Large-scale evaluation of shotgun triacylglycerol profiling for the fast detection of olive oil adulteration

Beatriz Quintanilla-Casas [a,b], Giulia Strocchi [a], Julen Bustamante [a,b], Berta Torres-Cobos [a,b], Francesc Guardiola [a,b], Wenceslao Moreda [c], José Manuel Martínez-Rivas [c], Enrico Valli [d], Alessandra Bendini [d], Tullia Gallina Toschi [d], Alba Tres [a,b,*], Stefania Vichi [a,b]

[a] Departament de Nutrició, Ciències de l'Alimentació i Gastronomia, Campus De l'Alimentació Torribera, Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Santa Coloma de Gramenet, Spain
[b] Institut de Recerca en Nutrició i Seguretat Alimentària (INSA-UB), Universitat de Barcelona, Santa Coloma de Gramenet, Spain
[c] Instituto de la Grasa (CSIC), Sevilla, Spain
[d] Department of Agricultural and Food Sciences, Alma Mater Studiorum - Università di Bologna, Bologna, Italy

ARTICLE INFO

ABSTRACT

Fast and effective analytical screening tools providing new suitable authenticity markers and applicable to a large number of samples are required to efficiently control the global olive oil (OO) production, and allow the rapid detection of low levels of adulterants even with fatty acid composition similar to OO. The present study aims to develop authentication models for the comprehensive detection of illegal blends of OO with adulterants including different types of high linoleic (HL) and high oleic (HO) vegetable oils at low concentrations (2–10%) based on shotgun triacylglycerol (TAG) profile obtained by Flow Injection Analysis-Heated Electrospray Ionisation-High Resolution Mass Spectrometry (FIA-HESI-HRMS) at a large-scale experimental design. The sample set covers a large natural variability of both OO and adulterants, resulting in more than one thousand samples analysed. A combined PLS-DA binary modelling based on shotgun TAG profiling proved to be a fit for purpose screening tool in terms of efficiency and applicability. The external validation resulted in the correct classification of the 86.8% of the adulterated samples (diagnostic sensitivity = 0.87), and the 81.1% of the genuine samples (diagnostic specificity = 0.81), with an 85.1% overall correct classification (efficiency = 0.85).

1. Introduction

Olive oil (OO) authenticity has become a focal point attracting the attention of producers, consumers and policy makers due to the high risk of fraud derived from the high economic value of OO compared to other edible oils (Everstine et al., 2013; European Commission, 2018). In this context, the blending with oils of lower economic value is among the most common economically motivated adulterations in OO (Everstine et al., 2013; Mailer & Gafner, 2020; Yan et al., 2020). Illegal blending of OOs with high-linoleic (HL) seed oils such as canola, sunflower and soybean oils or with high-oleic (HO) oils such as high-oleic sunflower or hazelnut oils has been reported, but adulterations with new OO adulterants (alone or combined), such as avocado oil, are also emerging (Mailer & Gafner, 2020). Apart from the economic damage to stakeholders and consumers, food safety concerns are among the

consequences of this fraudulent practice, especially when it is committed using substances from unknown sources. The latter might be toxic or have negative health effects, such as the presence of allergens from seed oils or the Spanish toxic oil syndrome outbreak (Gelpí et al., 2002; Arlorio et al., 2010). Improving the detection of illegal blends between OOs and oils of different botanical origin would contribute to reduce fraud within the sector, increasing its competitiveness and reducing the chances of food safety crisis.

The European Regulation (EEC Regulation 2568/91 and subsequent amendments, (European Commission Regulation, 1991)) and the International Olive Council (IOC) standards (IOC 2017; 2018a; 2018b; 2019) address this issue by means of the analysis of fatty acids (FAs), triacylglycerols (TAGs) and sterols. But fraudsters may remove some minor compounds to mask the adulteration, such in the case of desterolized seed oils. In this regard, major OO constituents, such as FAs and the corresponding TAGs are more difficult to alter to mask fraud. In

---

Abbreviations

AO	Avocado Oil
B	Behenic acid
EVOO	Extra Virgin Olive Oil
FA	Fatty Acid
FIA-HESI-HRMS	Flow Injection Analysis-Heated Electrospray-HRMS
G	Gondoic acid
HL	High-Linoleic
HO	High-Oleic
HOSFO	High-Oleic Sunflower Oil
HRMS	High Resolution Mass Spectrometry
HZO	Hazelnut Oil
L	Linoleic acid
Li	Lignoceric acid
Ln	Linolenic acid
O	Oleic acid
OO	Olive oil
OOc	Olive Oil commercial Category
P	Palmitic acid
Po	Palmitoleic acid
PC	Principal Component
PCA	Principal Component Analysis
PLS-DA	Partial Least Squares-Discriminant Analysis
RMSEcv	Root Mean Square Error of Cross Validation
S	Stearic acid
SBO	Soybean Oil
SFO	Sunflower Oil
SEcv	Standard Error of Cross-Validation
TAG	Triacylglycerol
VOO	Virgin Olive Oil

particular, TAG profile is expected to provide more information than FAs about the source of the oil because not only its composition but also the arrangement of FAs in the glycerol molecule are genetically determined and, thus, linked to the botanical origin (Sánchez & Harwood, 2002; Vichi et al., 2007). However, these official methods for OO purity control present some limitations (Conte et al., 2020). On the one hand, conventional methods are based on time- and reagent-consuming analytical procedures that make them unsuitable as screening methods, limiting the annual number of conformity checks to one per thousand tons of OO marketed in each Member State (EEC Regulation No 2568/91 and amendments, (European Commission Regulation, 1991)). These controls, which should be performed in all the relevant stages of the supply chain, do not guarantee a fully satisfactory level of consumer protection, and should be further improved (European Commission report, 2018). On the other hand, current methods do not always perform satisfactorily when the adulterants are present at low concentration or when they show compositions similar to that of OO, such as hazelnut oil or high-oleic sunflower oil (Conte et al., 2020; Mailer & Gafner, 2020). In fact, false positive results coming from these drawbacks caused the withdrawal from the EU legislation of the global method based on TAG analysis (Conte et al., 2020). Hence, faster and more effective analytical screening tools providing new suitable authenticity markers and applicable to a large number of samples are required for an efficient control of the global OO production, fulfilling the consumers' expectations.

Although several alternative methods have been proposed in the last years (Aparicio et al., 2013; Bajoub et al., 2018; Meenu et al., 2019), the rapid detection of adulterants at low levels (below 20%) is still challenging and further research is needed (Georgouli et al., 2017), particularly in the detection of HO oils such as hazelnut oil (HZO), high-oleic sunflower oil (HOSFO) or avocado oils (AO), whose composition is similar to OO, and that have been found in adulterations (Mailer & Gafner, 2020). In this sense, innovative high-dimensional methods are emerging besides the conventional targeted approach. These approaches are usually based on high-throughput analyses identifying specific patterns to distinguish authentic from adulterated samples (Bajoub et al., 2018; Cavanna et al., 2018), rather than only a few markers as target methods do (Esslinger et al., 2014) providing a higher efficiency in fraud detection. In this regard, technological advances in high resolution mass spectrometry (HRMS) (Xiang et al., 2012) combined to soft ionization techniques, have led to high-efficiency analysis such as shotgun lipid profiling. This application enabled the fast speciation of a large number of acylglycerol molecules by accurate mass measurement and elemental formulae assignment in a single run, with high selectivity and minimal sample preparation (Gómez-Ariza et al., 2006; Vaclavik et al., 2009; Cozzolino & De Giulio, 2011; Vichi et al., 2012; Longo et al., 2018). The

large number of potentially diagnostic minor ions not detectable by conventional methods offers valuable information for authentication purposes (Vichi et al., 2012). For a detailed profiling including minor TAGs, HRMS is mandatory to overcome the interference of multiple isobaric compounds. These fast and high-efficient features make shotgun TAG profiling a powerful screening tool for lipid food authentication, particularly if HRMS data are managed by chemometrics to obtain authentication models. Preliminary assays to detect low percentages of vegetable adulterant oils in OO revealed the potential of this authentication tool (Quintanilla-Casas et al., 2018). However, to evaluate the efficiency of any authentication methodology, robust models including the sufficient number of representative samples are required (McGrath et al., 2018). In this case, they should cover the actual variability of genuine OOs, including distinct harvest-seasons, geographical regions, olive cultivars and technological conditions to produce the oil, as well as containing different adulterant oils to consider the variability of both OO and adulterant, which is essential for the applicability of the future model. Although numerous studies in literature develop analytical tools to detect adulteration in OO, many of them are based on a limited OO and/or adulterant sample size (Bajoub et al., 2018; Meenu et al., 2019). Even if they provide successful results, they should be considered a proof of concept for the detection of OO adulteration, and thus, there is still a need for models that are tested on large sample sets, including the variability of both the OO and adulterant, and including emerging adulterants.

The present study aims to develop authentication models for detecting illegal blends of OO with different types of HL or HO vegetable oils at low concentrations (2–10%) based on shotgun TAG profile obtained by Flow Injection Analysis-Heated Electrospray Ionisation-High Resolution Mass Spectrometry (FIA-HESI-HRMS) at a large-scale experimental design. In fact, the sample set covers a high natural variability of both OO and adulterants, resulting in more than one thousand samples. Moreover, it also includes the analytical variability related to batch-to-batch variations that could hinder the robustness of the MS method (Rusilowicz et al., 2016), as the full sample set was analysed randomly in four main batches at intervals of 2–6 months. Finally, this study also includes the verification of model reliability to detect adulteration with either HL ($\geq$2%) or HO ($\geq$5%) oils by external validation.

2. Material and methods

2.1. Chemicals

Dichloromethane (SupraSolv® for GC-ECD/FID), methanol (SupraSolv® for gas chromatography) and NaCl (ACS reagent, $\geq$99.0%) were purchased from Merck (Darmstadt, Germany). Nitrogen (Alphagaz,

purity 99.999%, Air Liquide) was used in the Orbitrap-Exactive as the nebulization gas.

2.2. Samples

The present study includes 346 genuine and traceable samples from two sample banks: i) 170 extra virgin (EVOO) and virgin olive oil (VOO) samples produced in Catalonia (Spain), obtained in the framework of Autenfood project (ACCIÓ- Programa Operatiu FEDER Catalunya 2014–2020); and ii) 174 EVOOs and 2 samples of olive oil category (OOc), all obtained from different EU and non-EU countries in the framework of OLEUM project (EC H2020 Programme 2014–2020). Oils from both sample banks were produced during two consecutive crop years (2016/17 and 2017/18).

A total of 60 adulterant oils of different botanical origins -including vegetable oils of prevalence in emerging frauds-were provided by producers or by commercial suppliers. The set included high-linoleic (HL) oils: sunflower oils (SFO, refined, n = 15) and soybean oils (SBO, refined, n = 10), as well as high-oleic (HO) oils: hazelnut oils (HZO, virgin and refined, n = 10), high-oleic sunflower oil (HOSFO, refined, n = 15) and avocado oils (AO, virgin and refined, n = 10). They were used to prepare blends with OO, as described below.

Additionally, 32 blind samples were provided in the frame of Autenfood and OLEUM projects, consisting of 4 HL blends of OO and EVOO with a HL adulterant (SFO) at 10–15% (10%, n = 2; 15%, n = 2) and 28 blends of OO and EVOO with 5 types of HO adulterants (virgin and refined HZO; virgin and refined AO; and HOSFO) at 5–20% (5%, n = 5; 10%, n = 5; 15%, n = 13; 20%, n = 5).

2.3. Experimental design

A total of 150 genuine OO samples were randomly selected, including 90 from the Autenfood study and 60 from the OLEUM study, and were used for the in-house preparation of blends with 2% and 5% of HL adulterants and with 5% and 10% of HO adulterants, following a balanced incomplete Latin squares experimental design (Supplementary information, S1). Additionally, 28 blends from a preliminary trial, obtained from one OOc and one EVOO blended with one HL (SFO) and six HO adulterants (HZO, AO, HOSFO), at the same concentrations were included. The fact that HO adulterants and OO have a high similarity in TAG profile, made us to test higher concentrations of HO adulterants (5% and 10%) than HL adulterants (2% and 5%). The rest of genuine samples were analysed to enlarge the genuine samples pool and increase the representativeness of the model. Finally, the full sample set consisted of 1006 samples: 628 blends (obtained from 152 genuine oils and 60 adulterants), 346 genuine oils (those 152 used to prepare the blends plus

Table 1
Full sample set, training set and validation set analysed for the HL and the HO submodels.

	Full set	Outliers	Training set	Validation set
HL blends				
2%	152	–	122	30
5%	152	2	120	30
Total HL blends	304	2	242	60
Blind HL blends	4	–	–	4
HO blends				
5%	162	6	124	32
10%	162	8	122	32
Total HO blends	324	14	246	64
Blind HO blends	28	–	–	28
Total samples				
Total blends	628	–	488	161
Genuine samples	346	7	272	67
Blind blends	32	–	–	32
Total samples	1006	23	760	223

HL: high linoleic; HO: high oleic.

other 194 oils) and 32 blind blends (Table 1). This full sample set was randomly analysed in four main batches at intervals of 2–6 months between them.

2.4. Flow Injection Analysis-Heated Electrospray-high resolution mass spectrometry

TAG profiling was carried out according to Vichi et al. (2012). Samples were prepared as follows: 30 mg of oil dissolved in 3 mL of dichloromethane:methanol (70:30, v/v) were diluted to 1:100 using the same solvent mixture, and the solution was saturated with NaCl as the cationisation agent. After vortex mixing during 30 s, 10 μL of the supernatant was further diluted to 1:50 with the same solvent mixture and, finally, analysed.

Flow injection analysis of 5 μL of the samples was carried out with an Orbitrap-Exactive HCD (Thermo Fisher Scientific, Bremen, Germany) equipped with an electrospray source (H-ESI II) and coupled to an Accela pump (Thermo Fisher Scientific, San Jose, CA). The mobile phase was methanol:dichloromethane (80:20, v/v) at a flow rate of 50 μL/min. Mass spectra were acquired in full scan positive ionization mode under the following conditions: spray voltage 3.00 kV, capillary voltage 37 V, tube lens 150 V and skimmer voltage 40 V. The sheath gas flow rate was at 35 au (arbitrary units), and the aux gas flow rate was 5 au. Capillary and heater temperatures were set at 400 °C and 30 °C, respectively. The mass range was set to m/z 200–1200, and the acquisition time was 1.5 min. Two blanks were analysed before each sample. The automatic gain control was used to fill the C-trap and gain accuracy in mass measurements (ultimate mass accuracy mode, 5×10^5 ions). Resolving power defined as R: 100,000 (m/z 200, FWHM) was set. The instrument was calibrated before the analysis of each sequence of samples (every 24–48 h).

The mass peaks considered were single positive charged sodium molecular ions matching with the following criteria, set to generate reliable elemental formulae that could agree with TAG compounds: C ≤ 200, H ≤ 400, O = 6, Na = 1 and rings and double bond equivalents (RDBE): 2.5–19.5. Mass error tolerance was set at 5 ppm. The molecular formulae calculation was performed with Xcalibur 4.1 (Thermo Fisher Scientific, Bremen, Germany), and the subsequent data analysis was done using excel files. Relative abundances obtained for each TAG signal were normalized and expressed as percentage.

2.5. Chemometrics

2.5.1. Data pre-processing and exploration

Multivariate analysis was performed with SIMCA software v13.0© (Umetrics AB, Sweden). After data pre-processing (mean centring and scaling to unit variance) and variable selection (exclusion of those variables with a number of non-median values under 50), a Principal Component Analysis (PCA) was developed (3 principal components (PC), 34.5% explained variance) to explore the model and to detect outlying samples according to Hotelling's T^2 and Q-residuals. After exclusion of outliers (n = 23), the data matrix of the full sample set consisted of 951 samples x 292 variables.

2.5.2. Classification strategy by binary modelling

Partial Least Squares-Discriminant Analysis (PLS-DA) was applied to develop two independent binary models that discriminated between (i) genuine oils and blends with ≥2% of HL adulterants, and (ii) genuine oils and blends with ≥5% of HO adulterants (Fig. 1). In PLS-DA binary models, classes are expressed as PLS dummy variables (here being 0 for non-adulterated, and 1 for adulterated class). Then, the PLS predicted value (PV) of each sample is used to classify it into one class or the other according to a classification threshold (here, predicted value = 0.5). Thus, test samples were classified as non-adulterated (with HO or HL oils) when both models classified them as non-adulterated (predicted
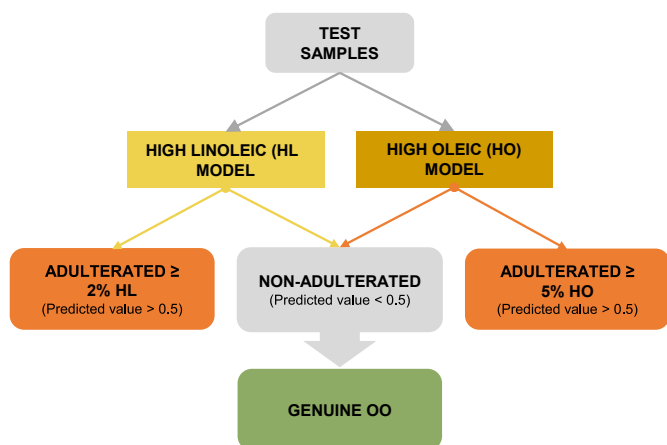
Fig. 1. Classification model diagram based on two combined binary PLS-DAs to predict the presence of HL (≥2%) and HO (≥5%) adulterants in olive oils. Classification threshold set at prediction value = 0.5.

value for HL and HO class < 0.5), and as adulterated (predicted value for HL and HO class > 0.5) when classified in this class at least by one of the models (HL or HO).

First, the sample set (n = 951) was randomly split three times into a training set containing the 80% of samples (n = 760), and an external validation set with the remaining 20% (n = 191). In each validation set, samples were selected at random but the number of samples of each class (genuine samples, n = 67; HL blends, n = 60; HO blends, n = 64) was proportional to the number of training samples of that class. A balance between the training set and the validation set was also kept in the sample proportions of each analytical batch and harvest season (2016/17, 2017/18), and in the percentage and type of adulterant of the blends (2% and 5% of HL, both n = 30; 5% and 10% of HO, both n = 32; SFO, n = 30; SBO, n = 30; HO, n = 27; HZO, n = 27; AO, n = 11).

Then, PLS-DA was applied to each training set of samples to develop two binary classification models: to discriminate between genuine OO and blends with ≥2% of HL (n = 514) or between OO and blends with ≥5% of HO (n = 518). Root Mean Square Error of Cross Validation (RMSEcv) and Q^2 values were evaluated to select the number of latent variables (LVs) for each model. These training models were internally validated by leave-10%-out cross-validation and the percent of correct classifications was used to evaluate the performance of the models. The risk of model overfitting was assessed by examining the RMSEcv, the ANOVA of cross-validated residuals and by performing permutation tests (n = 20).

2.5.3. External validation

The external validation of each of the three training models was carried out by predicting the authenticity of the samples in the corresponding validation sets, as they had not been used to develop that model. Additionally, to ensure that the models could also correctly classify blends with slightly higher percentages of adulterants, 32 blind samples described in section 2.2 were included, resulting in a total of 223 external validation samples (genuine, n = 67; HL, n = 64; HO, n = 92).

The reliability of the model to detect adulterations was evaluated by the percent of correct classification, the diagnostic sensitivity, defined as the ratio of true positives/(true positives + false negatives), the diagnostic specificity, as the ratio true negatives/(true negatives + false positives) (Magnusson & Ornemark, 2014) and the efficiency, defined as (true positives + true negatives)/(all positives + all negatives). All of them were expressed as mean values obtained from three randomly selected validation sets.

2.5.4. Exploration of regression coefficients

The TAGs that most contributed to the discriminant capacity of each binary model were identified based on the regression coefficients of the PLS-DA models developed with the full sample set. Relevant variables considered for examination were those that presented significant PLS-DA regression coefficients higher than 0.03 with a jack-knife standard error of cross-validation (SEcv) lower than the given coefficient value.

3. Results and discussion

3.1. Flow Injection Analysis-Heated Electrospray-high resolution mass spectrometry outcome

The output of lipidomic approaches is fast and highly efficient; however, it also has some limitations such as the ion suppression effects, the in-source generation of fragments that can bias the mass spectrum and the difficulties in resolving the isobaric/isomeric mass overlap (Hu et al., 2019). Here, we optimized the analytical conditions to minimize these limitations (Vichi et al., 2012). Firstly, the present method focused on a single class of compounds that are expected to present similar ionization efficiency, so the effect of ion suppression is likely to be less critical than when analysing distinct lipid classes. Secondly, promoting the formation of stable sodium adducts overcame the in-source fragmentation and provided almost fragment-free spectra (Fig. 2a). Sodium adducts of major OO TAGs stand out in the mass spectra generated by HRMS in positive mode, and a large number of minor TAG clusters can be observed at a much lower scale (Fig. 2b). Finally, the use of HRMS was essential to resolve signals of minor TAGs from isobaric interferences (Fig. 2c and d). Moreover, it provided accurate mass measurements that, together with proper restrictions criteria, led to reliable elemental formulae conforming an accurate and extremely detailed profile. Direct HRMS analysis of OO and their blends with HO and HL oils allowed the identification of more than one thousand elemental formulae matching with the fixed formula restrictions. Despite this approach did not distinguish the isomeric TAGs because they present the same molecular formula, direct HRMS analysis allowed the detection of a much higher number of species than conventional chromatographic methods. The discrimination capacity of minor TAGs has been scarcely explored due to their difficult determination by conventional methods, but the differences evidenced by FIA-HESI-HRMS in the minor TAGs profiles of distinct vegetable oils (Vichi et al., 2012) suggested a promising role of these molecules in the differentiation between genuine and adulterated OOs.

3.2. Combined classification strategy

The classification strategy aimed at distinguishing genuine from adulterated samples. Here, two types of adulterated samples were included, HL and HO, being HO and genuine samples very similar in their TAG composition. Previous studies demonstrated that in scenarios where two categories are very similar with respect to the other, binary PLS-DA models led to higher prediction efficiencies than multi-class models (Tres et al., 2013; Quintanilla-Casas et al., 2020). According to this, we developed two independent and complementary binary models discriminating between genuine OOs and their blends with HL and HO adulterants, respectively. In this classification strategy, the first model aimed to identify illegal blends of OOs with HL oils, while the second model, focused on more subtle differences between TAGs of OOs and other HO oils, aimed to differentiate genuine OOs from their blends with HO adulterants. In this way, by predicting the class of test samples by both complementary models and by combining their outputs, genuine OOs are expected to be classified as non-adulterated by both models while the presence of HL or HO adulterants is expected to be revealed by the corresponding binary model (Fig. 1).

After data pre-treatment and removal of outlying samples, the full dataset (n = 951) was randomly split three times into a training set (n =
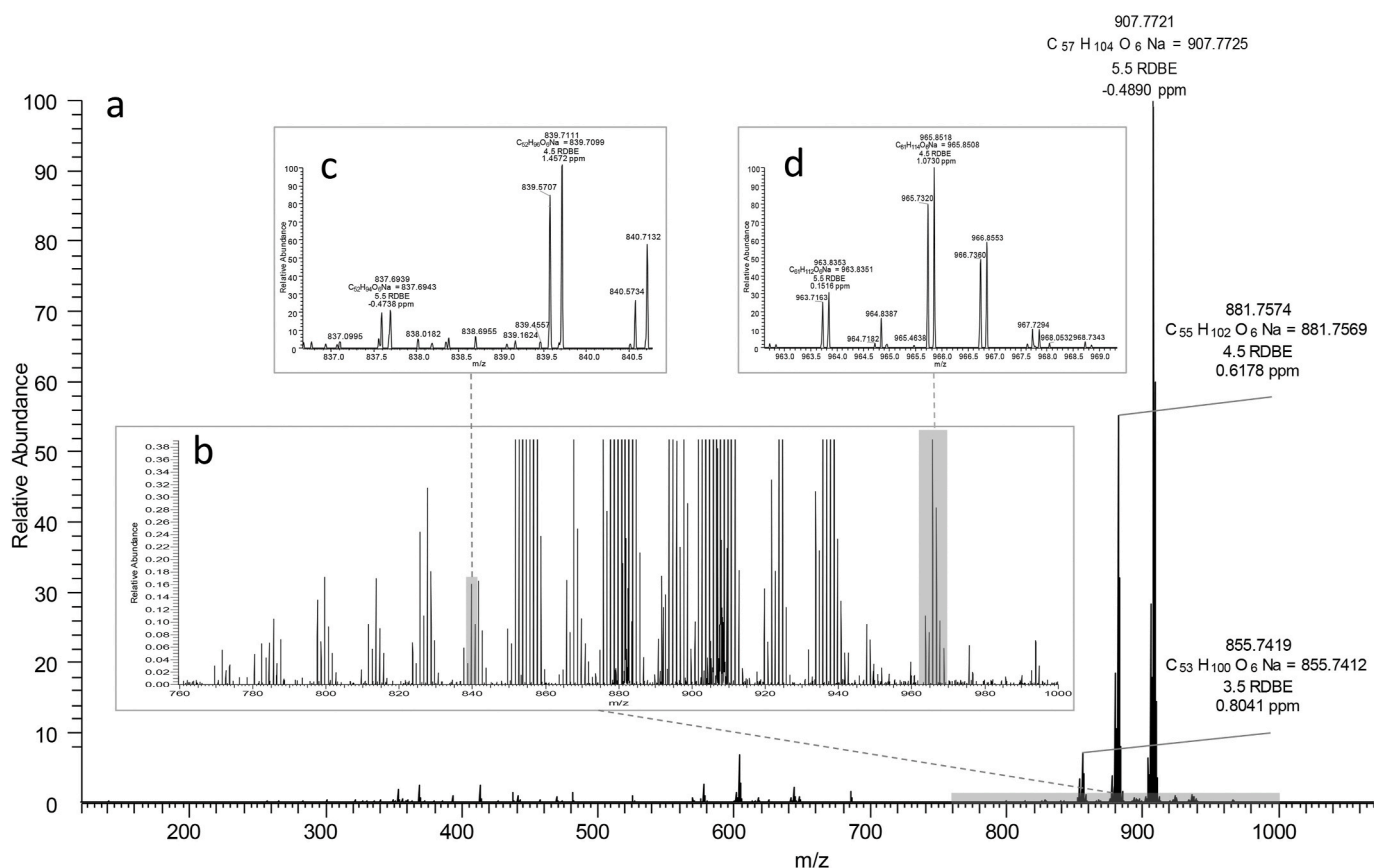
Fig. 2. a) FIA-HESI-HRMS spectrum (*m/z* 120–1200) of an OO sample, obtained by Orbitrap setting the R: 100,000 (*m/z* 200, FWHM), and displaying $[M + Na]^+$ adducts of TAGs. Experimental and theoretic exact mass, molecular formula, RDBE and mass error are reported for main signals of principal TAG clusters ($C_{57}H_{104}O_6Na$: OOO; $C_{55}H_{102}O_6Na$: POO; $C_{53}H_{100}O_6Na$: PPO, where O: oleic acid, P: palmitic acid). b) Expansion of the region in the mass range *m/z* 760–1000, evidencing minor clusters. c), d) Expanded mass regions showing how minor C49 and C58 TAGs, respectively, are successfully resolved from isobaric interferences.

80% of sample set) and a validation set (n = 20% of the sample set) as described above.

The training sets (n = 760) were used to build the OO/HL (n = 514) and the OO/HO (n = 518) PLS-DA binary models. Leave 10%-out cross-validation was applied to evaluate the discrimination capacity of the models, resulting in 90.8% and 88.3% of correct classification, respectively (mean results of three randomly selected training sets) (Table 2; plots showing the predicted values from one of the training sets are also available at Supplementary information, Figure S1). In both cases,

Table 2
Cross-validation (leave 10%-out) prediction results of each PLS-DA classification model: OO vs blends with HL adulterants (≥2%) and OO vs blends with HO adulterants (≥5%). Results are mean values obtained from three randomly selected training sets. Classification threshold = 0.5.

	n	Adulterated	Genuine	Correct class (%)
PLS-DA (OO vs HL)				
Adulterated	242	208 ± 7	34 ± 7	85.8 ± 2.8
Genuine	272	13 ± 5	259 ± 5	95.3 ± 1.8
Total	514			90.8 ± 2.1
N = 514, variables = 268, 6 LVs, $Q^2 = 0.5360^a$, RMSEcv = 0.3429^a, ANOVA p-value <0.05				
PLS-DA (OO vs HO)				
Adulterated	246	211 ± 3	35 ± 3	85.9 ± 1.0
Genuine	272	22 ± 4	250 ± 4	91.8 ± 1.3
Total	518			88.3 ± 1.8
N = 518, variables = 262, 5 LVs, $Q^2 = 0.3850^a$, RMSEcv = 0.3959^a, ANOVA p-value <0.05				

OO: olive oil; HL: high linoleic; HO: high oleic.
[a] Mean Q^2 and RMSEcv values of the three random training sets.

ANOVA results (p < 0.05) and prediction capacity of 20 random models ($Q^2 < 0$) obtained from the permutation tests indicated that the models had a high discrimination capacity and were not over-fitted.

3.3. External validation

The external validation was carried out by predicting the class of the 223 samples conforming the validation set (20% of the total sample set, n = 191, plus additional blind samples, n = 32). The complementary binary models allowed the correct classification of the 86.8% of the adulterated samples (diagnostic sensitivity = 0.87), and the 81.1% of the genuine samples conforming the validation set (diagnostic specificity = 0.81), with an 85.1% overall correct classification (efficiency = 0.85) (Table 3). Specifically, 88.5% and 85.5% of OO blends with HL (≥2%) and HO (≥5%) oils were detected, respectively.

To evaluate to what extent the performance of the present method depended upon the amount of the adulterant present in the blend, the sensitivity was calculated at each concentration tested with the validation set (Table 4). As expected, the diagnostic sensitivity was proportional to the amount of adulterant in the blend, but even at the lowest concentrations tested (2% of HL adulterants, 5% of HO adulterants), it was always above 80%. These high sensitivity and specificity values were reflected by suitable rates of (i) false negatives and (ii) false positives (Table 3), which in qualitative binary methods can be defined as (i) the proportion of positive samples classified as negative and (ii) the proportion of negative samples classified as positive (Valcárcel et al., 2002; Magnusson & Ornemark, 2014). The outputs of the complementary models allowed a false negative rate below 15% for the global model and for each of the independent binary models; and a false

Table 3

Global outcome of external validation of the classification approach based on combining two consecutive binary PLS-DAs. Results are mean values obtained from three randomly selected validation sets.

	N	Predicted category		Correct class (%)	FN[a] (%)	FP[b] (%)
		ADULTERATED	GENUINE			
ADULTERATED [c]	156	135 ± 7	21 ± 7	**86.8 ± 4.4**	13.4	
HL blends [d]	64	57 ± 6	7 ± 6	88.5 ± 8.6	11.6	
HO blends [e]	92	79 ± 2	13 ± 2	85.5 ± 1.7	14.1	
GENUINE	67	13 ± 1	54 ± 1	**81.1 ± 0.9**		19.4
Total	223			**85.1 ± 3.0**		

OO: olive oil; HL: high linoleic; HO: high oleic.

[a] False negatives (percent of adulterated samples classified as genuine).

[b] False positives (percent of genuine samples classified as adulterated).

[c] Sum of blends of OO with HL ($\geq 2\%$) and HO ($\geq 5\%$) adulterants.

[d] Blends of OO with HL ($\geq 2\%$).

[e] Blends of OO with HO ($\geq 5\%$) adulterants.

Table 4

Method sensitivity according to the adulteration rate of OOs with HL and HO adulterants, based on external validation.

	Adulterant concentration (%)	n[a]	Diagnostic sensitivity
HL	2	30	0.83
	5	30	0.92
	10	2	1
	15	2	1
HO	5	37	0.81
	10	37	0.86
	15	13	0.92
	20	5	1

OO: olive oil; HL: high linoleic; HO: high oleic.

[a] Number of blends in the external validation set at each concentration of adulterant.

positive rate below 20% (Table 3).

In comparison with other reported methods, the classification models based on shotgun TAG profiling proved to be a fit for purpose screening tool in terms of efficiency and applicability. The efficiency of the present method was higher than that reported for an interlaboratory study of the chromatographic method adopted by IOC, based on the difference between theoretical and experimental TAGs, which was 76.1–83.3% at HZO concentrations between 5 and 10% (García--González et al., 2007). Besides the better diagnostic capacity, the present authentication method is based on a much faster analytical procedure, suitable to be applied to a high number of samples as a screening tool. Furthermore, the highly representative sampling, including a high number of olive and adulterant oils from two harvest years covering a large natural sample variability, reinforces the present results in comparison to previous application of direct MS analysis, chromatographic and even spectroscopic methods. In fact, only a few previous studies reported the detection of low amounts of vegetable oil in OO based on in-house adulteration schemes using more than 25–50 genuine OO samples (Downey et al., 2002; Zou et al., 2009; Zhang et al., 2011; Sun et al., 2015; Yang et al., 2013). However, it has to be highlighted that in some of them, the number of adulterant oils used to prepare the blends was low or they lacked of an external validation, and all of them only focused on the detection of HL adulterants in OO. The application of some rapid spectroscopic techniques had reported the suitable detection of hazelnut oils up to 8%–10% in OOs (Aparicio et al., 2013), but also at a lower sampling scale than in the present study.

3.4. Exploration of PLS-DA regression coefficients

As explained in the Methods section, the models were developed on the MS signals (those that met the criteria set to obtain the formulae that could agree with TAG compounds), and thus, as in other profiling approaches, peak identification were not necessary at that stage. Once

models were successful in classifying samples, the PLS-DA significant regression coefficients of each binary model were examined to identify the variables (molecular formulae) that contributed the most to the discrimination between genuine and adulterated OOs (Fig. 3).The molecular formulae assigned to each exact mass were tentatively identified as the most plausible TAGs according to the previous knowledge on TAG profile of OO and of the adulterants used in this study. These tentatively identified TAGs were named according to their carbon number (CN), defined as the sum of carbon atoms conforming the FA moiety. In both models, the highest regression coefficients corresponded to molecular formulae matching with several TAG minor species distributed within the entire experimental mass range, thus showing the key role of minor TAGs to detect foreign oils in OO. As expected, the HL-adulterated class (Fig. 3-I) was mainly discriminated from genuine OOs by TAGs containing linoleic (L, C18:2) and linolenic (Ln, C18:3) acids, such as C54 TAGs LLL ($C_{57}H_{98}O_6Na$) and LLLn ($C_{57}H_{96}O_6Na$), and C52 TAGs such as PLL (P: Palmitic acid, C16:0) ($C_{55}H_{98}O_6Na$) and PoLL/PLLn (Po: palmitoleic acid, C16:1) ($C_{55}H_{96}O_6Na$). Moreover, other minor signals showed significant regression coefficients, matching with the molecular formulae of C26:0-C29:0 TAGs ($C_{29}H_{54}O_6Na$–$C_{31}H_{58}O_6Na$) that included short-chain FAs, or C58:2 TAGs compatible with the presence of long-chain FAs such as behenic (B, C22:0) or lignoceric (Li, C24:0) acids, i.e. LAA (A: arachidic acid, C20:0), SLB (S: stearic acid, C18:0), OOB (O: oleic acid, C18:1) or PLLi ($C_{61}H_{114}O_6Na$), in agreement with FA and TAG compositions reported in SFO and SBO (Christopoulou et al., 2004; Gao et al., 2017).

In the HO-adulterated class (Fig. 3-II), in which OO and the adulterants had similar proportions of major FAs, the most discriminant variables were signals matching with molecular formula of long-chain C58:2 and C60:2 TAGs ($C_{61}H_{114}O_6Na$; $C_{63}H_{118}O_6Na$), which according to HZO and AO FA composition (Lísa & Holčapek, 2008) are compatible with LAA/OGA (G: gondoic acid, C20:1 n-9) and LAB/OGB structure, respectively, followed by SOO ($C_{57}H_{106}O_6Na$). Also, in the case of HO adulterants, some TAGs containing L, such as LLLn ($C_{57}H_{96}O_6Na$) contributed significantly to classify adulterated samples, as well as short-chain TAGs such as C23:1 ($C_{26}H_{46}O_6Na$), C24:0 ($C_{27}H_{48}O_6Na$) and C26:0 ($C_{29}H_{54}O_6Na$).

These results revealed the high diagnostic role of minor compounds conforming OO's TAG profile for the detection of HL and HO adulterant oils in OO. These diagnostic minor TAGs, formed by short-chain, long-chain or odd-chain FAs, are hardly detectable by conventional methods. Moreover, the high number of variables showing significant contributions to each model confirmed the importance of disposing of high-dimensional multi-component approaches for the identification of specific patterns to distinguish authentic from adulterated samples.

4. Conclusions

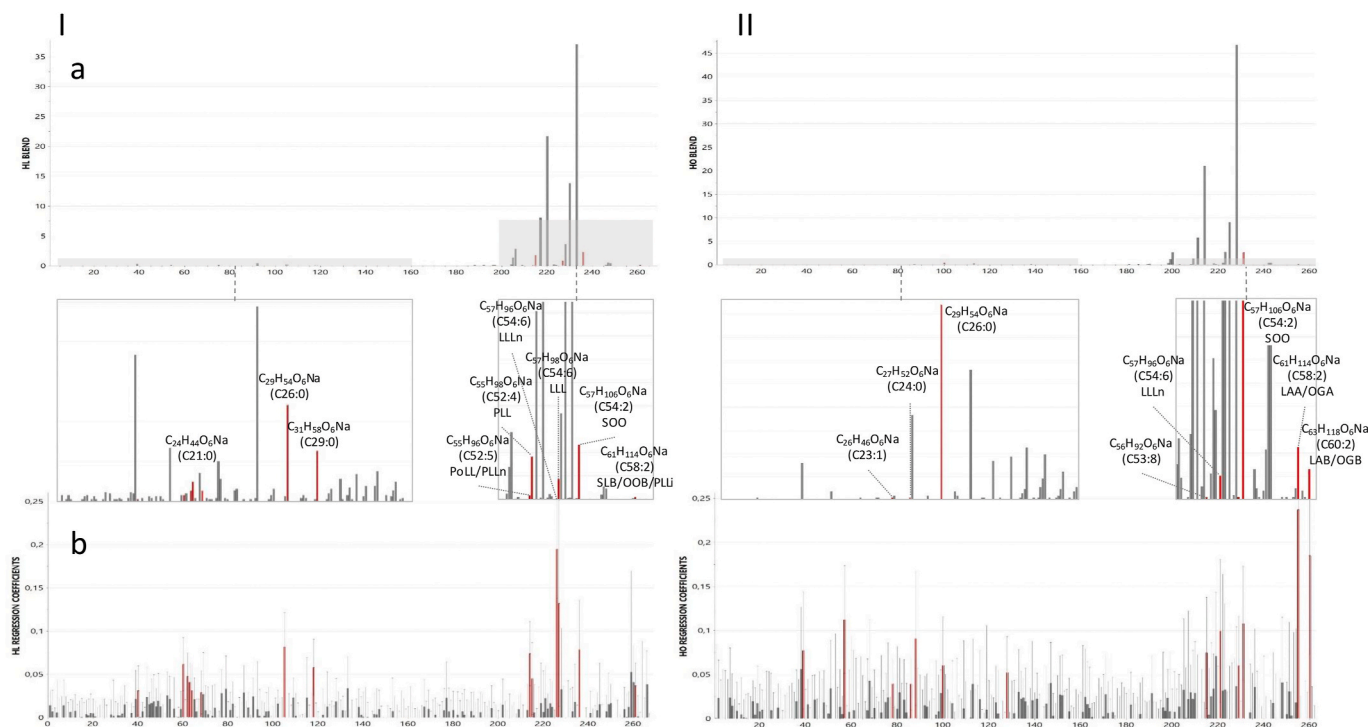This study represents a significant improvement in the development

Fig. 3. Exploration of most relevant variables in each PLS-DA binary model: I) OO vs blends with HL adulterants (≥2%) and II) OO vs blends with HO adulterants (≥5%). a) Relative intensities of variables obtained by FIA-HESI-HRMS in: I) OO with 2% SFO, II) OO with 10% AO. b) PLS-DA regression coefficients referred each adulterated class, calculated for each variable (the highest significant coefficients coloured in red). P: palmitic acid (C16:0); Po: palmitoleic acid (C16:1); S: stearic acid (C18:0); O: oleic acid (C18:1); L: linoleic acid (C18:2); Ln: linolenic acid (C18:3); A: arachidic acid (C20:0); G: gondoic acid (C20:1); B: behenic acid (C22:0); Li: lignoceric acid (C24:0). X-axis represents the TAG used to build up each classification model, sorted by its exact mass as showed in Table S2 (Supplementary information, S2). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

of screening tools to detect illegal blends of OOs with HO and HL vegetable oils. This fast method allows the comprehensive detection of both HO and HL adulterants at ≥2% and ≥5%, respectively, with an 85.1% overall correct classification, based on a large-scale sampling and supported by the external validation of the classification model. Although the application of non-targeted approaches is currently missing validation strategies and still restricted for official control purposes (Esslinger et al., 2014), the proposed method could serve as a valuable screening tool to implement risk assessment procedures to orientate official controls, making them more effective and helping the accomplishment of consumer protection actions.

CRediT authorship contribution statement

Beatriz Quintanilla-Casas: Formal analysis, Data curation, Investigation, Methodology, Validation, Writing - original draft. **Giulia Strocchi:** Formal analysis. **Julen Bustamante:** Formal analysis, Writing - review & editing. **Berta Torres-Cobos:** Data curation, Formal analysis. **Francesc Guardiola:** Investigation, Supervision, Writing - review & editing. **Wenceslao Moreda:** Project administration, Conceptualization, Writing - review & editing. **José Manuel Martínez-Rivas:** Project administration, Conceptualization, Writing - review & editing. **Enrico Valli:** Project administration, Resources, Writing - review & editing. **Alessandra Bendini:** Conceptualization, Funding acquisition, Project administration, Resources, Writing - review & editing. **Tullia Gallina Toschi:** Conceptualization, Funding acquisition, Project administration, Resources, Writing - review & editing. **Alba Tres:** Investigation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing - review & editing. **Stefania Vichi:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Validation, Writing - review & editing, Supervision.

References

Aparicio, R., Morales, M. T., Aparicio-Ruiz, R., Tena, N., & García-González, D. L. (2013). Authenticity of olive oil: Mapping and comparing official methods and promising

alternatives. *Food Research International, 54*, 2025–2038. https://doi.org/10.1016/j.foodres.2013.07.039

Arlorio, M., Coisson, J. D., Bordiga, M., Travaglia, F., Garino, C., Zuidmeer, L., Van Ree, R., Giuffrida, M. G., Conti, A., & Martelli, A. (2010). Olive oil adulterated with hazelnut oils: Simulation to identify possible risks to allergic consumers. *Food Additives & Contaminants: Part A, 27*, 11–18. https://doi.org/10.1080/02652030903225799

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*, 832–857. 10.0.4.56/10408398.2016.1225666.

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. https://doi.org/10.1016/j.tifs.2018.08.007

Christopoulou, E., Lazaraki, M., Komaitis, M., & Kaselimis, K. (2004). Effectiveness of determinations of fatty acids and triglycerides for the detection of adulteration of olive oils with vegetable oils. *Food Chemistry, 84*, 463–474. https://doi.org/10.1016/S0308-8146(03)00273-5

Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., Lacoste, F., Brereton, P., García-González, D. L., Moreda, W., & Gallina Toschi, T. (2020). 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, 105*, 483–493. https://doi.org/10.1016/j.tifs.2019.02.025

Cozzolino, R., & De Giulio, B. (2011). Application of ESI and MALDI-TOF MS for triacylglycerols analysis in edible oils. *European Journal of Lipid Science and Technology, 113*, 160–167. https://doi.org/10.1002/ejlt.201000429

Downey, G., Mcintyre, P., & Davies, A. N. (2002). Detecting and quantifying sunflower oil adulteration in extra virgin olive oils from the eastern Mediterranean by visible and near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry, 50*, 5520–5525. https://doi.org/10.1021/jf0257188

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. https://doi.org/10.1016/j.foodres.2013.10.015

European Commission Regulation. (1991). 2568/91 on the characteristics of olive oil and olive residue oil and on the relevant methods of analysis, and subsequent amendments. *Official Journal of European Community, L248*, 1–102. July 11.

European Commission report. (2018). *Study on the implementation of conformity checks in the olive oil sector throughout the European Union*. Contract AGRI-2018-0485. available at: https://op.europa.eu/en/publication-detail.

Everstine, K., Spink, J., & Kennedy, S. (2013). Economically motivated adulteration (EMA) of food: Common characteristics of EMA incidents. *Journal of Food Protection, 76*, 723–735. https://doi.org/10.4315/0362-028X.JFP-12-399

Gao, B., Luo, Y., Lu, W., Liu, J., Zhang, Y., & Yu, L. (2017). Triacylglycerol compositions of sunflower, corn and soybean oils examined with supercritical CO_2 ultra-performance convergence chromatography combined with quadrupole time-of-flight mass spectrometry. *Food Chemistry, 218*, 569–574. https://doi.org/10.1016/j.foodchem.2016.09.099

García-González, D., Vieira, M., Tena, N., & Aparicio, R. (2007). Evaluation of the methods based on triglycerides and sterols for the detection of hazelnut oil in olive oil. *Grasas y Aceites, 58*, 344–350. https://doi.org/10.3989/gya.2007.v58.i4.445

Gelpí, E., Posada de la Paz, M., Terracini, B., Abaitua, I., Gómez de la Cámara, A., Kilbourne, E. M., Lahoz, C., Nemery, B., Philen, R. M., Soldevilla, L., & Tarkowski, S. (2002). The Spanish toxic oil syndrome 20 years after its onset: A multidisciplinary review of scientific knowledge. *Environmental Health Perspectives, 110*, 457–464. https://doi.org/10.1289/ehp.110-1240833

Georgouli, K., Martinez Del Rincon, J., & Koidis, A. (2017). Continuous statistical modelling for rapid detection of adulteration of extra virgin olive oil using mid infrared and Raman spectroscopic data. *Food Chemistry, 217*, 735–742. https://doi.org/10.1016/j.foodchem.2016.09.011

Gómez-Ariza, J. L., Arias-Borrego, A., & García-Barrera, T. (2006). Use of atmospheric pressure photoionization quadrupole time-of-flight mass spectrometry for fast olive oil fingerprinting. *Rapid Communications in Mass Spectrometry, 20*, 1181–1186. https://doi.org/10.1002/rcm.2432

Hu, C., Duan, Q., & Han, X. (2019). Strategies to improve/eliminate the limitations in shotgun lipidomics. *Proteomics*, 1900070. https://doi.org/10.1002/pmic.201900070

International Olive Council IOC/T.15/Doc No 3/Rev 14. (2019). *Trade standard applying to olive oils and olive pomace oils*. November.

International Olive Council IOC/T.20/Doc No 20/Rev 4. (2017). *Determination of the difference between actual and theoretical content of triacylglycerols with ECN 42*.

International Olive Council IOC/T.20/Doc No 25/Rev 2. (2018a). *Evaluation of the coherence of TAG composition with the fatty acid composition*.

International Olive Council IOC/T.20/Doc No 26/Rev 4. (2018b). *Determination of the sterol composition and content and alcoholic compounds by capillary gas chromatography*.

Lísa, M., & Holčapek, M. (2008). Triacylglycerols profiling in plant oils important in food industry, dietetics and cosmetics using high-performance liquid chromatography–atmospheric pressure chemical ionization mass spectrometry. *Journal of Chromatography A, 1198–1199*, 115–130. https://doi.org/10.1016/j.chroma.2008.05.037

Longo, E., Morozova, K., Yener, S., Boselli, E., Biasioli, F., & Scampicchio, M. (2019). Direct flow injection profiling of acylglycerols from food products using isopropanol as solvent. *Journal of Mass Spectrometry, 54*, 412–421. https://doi.org/10.1002/jms.4346

Magnusson, B., & Ornemark, U. (2014). *Eurachem Guide: The fitness for purpose of analytical methods – a laboratory guide to method validation and related topics* (2nd ed.), ISBN 978-91-87461-59-0 https://www.eurachem.org/images/stories/Guides/pdf/MV_guide_2nd_ed_EN.pdf.

Mailer, R. J., & Gafner, S. (2020). Adulteration of olive (*Olea europaea*) oil. *Botanical Adulterants Prevention Bulletin, 19*, 1–14. http://cms.herbalgram.org/BAP/BAM/issue19.html.

McGrath, T. F., Haughey, S. A., Patterson, J., Fauhl-Hassek, C., Donarski, J., Alewijn, M., van Ruth, S., & Elliott, C. T. (2018). What are the scientific challenges in moving from targeted to non-targeted methods for food fraud testing and how can they be addressed? –Spectroscopy case study. *Trends in Food Science & Technology, 76*, 38–55. https://doi.org/10.1016/j.tifs.2018.04.001

Meenu, M., Cai, Q., & Xu, B. (2019). A critical review on analytical techniques to detect adulteration of extra virgin olive oil. *Trends in Food Science & Technology, 91*, 391–408. https://doi.org/10.1016/j.tifs.2019.07.045

Quintanilla-Casas, B., Bustamante, J., Guardiola, F., García-González, D. L., Barbieri, S., Bendini, A., Gallina Toschi, T., Vichi, S., & Tres, A. (2020). Virgin olive oil volatile fingerprint and chemometrics: Towards an instrumental screening tool to grade the sensory quality. *Lebensmittel-Wissenschaft und -Technologie- Food Science and Technology, 121*, 108936. https://doi.org/10.1016/j.lwt.2019.108936

Quintanilla-Casas, B., Vichi, S., Bustamante, J., Simon, M., Guardiola, F., & Tres, A. (2018). Shotgun triacylglycerol profiling and chemometrics: Towards a fast authentication of olive oil adulteration. *Annals of Nutrition & Metabolism, 73*, 77. https://doi.org/10.1159/000490752

Rusilowicz, M., Dickinson, M., Charlton, A., O'Keefe, S., & Wilson, J. (2016). A batch correction method for liquid chromatography–mass spectrometry data that does not depend on quality control samples. *Metabolomics, 12*, 56. https://doi.org/10.1007/s11306-016-0972-2

Sánchez, J., & Harwood, J. L. (2002). Biosynthesis of triacylglycerols and volatiles in olives. *European Journal Of Lipid Science And Technology, 104*, 564–573. https://doi.org/10.1002/1438-9312(200210)104:9/10%3C564::AID-EJLT564%3E3.0.CO;2-5

Sun, X., Lin, W., Li, X., Shen, Q., & Luo, H. (2015). Detection and quantification of extra virgin olive oil adulteration with edible oils by FT-IR spectroscopy and chemometrics. *Analytical Methods, 7*, 3939–3945. https://doi.org/10.1039/C5AY00472A

Tres, A., van der Veer, G., & van Ruth, S. M. (2013). Vegetable oils. In M. de la Guardia, & A. Gonzalvez (Eds.), *Food protected designations of origin. Methodologies and applications* (pp. 543–573). Oxford: Elsevier B.V, ISBN 9780444595621.

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. https://doi.org/10.1016/j.aca.2009.04.043

Valcárcel, M., Cárdenas, S., Barceló, D., Buydens, L., Heydorn, K., Karlberg, B., Klemm, K., Lendl, B., Milman, B., Neidhart, B., Ríos, A., Stephany, R., Townshend, A., & Zschunke, A. (2002). *Metrology of qualitative chemical analysis*. Luxembourg: *BCR Information*, European Commission, ISBN 92-894-5194-7.

Vichi, S., Cortés-Francisco, N., & Caixach, J. (2012). Ultrahigh resolution mass spectrometry and accurate mass measurements for high-throughput food lipids profiling. *Journal of Mass Spectrometry, 47*, 1177–1190. https://doi.org/10.1002/jms.3017

Vichi, S., Pizzale, L., & Conte, L. S. (2007). Stereospecific distribution of fatty acids in triacylglycerols of olive oils. *European Journal of Lipid Science and Technology, 109*, 72–78. https://doi.org/10.1002/ejlt.200600199

Xiang, F., Hendrikson, C. L., & Marshall, A. G. (2012). High resolution mass spectrometry. *Analytical Chemistry, 84*, 708–719. https://doi.org/10.1021/ac203191t

Yan, J., Erasmus, S. W., Aguilera Toro, M., Huang, H., & van Ruth, S. M. (2020). Food fraud: Assessing fraud vulnerability in the extra virgin olive oil supply chain. *Food Control, 111*, 107081. https://doi.org/10.1016/j.foodcont.2019.107081

Yang, Y., Duarte Ferro, M., Cavaco, I., & Liang, Y. (2013). Detection and identification of extra virgin olive oil adulteration by GC-MS combined with chemometrics. *Journal of Agricultural and Food Chemistry, 61*, 3693–3702. https://doi.org/10.1021/jf4000538

Zhang, X. F., Zou, M. Q., Qi, X. H., Liu, F., Zhang, C., & Yin, F. (2011). Quantitative detection of adulterated olive oil by Raman spectroscopy and chemometrics. *Journal of Raman Spectroscopy, 42*, 1784–1788. https://doi.org/10.1002/jrs.2933

Zou, M.Q., Zhang, X.F., Qi, X.H., Ma, H.L., Dong, Y., Liu, C.W., Guo, X., & Wang, H. Rapid authentication of olive oil adulteration by Raman Spectrometry. Journal of Agricultural and Food Chemistry, 57, 6001–6006.10.1021/jf900217s.