garrido, J., & Borges, F. (2013). Wine and grape polyphenols - A chemical perspective. *Food Research International, 54*(2), 1844–1858. https://doi.org/10.1016/j.foodres.2013.08.002

Gosetti, F., Bolfi, B., Manfredi, M., Calabrese, G., & Marengo, E. (2015). Determination of eight polyphenols and pantothenic acid in extra-virgin olive oil samples by a simple, fast, high-throughput and sensitive ultra high performance liquid chromatography with tandem mass spectrometry method. *Journal of Separation Science, 38*(18), 3130–3136. https://doi.org/10.1002/jssc.201500452

Gutiérrez-Escobar, R., Aliaño-González, M. J., & Cantos-Villar, E. (2021). Wine Polyphenol Content and Its Influence on Wine Quality and Properties: A Review. *Molecules, 26*(3), 718. https://doi.org/10.3390/molecules26030718

Hroboňová, K., & Sádecká, J. (2020). Coumarins content in wine: Application of HPLC, fluorescence spectrometry, and chemometric approach. *Journal of Food Science and Technology, 57*(1), 200–209. https://doi.org/10.1007/s13197-019-04048-2

Ibáñez, C., Simó, C., García-Cañas, V., Acunha, T., & Cifuentes, A. (2015). The role of direct high-resolution mass spectrometry in foodomics. *Analytical and Bioanalytical Chemistry, 407*(21), 6275–6287. https://doi.org/10.1007/s00216-015-8812-1

Ivanova-Petropulos, V., Naceva, Z., Sándor, V., Makszin, L., Deutsch-Nagy, L., Berkics, B., … Kilár, F. (2018). Fast determination of lactic, succinic, malic, tartaric, shikimic, and citric acids in red Vranec wines by CZE-ESI-QTOF-MS. *Electrophoresis, 39*(13), 1597–1605. https://doi.org/10.1002/elps.201700492

Kuo, T. H., Dutkiewicz, E. P., Pei, J., & Hsu, C. C. (2020). Ambient Ionization Mass Spectrometry Today and Tomorrow: Embracing Challenges and Opportunities. *Analytical Chemistry, 92*(3), 2353–2363. https://doi.org/10.1021/acs.analchem.9b05454

Massart, D. L., Vandeginste, B. G. M., Buydens, L. M. C., de Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1997). *Handbook of chemometrics and qualimetrics* (1st ed.). Amsterdam, The Netherlands: Elsevier Science.

Medina, S., Perestrelo, R., Silva, P., Pereira, J. A. M., & Câmara, J. S. (2019). Current trends and recent advances on food authenticity technologies and chemometric approaches. *Trends in Food Science and Technology, 85*, 163–176. https://doi.org/10.1016/j.tifs.2019.01.017

Morin, J.-F., & Lees, M. (2018). *FoodIntegrity handbook. A guide to food authenticity issues and analytical solutions* (J.-F. Morin & M. Lees, eds.). https://doi.org/10.32741/fihb.

Nanita, S. C., & Kaldon, L. G. (2016). Emerging flow injection mass spectrometry methods for high-throughput quantitative analysis. *Analytical and Bioanalytical Chemistry, 408*(1), 23–33. https://doi.org/10.1007/s00216-015-9193-1

Pérez-Gálvez, A., Garrido-Fernández, J., Mínguez-Mosquera, M. I., Lozano-Ruiz, M., & Montero-de-Espinosa, V. (1999). Fatty Acid Composition of Two New Pepper Varieties (Capsicum annuum L. cv. Jaranda and Jariza). Effect of Drying Process and Nutritional Aspects. *Journal of the American Oil Chemists' Society, 76*(2), 205–208. https://doi.org/10.1007/s11746-999-0219-8

Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics, 11*, 395. https://doi.org/10.1186/1471-2105-11-395

Points, J., Manning, L., & Group, A. (2020). Facing up to food fraud in a pandemic. *Food Science and Technology, 34*(2), 16–20. https://doi.org/10.1002/fsat.3403_4.x.

Rodionova, O. Y., Titova, A. V., & Pomerantsev, A. L. (2016). Discriminant analysis is an inappropriate method of authentication. *TrAC - Trends in Analytical Chemistry, 78*, 17–22. https://doi.org/10.1016/j.trac.2016.01.010

Rothwell, J. A., Perez-Jimenez, J., Neveu, V., Medina-Remón, A., M'Hiri, N., García-Lobato, P., … Scalbert, A. (2013). Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content. *Database*, *2013*. https://doi.org/10.1093/database/bat070.

Rubert, J., Zachariasova, M., & Hajslova, J. (2015). Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 32*(10), 1685–1708. https://doi.org/10.1080/19440049.2015.1084539

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environmental Science and Technology, 48*(4), 2097–2098. https://doi.org/10.1021/es5002105

Smith, C. A., Maille, G. O??., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., … Siuzdak, G. (2005). METLIN: A metabolite mass spectral database. *Therapeutic Drug Monitoring, 27*(6), 747–751. https://doi.org/10.1097/01.ftd.0000179845.53213.39

Sun, J., Zhang, M., Kubzdela, N., Luo, Y., Harnly, J. M., & Chen, P. (2018). Determination of Variance of Secondary Metabolites in Lettuces Grown Under Different Light Sources by Flow Injection Mass Spectrometric (FIMS) Fingerprinting and ANOVA–PCA. *Journal of Analysis and Testing, 2*(4), 312–321. https://doi.org/10.1007/s41664-018-0072-6

Takáts, Z., Wiseman, J. M., Gologan, B., & Cooks, R. G. (2004). Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science, 306*(5695), 471–473. https://doi.org/10.1126/science.1104404

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., … Scalbert, A. (2018). HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Research*, *46*(D1), D608–D617. https://doi.org/10.1093/nar/gkx1089.

Wold, S. (1976). Pattern recognition by means of disjoint principal components models. *Pattern Recognition, 8*(3), 127–139. https://doi.org/10.1016/0031-3203(76)90014-5