

## Profiling versus fingerprinting analysis of sesquiterpene hydrocarbons for the geographical authentication of extra virgin olive oils

Beatriz Quintanilla-Casas<sup>a,b</sup>, Sofia Bertin<sup>a</sup>, Kerstin Leik<sup>a</sup>, Julen Bustamante<sup>a,b</sup>, Francesc Guardiola<sup>a,b</sup>, Enrico Valli<sup>c</sup>, Alessandra Bendini<sup>c</sup>, Tullia Gallina Toschi<sup>c</sup>, Alba Tres<sup>a,b,\*</sup>, Stefania Vichi<sup>a,b</sup>

<sup>a</sup> Departament de Nutrició, Ciències de l'Alimentació i Gastronomia, Campus de l'Alimentació de Torribera, Universitat de Barcelona (UB), Santa Coloma de Gramenet, Spain

<sup>b</sup> Institut de Recerca en Nutrició i Seguretat Alimentària (INSA-UB), Universitat de Barcelona (UB), Santa Coloma de Gramenet, Spain

<sup>c</sup> Department of Agricultural and Food Sciences, Alma Mater Studiorum – Università di Bologna, Bologna, Italy

### ARTICLE INFO

#### Keywords:

Fingerprinting  
Geographical origin  
Virgin olive oil  
Sesquiterpene  
Food authentication

### ABSTRACT

The verification of the geographical origin of extra virgin (EVOO) and virgin olive oil (VOO) is crucial to protect consumers from misleading information. Despite the large number of studies performed, specific markers are still not available. The present study aims to evaluate sesquiterpene hydrocarbons (SHs) as markers of EVOO geographical origin and to compare the discrimination efficiency of targeted profiling and fingerprinting approaches. A prospective study was carried out on 82 EVOOs from seven countries, analyzed by Headspace Solid Phase Microextraction-Gas Chromatography-Mass Spectrometry (HS-SPME-GC-MS). Classification models were developed by Partial Least Square-Discriminant Analysis (PLS-DA) and internally validated (leave 10%-out cross-validation). The percentage of correct classification was higher for the fingerprinting (100%) than for the profiling approach (45.5–100%). These results confirm the suitability of SHs as EVOO geographical markers and establish the fingerprinting as the most efficient approach for the treatment of SH analytical data with this purpose up to date.

### 1. Introduction

As reported by EU Parliament (Parliament Resolution EU No 2013/2091 (IN1)), the cases of food fraud reduce the confidence of consumers in the food chain, compromising its global image and causing a negative influence in the food sector. EU Regulation No 29/2012 states as mandatory the country of origin in labeling extra virgin olive oil (EVOO) and virgin olive oil (VOO) to inform the consumer regarding their geographical origin. The geographical origin reported in the label refers to i) a single EU Member State or third country, ii) oil blends of European Union or non-European Union origin, or iii) certain protected designations of origin or protected geographical indications according to EU Regulation (Regulation (EU) No 1151/2012). The verification of conformity of the label-declared geographical origin of EVOO and VOO plays a key role, not only to protect consumers from misleading information and restore their confidence in the product, but also to detect and prevent fraudulent practices and increase the competitiveness of the sector. A large number of studies have been performed trying to face

up the EVOO geographical authentication. They have been based on several chemical compounds such as triacylglycerols, fatty acids, phenolic compounds, pigments, sterols and volatile compounds, by applying different analytical techniques as well as chemometric approaches (Bajoub, Bendini, Fernández-Gutiérrez, & Carrasco-Pancorbo, 2018; Conte et al., 2019). However, it is known that the levels of some of these analytes change along EVOO shelf life (i.e. phenols and pigments) and others are related to olive oil quality/purity (i.e. volatile compounds), meaning that they can be affected by storage and processing factors (García-González & Aparicio, 2010). Other studies focused on the olive oil chemical fingerprint by stable Isotope Ratio Mass Spectrometry and Nuclear Magnetic Resonance (Alonso-Salces et al., 2015; Camin et al., 2016). Even though their results were promising by combining data from both analyses, they require smart instrumentation that is not often affordable for common control laboratories. For these reasons, we can state that there is room still for improvement in the development of EVOO and VOO geographical markers.

To develop efficient tools for the geographical authentication of

\* Corresponding author at: Departament de Nutrició, Ciències de l'Alimentació i Gastronomia, XaRTA, INSA-UB Campus De l'Alimentació Torribera, Av Prat de la Riba, 171, 08921 Santa Coloma de Gramenet, Spain

E-mail address: [atres@ub.edu](mailto:atres@ub.edu) (A. Tres).

<https://doi.org/10.1016/j.foodchem.2019.125556>

Received 1 March 2019; Received in revised form 13 September 2019; Accepted 16 September 2019

Available online 30 September 2019

0308-8146/ © 2019 Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

EVOO and VOO, it is necessary to identify the most robust markers and analytical approaches. To be reliable, geographical markers of food products should depend mainly climatic and agronomic factors linked to a specific area, while keeping the influence of other factors to a minimal degree (Vichi, Tres, Quintanilla-Casas, Bustamante, & Guardiola, 2018). Additionally, the determination of such markers for routine analysis should imply low cost, short times and automatable procedures.

Recent studies reveal that sesquiterpene hydrocarbons (SHs) might act as valid markers to address the genetic and geographical origin of EVOO and VOO (Bortolomeazzi, Berno, Pizzale, & Conte, 2001; Zunin, Boggia, Salvadeo, & Evangelisti, 2005; Vichi, Guadayol, Caixach, Lopez-Tamames & Buxaderas, 2006; Vichi, Lazzez, Grati-Kamoun, Lopez-Tamames & Buxaderas, 2010; Damascelli & Palmisano, 2013). SHs are semi-volatile plant metabolites comprising an extremely wide number of compounds in nature. In EVOO and VOO, SH composition is highly dependent on the olive trees' cultivar and growing area, and scarcely influenced by other factors such as oil extraction conditions and storage (Vichi et al., 2018). The effect of agronomic and pedoclimatic conditions on olive oil SHs has been proven by the fact that significant differences in the SH composition have been found between samples from the same cultivar produced in different geographical areas (Ben Temime, Campeol, Cioni, Daoud, & Zarrouk, 2006; Youssef et al., 2011; Vichi, Lazzez, Méndez, & Caixach, 2015) and also between EVOOs from different cultivars grown in the same parcel did (Vichi, Lazzez, Grati-Kamoun, López-Tamames, & Buxaderas, 2010). However, the suitability of SHs as geographical markers in a realistic scenario should be tested with olive oils from different geographical areas under the usual production practices, implying the use of monovarietal oils from typical olive cultivars as well as their usual market blends, as addressed by some studies (Zunin et al., 2005; Damascelli & Palmisano, 2013).

In the last years, the analysis of SHs has evolved from time-consuming methods (Bortolomeazzi et al., 2001) to simpler methods based on the analysis of the volatile fraction such as solid phase microextraction (SPME) (Vichi, Guadayol, Caixach, López-Tamames, & Buxaderas, 2006), allowing further studies of these compounds in EVOOs and VOOs and considering their use as possible authenticity markers.

Concerning the analytical approach, the traditional way to assess these semi-volatile compounds is based on a target-type analysis to identify and determine the SH profile of samples. This approach involves a peak identification step, which presents some difficulties because the mass spectra of these analytes contain the same specific ions in different proportions, which causes that many SHs have not been identified yet. Consequently, under a targeted profiling approach, as defined by Ballin and Laursen (2019), part of the information is ignored.

Nowadays, the emerging strategy in food authentication consists in finding specific patterns in highly dimensional analytical data, known as fingerprints, which might be based directly in raw analytical signals such as a chromatogram (Berrueta, Alonso-Salces & Heberger, 2007; Bosque-Sendra, Cuadros-Rodríguez, Ruiz-Samblas & de la Mata, 2012; Melucci et al., 2016; Ballin & Laursen, 2019). When these distinctive patterns are specific to a given food category, such as a particular geographical origin, then can be used to verify its authenticity. Under the fingerprinting approach, since peak identification and quantitation are not necessary, some of the drawbacks related with the targeted profiling approach mentioned above are overcome. Besides, since the full analytical data is used, more information is considered and misclassifications are revealed easier.

With the aim to evaluate the suitability of SHs as geographical markers for EVOO and VOO under real production conditions we carried out a prospective study on EVOOs from seven different geographical origins, comprising monovarietal oils as well as market blends of oils from various cultivars typically produced in these origins.

The SHs were determined by HS-SPME and gas chromatography-mass spectrometry (GC-MS) and data was evaluated under targeted (profiling) and non-targeted (fingerprinting) analytical approaches with the aim to compare their discrimination-efficiency in the verification of the geographical origin.

## 2. Material and methods

### 2.1. Sampling

A total of 82 authentic and traceable samples, declared as EVOO by the suppliers, were obtained in the framework of OLEUM project (EC H2020 Programme 2014–2020) from seven different EU and non-EU countries: Croatia (HRV) (n = 11); Slovenia (SVN) (n = 8); Spain (ESP) (n = 17); Italy (ITA) (n = 15); Greece (GRC) (n = 6); Morocco (MAR) (n = 15) and Turkey (TUR) (n = 10). With the aim of reflecting the real production scenario, EVOO samples in this prospective study were obtained under usual production practices for commercial purposes, and thus consisted of both monovarietal oils as well as market blends of olive cultivars typical of each geographical origin (Supplementary material, Table S1).

### 2.2. Headspace-solid phase microextraction (HS-SPME)

SHs present in EVOO were analyzed using a Triplus autosampler (Thermo Fischer Scientific, Bremen, Germany) at the conditions reported by Vichi et al. (2006). Shortly, 2 g of oil was weighed into a 10 mL vial fitted with a silicone septum and kept at 70 °C under agitation. After 10 min of sample conditioning, a divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fiber (2 cm length, 50/30 µm film thickness) was exposed during 60 min to the sample headspace and then desorbed for 10 min in the GC injection port (260 °C). The thermal stability of SHs at these SPME conditions was previously verified (Vichi et al., 2006). During the desorption step, the injector was maintained in split-less mode during 5 min. Oil samples were previously spiked with a standard solution of indene to a final concentration of 10 µg/kg.

### 2.3. Gas chromatography-mass spectrometry (GC-MS)

Separation and detection of volatile compounds was performed by GC coupled to an ion trap mass selective spectrometry using a ThermoFinnigan Trace GC equipped with an ITQ MS (Thermo Fisher Scientific, Waltham, MA) using helium as carrier gas at a constant flow of 1.3 mL min<sup>-1</sup>. Analytes were separated on a Supelcowax-10 (Supelco, Bellefonte, PA) 60 m × 0.25 mm i.d., 0.25 µm film thickness. Column temperature was held at 40 °C for 3 min, increased to 75 at 4 °C min<sup>-1</sup>, then to 200 at 8 °C min<sup>-1</sup> and to 260 °C at 15 °C min<sup>-1</sup>, holding the last temperature for 2 min. The temperatures of the ion source and the transfer line were 200 and 275 °C, respectively. Mass spectra were recorded with a scan event time of 0.37 s; electron energy was 70 eV. Acquisition in the complete scanning mode (SCAN) was in the range m/z 40–300, to allow the identification of compounds in EVOO samples.

### 2.4. Data processing

#### 2.4.1. Profiling approach

Compounds were identified by comparing their mass spectra and retention times to those of the standard compounds, or the ones available in the NIST 2.0 mass spectrum library and in the literature. Non-isothermal linear retention indices (LRI), using the definition of Van den Dool and Kratz (1963), were calculated and compared with those available in the literature (Supplementary material, Table S2).

Several common ions, only differing in their proportions, characterize the mass spectra of SHs. Therefore, a quantitative assessment of

SHs was carried out in Extracted Ion Chromatogram (EIC) by selecting the following ions:  $m/z$  69, 93, 107, 119, 135, 157, 159, 161, 189, 200, 202 and 204. The selection of quantification ions was done according to Vichi et al. (2006) and the confirmation ions were the molecular ions  $m/z$  204, 202 or 200. Quantification was carried out by internal standard, considering a response factor equal to 1, and expressed as  $\mu\text{g}$  equivalents of IS/kg of oil (Supplementary material, Table S2). According to Vichi et al. (2006), both SH and monoterpenes may be taken into consideration to be studied as genetic or geographic markers of virgin olive oil origin. However, monoterpene content suffers higher variability due to their low-boiling point compared with sesquiterpenes, introducing variability not related to the origin into the model (data not shown). Since models developed with sesquiterpene data were successful, it was not considered necessary to also include monoterpenes.

#### 2.4.2. Fingerprinting approach

The EIC of specific SH ions ( $m/z$  93, 107, 119, 135, 157, 159, 161, 189 and 204) were obtained from the Total Ion Current (TIC). The intensities of scans comprised from 18th to the 30th minute (2467 scans) were considered for each ion ( $2467 \text{ scans} \times 9 \text{ ions} = 22,203$  variables per sample). To solve the retention time shifting, for each selected ion the EICs of the 82 samples were aligned by *icoshift* algorithm in Matlab® (Tomasí, Savorani, & Engelsen, 2011). Once aligned, the 9 matrices of the 9 aligned EICs were concatenated conforming a two-way unfolded matrix ( $82 \text{ samples} \times 22,203$  variables).

#### 2.4.3. Chemometrics

Univariate statistical analysis for the profiling approach was carried out with SPSS software v25© (IBM Corp., NY USA). A one-way ANOVA was applied: F test and Tukey multiple comparisons test were used when variances were equal between groups. Instead, Welch test and Games-Howell multiple comparisons test were applied when groups presented unequal variances.  $P < 0.05$  was considered significant.

Multivariate analysis of profiling and fingerprinting approaches was performed with SIMCA software v13.0© (Umetrics AB, Sweden). After data pre-processing (log10, mean centering and scaling for the target data; first derivative, log10, mean centering and scaling for the fingerprint data), a Principal Component Analysis (PCA) was developed for both profile and fingerprint data to explore the natural grouping of samples and detect potential outliers (according to Hotelling's  $T^2$  range and distance to the model parameters). Partial Least Square-Discriminant Analysis (PLS-DA) classification models were built with data obtained by profiling (34 variables) and fingerprinting analysis (22,203 variables) to verify the geographical origin of EVOO samples coming from 7 different countries: HRV, SVN, ESP, ITA, GRC, MAR and TUR. PLS-DA is a supervised discriminant technique based on finding the maximum correlation between the data (the SH profile or the SH fingerprint) and each of the categories (each of the seven countries of origin). By doing this, PLS-DA finds the most different features between categories while minimizing those variables not related with a given category. The models were internally validated by leave 10% out cross-validation and the number of latent variables of PLS-DA models were selected according to the lowest RMSEcv value. Model successfulness was evaluated by their prediction power ( $Q^2$  value) and the % of correct classifications. Random behavior and model over-fitting were assessed through the ANOVA on the cross-validated predictive residuals ( $p$ -value) and the permutation test, in which the prediction power ( $Q^2$  value) of 20 models developed after randomizing sample categories (countries) was compared with that of the original model.

### 3. Results and discussion

#### 3.1. Profiling approach

The chromatograms obtained extracting typical SH ions from the TIC, showed an extremely complex fraction (Fig. 1). As commented

above, the identification of SHs is a challenging task because they present very similar mass spectra. Despite this fact, a total of 34 peaks were included in the SH profile; 23 of them were assigned to previously reported SH (Bortolomeazzi et al., 2001; Vichi et al., 2006) while the remaining ones were not found in literature but could be related to SH compounds based on their mass spectra. The quantitative data of these SHs, expressed as  $\mu\text{g}$  equivalents of IS/kg of oil, were used to perform the univariate statistical analysis by a one-way ANOVA (Supplementary material, Table S2). Although some differences were found for some SHs, the high intra-class and inter-class variability caused that this univariate approach was not successful in distinguishing the various origins and that specific markers of origin could not be directly found.

Multivariate techniques under a profiling and a fingerprinting approach were assayed in order to better explore the differences between samples from different countries. In the profiling approach, after data pre-treatment and PCA exploration, no outliers were detected. Therefore, the PLS-DA classification model for the targeted data was developed with all the samples ( $n = 82$ ) (Fig. 2a). After various pre-processing techniques assayed, the model on the log10, mean centering and data scaling to unit variance was the most successful, and with 8 latent variables it achieved the lowest global RMSEcv for most of the categories.

Table 1 shows the classification results obtained from cross-validation by leave 10%-out and the respective RMSEcv values for each class. The model rendered good percentages of correct classification for samples from certain geographical origin, such as SVN (100%), TUR (100%) and MAR (93.3%). However, in the case of oils from the rest of the countries, it generated some misclassifications, particularly in the case of HRV (45.5%), resulting in a non-satisfactory model. This agrees with the fact that the global  $Q^2$  score (0.351) was low, which indicates a low prediction power of the present classification model. On the other hand, the ANOVA  $p$ -value (0.013) indicates that the model is significant and thus, that the classification is not at random. Also, the  $Q^2$  values of the permutation test for each category were below 0 indicating the absence of a random classification and of model overfitting.

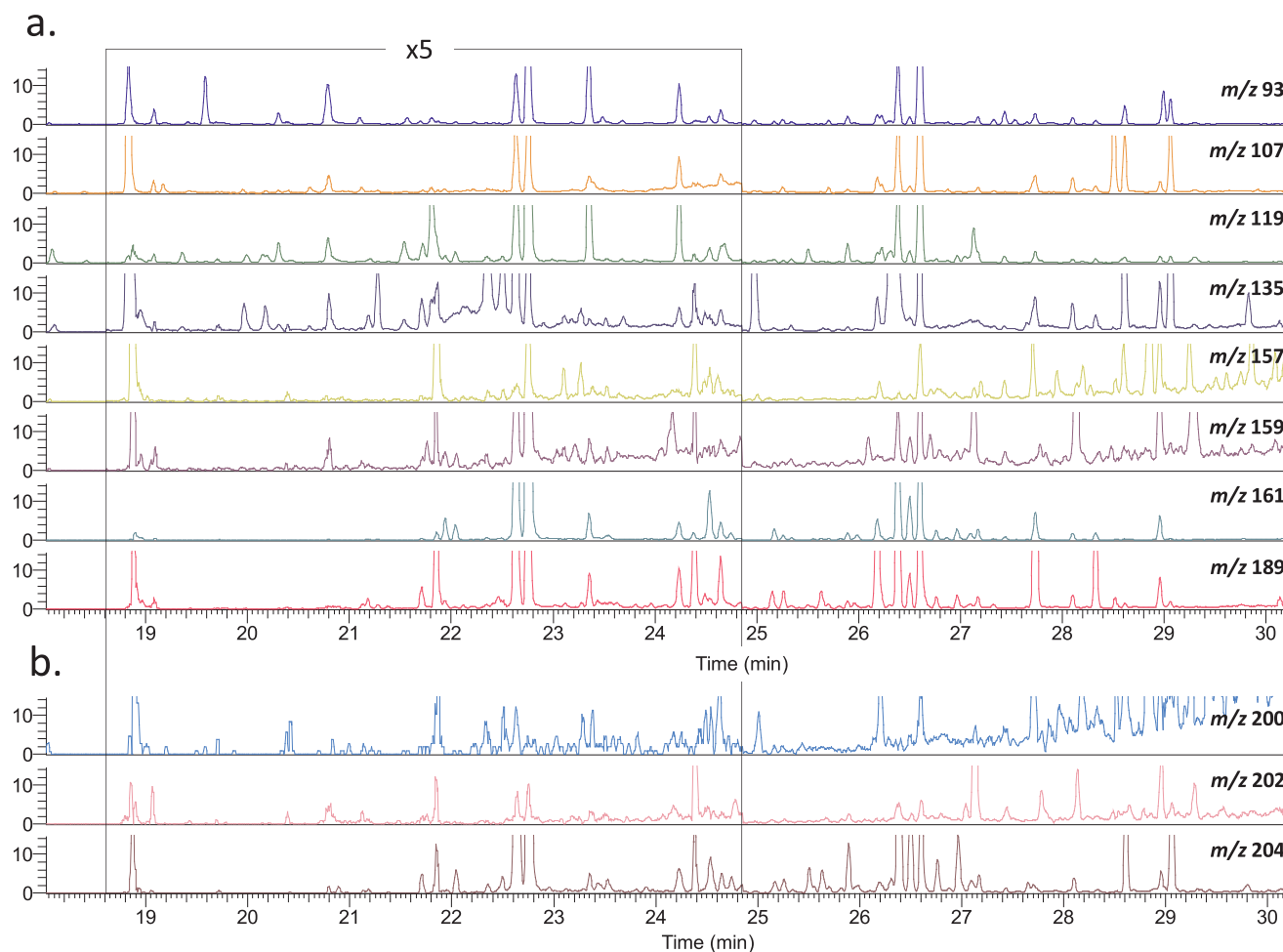
As aforementioned, the target analysis is limited to the number of compounds that can be identified or tentatively identified based on their mass spectrum and linear retention index (LRI). However, the chromatograms obtained by extracting typical terpene fragment ions (Fig. 1) show that the SH fraction is much more complex, and that many SHs might have not been considered, meaning that the profiling approach might have missed part of the information of the SHs profile.

#### 3.2. Fingerprinting approach

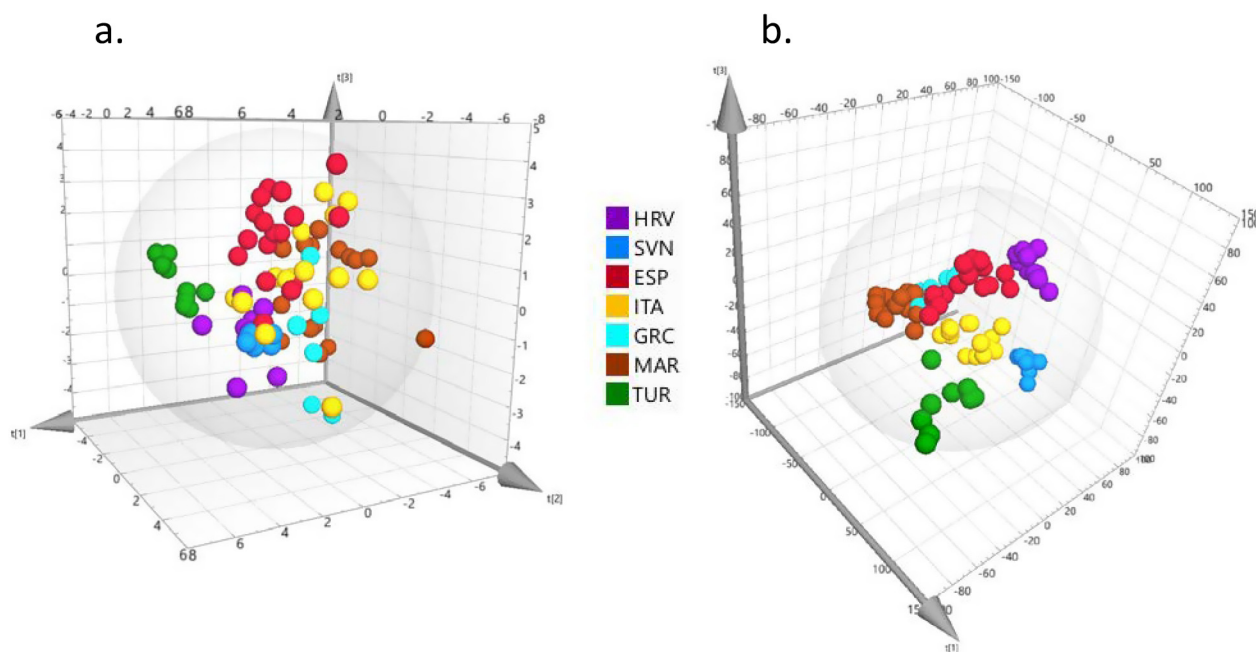
With the aim to extract exhaustive information from the SH fraction in EVOO, a non-targeted fingerprinting analysis was evaluated. All data points obtained from the selected region of each SH specific EIC were used as variables so that every signal related to SH was taken into account by the model.

The two-way unfolded matrix obtained ( $82 \text{ samples} \times 22,203$  variables) was subjected to data pre-processing and PCA exploration, in which any outlier was detected. Then, a PLS-DA classification model was performed. The model leading to the lowest RMSEcv used 6 latent variables (Fig. 2b). In this case, the sample grouping according to the origin was drastically improved compared to the profiling model. A 100% of correct classification (by leave 10%-out cross-validation) was obtained for each of the 7 countries of origin (Table 1). ANOVA  $p$ -value ( $1.6 \times 10^{-18}$ ) indicated that the model was significant and excluded a random classification. Results from the permutation test were very satisfactory, with  $Q^2$  values below 0.2, suggesting that the optimized classification model was not over-fitted.

The successful classification results obtained under this approach agreed with the fact that the sub-models for each geographical origin found patterns of the SH fingerprint that were character of each of them, as revealed by the regression coefficient plots (Supplementary



**Fig. 1.** Extracted ion chromatograms of sesquiterpene hydrocarbons: a) Quantification ions; b) Confirmation ions (molecular ions), obtained by analysing an extra virgin olive oil from Spain by HS-SPME-GC-MS.



**Fig. 2.** Score scatter plot (first 3 latent variables) of classification models (PLS-DA) developed by country of origin, based on extra virgin olive oil sesquiterpene data by applying a) profiling approach (34 variables); b) fingerprint approach (22,203 variables). HRV: Croatia, SVN: Slovenia, ESP: Spain, ITA: Italy, GRC: Greece, MAR: Morocco and TUR: Turkey.



**Table 1**

Misclassification results of classification models (PLS-DA) developed with Extra Virgin Olive Oil sesquiterpene profile (34 variables; log10, mean centering and scaling to unit variance; 8 latent variables) and extra virgin olive oil sesquiterpene fingerprint (22,203 variables; 1st derivative, log10, mean centering and scaling to unit variance; 6 latent variables), cross-validated by leave 10%-out.

	Members	Correct classification	HRV	SVN	ESP	ITA	GRC	MAR	TUR	No class (YPred < 0.5)	RMSEcv
<b>Profiling<sup>a</sup></b>											
HRV	11	45.5%	5	0	0	0	0	0	0	6	0.28
SVN	8	100%	0	8	0	0	0	0	0	0	0.22
ESP	17	58.8%	0	0	10	0	0	0	0	7	0.38
ITA	15	53.3%	0	0	1	8	0	0	0	5	0.39
GRC	6	50%	0	0	0	0	3	0	0	3	0.25
MAR	15	93.3%	0	0	0	0	0	14	0	1	0.26
TUR	10	100%	0	0	0	0	0	0	10	0	0.17
Total	82	73.7%	5	8	11	8	3	14	10	22	
<b>Fingerprinting<sup>b</sup></b>											
HRV	11	100%	11	0	0	0	0	0	0	0	0.25
SVN	8	100%	0	8	0	0	0	0	0	0	0.23
ESP	17	100%	0	0	17	0	0	0	0	0	0.32
ITA	15	100%	0	0	0	15	0	0	0	0	0.33
GRC	6	100%	0	0	0	0	6	0	0	0	0.23
MAR	15	100%	0	0	0	0	0	15	0	0	0.26
TUR	10	100%	0	0	0	0	0	0	10	0	0.19
Total	82	100%	11	8	17	15	6	15	10	0	

Abbreviations used: HRV: Croatia, SVN: Slovenia, ESP: Spain, ITA: Italy, GRC: Greece, MAR: Morocco; TUR: Turkey; RMSEcv: Root Mean Square Error of cross-validation.

<sup>a</sup> Profiling PLS-DA model:  $Q^2$ : 0.351; ANOVA p-value: 0.013;

<sup>b</sup> Fingerprinting PLS-DA model  $Q^2$ : 0.561; ANOVA p-value:  $1.6 \times 10^{-18}$ .

material, Fig. S1). To illustrate this, a section of EIC for  $m/z$  119 of TUR samples (Fig. 3a) is plotted against the corresponding regression coefficients of the SHs fingerprint of TUR sub-model (Fig. 3b). It reveals that some of the highest regression coefficients corresponded to peaks (i.e. peaks 7, 9, 13 and 17) that had been quantified with the  $m/z$  119 and included in the profiling model. Nevertheless, other significant regression coefficients were related with parts of the EIC that had not been included in the profiling approach, such as minor SHs or not well-resolved peaks. Thus, this explains the higher discrimination power of the fingerprinting approach compared to the profiling approach.

This prospective study sets SHs as successful EVOO geographical markers because even if various monovarietal EVOOs and EVOO cultivar blends were included for each geographical origin (Supplementary material, Table S1), the country of origin was correctly verified. This is because PLS-DA was supervised per geographical origin (country), and thus the model was addressed to focus on the SHs features more related to the geographical area, beyond the cultivar. This means that the PLS-DA model finds features that are common between samples from the same region even if they are from different cultivars. In this way, even if in some cases the same cultivar was present in different countries ['Arbequina', 'Leccino' and 'Istrska belica' (Fig. 4)], the model correctly classified the samples into the country of origin. This is especially relevant because it is known that genetic factors influence EVOO's SH profile (Guinda, Lanzon & Albi, 1996; Osorio-Bueno, Sanchez-Casas, Montano García & Gallardo González, 2005; Vichi et al., 2010). However, here, thanks to the sampling design and to the ability of PLS-DA to extract information from the fingerprint correlated with the discriminated characteristic (origin in this case), the influence of pedo-climatic aspects on SHs could be exploited.

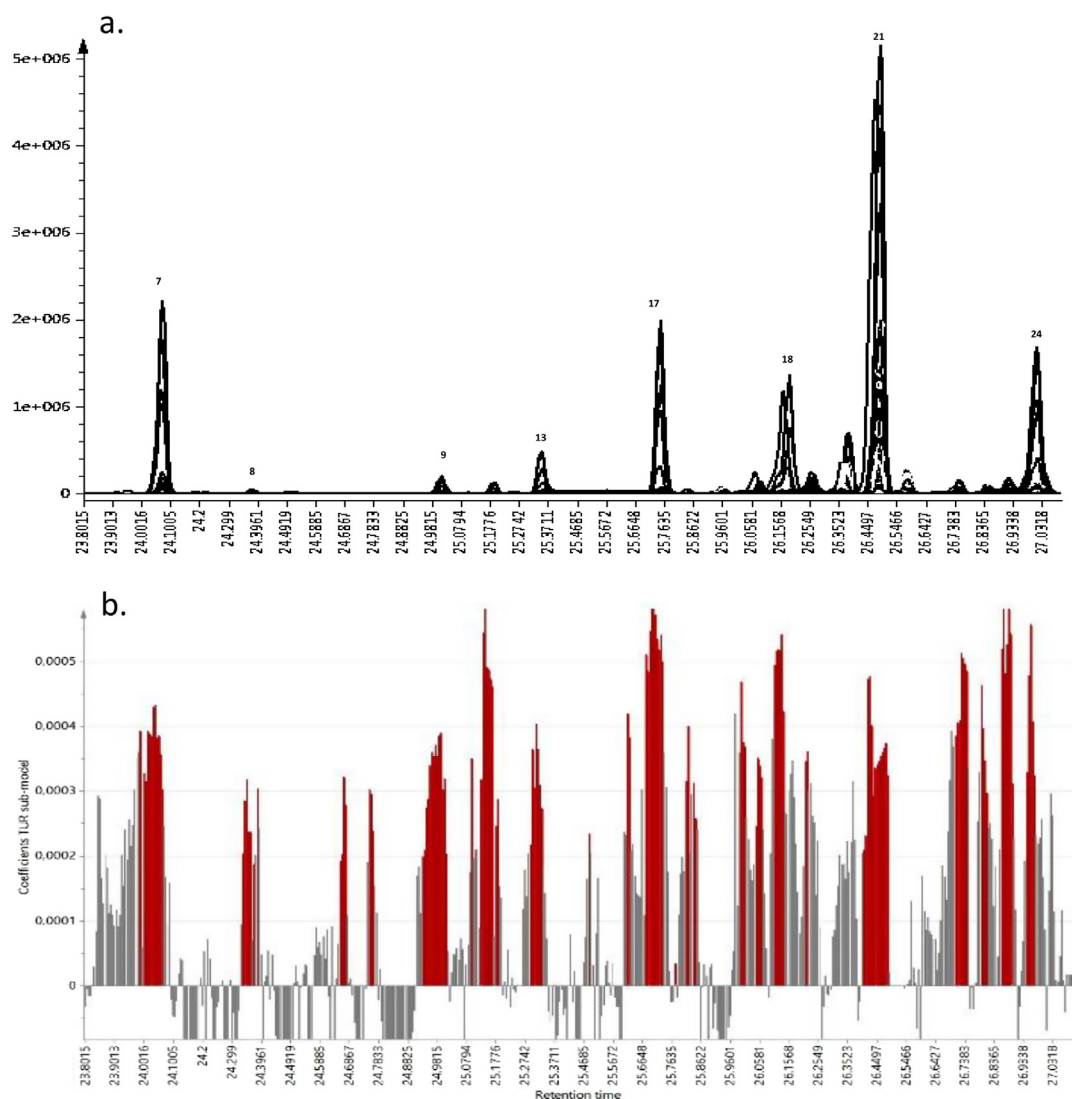
On the other hand, it is noteworthy that although the model was supervised per country of origin, it naturally grouped samples into smaller sub-regions within the same country (although the sub-region information had not been provided to the model). Fig. 5 illustrates this behavior by exemplifying the case of Italian and Turkish oils, where samples from Tuscany, Sicily and Apulia (Fig. 5a), and samples from North Aegean, Germencik and Antakya (Fig. 5b), respectively, conform independent clusters within each class. This entails that the SH fingerprint holds similar traits among samples from regions smaller than a country and sets a promising scenario for downscaling the model to

verify the geographical origin of EVOO produced in smaller regions of interest such as those from protected designations of origin (PDO) or protected geographical indications (PGI).

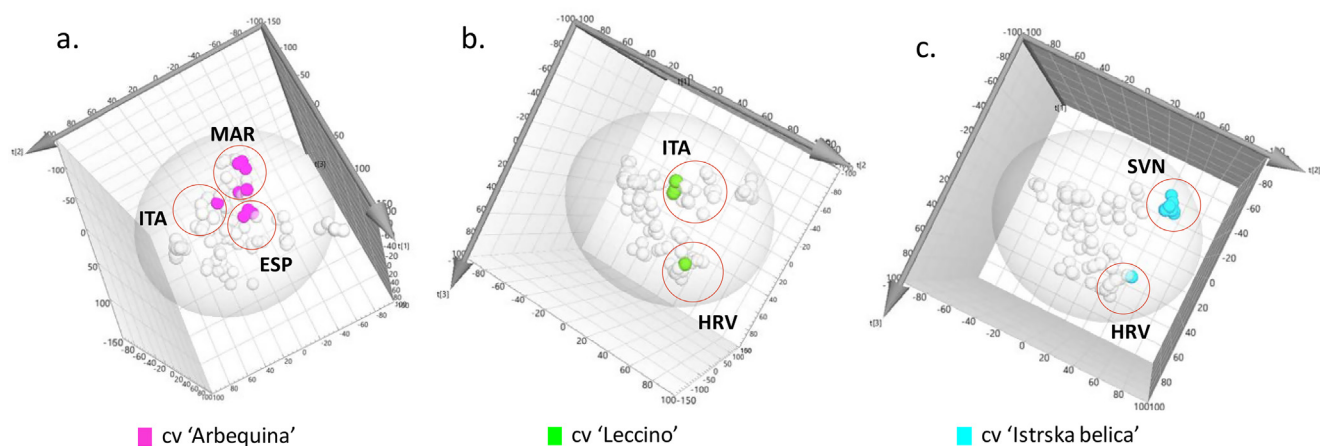
#### 4. Conclusions

This prospective study focused on the suitability of SHs as EVOO geographical markers and the evaluation of the best approach for data processing, allowed us i) to confirm that SH can be successfully used for the verification of EVOO geographical origin, ii) to state that the fingerprinting approach provided a model with a higher discrimination capacity (100% correct classification) with respect to the targeted profiling one (from 46 to 100% correct classification, depending on the country). It is remarkable that this classification rate was achieved under a real scenario of EVOO global production, which implied the use of various monovarietal and blends of oils from cultivars typically produced and marketed in each country. Also, samples from the same olive cultivar coming from different countries were correctly classified according to the geographical origin. Moreover, as the SH fingerprint holds similar traits among samples from sub-regions within a country, it sets a promising scenario for downscaling the model to smaller regions of interest such as PDO or PGI oils, as well as for challenging model robustness with samples for various harvest years. Actually, evaluating the effect of the harvest year has been shown to be crucial for some authentication models developed for EVOO verification, because as reviewed by Tres, van der Veer, and van Ruth (2013) the differences in the climatic conditions might affect EVOO composition.

Overall, we can conclude that the successfulness of the model is the result of a conjunction of factors: i) sesquiterpenes are suitable geographical markers, ii) the use of the sesquiterpene fingerprint permits to exploit all the information obtained during the analysis in contrast of the target approach, and iii) PLS-DA finds features in the sesquiterpene fingerprint that are common between samples from the same region even if they belong to different cultivars. Although we are aware that an increment of samples (with more samples from these and other origins, and from different harvest years) and external validation are still necessary to develop a more robust and elaborated model for the classification of samples according to their geographical origin, these preliminary results confirm the suitability of SHs as geographical markers



**Fig. 3.** a) Section of  $m/z$  119 EIC (from 23.8 to 27 min) of Turkish extra virgin olive oils by HS-SPME-GC-MS; b) PLS regression coefficients of the fingerprinting classification model, resulting from each data point in Figure a vs. 'Turkey' category (the highest coefficients are in red). Peaks considered in the profiling approach are: 7:  $\alpha$ -bergamotene; 8:  $\beta$ -gurjunene; 9:  $\beta$ -caryophyllene; 13: non-identified sesquiterpene; 17:  $\alpha$ -zingiberene; 18: germacrene D; 21: (E,E)- $\alpha$ -farnesene; 24:  $\delta$ -cadinene. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 4.** Score scatter plot (first 3 latent variables) of sesquiterpene fingerprint classification model (PLS-DA) supervised by geographical origin, showing how extra virgin olive oils (EVOO) from the same olive cultivar cluster according to the country of origin: a) 'Arbequina' EVOOs produced in Italy (ITA), Spain (ESP) and Morocco (MAR); b) 'Leccino' EVOOs produced in Italy (ITA) and Croatia (HRV); c) 'Istrska belica' EVOOs produced in Croatia (HRV) and Slovenia (SVN).

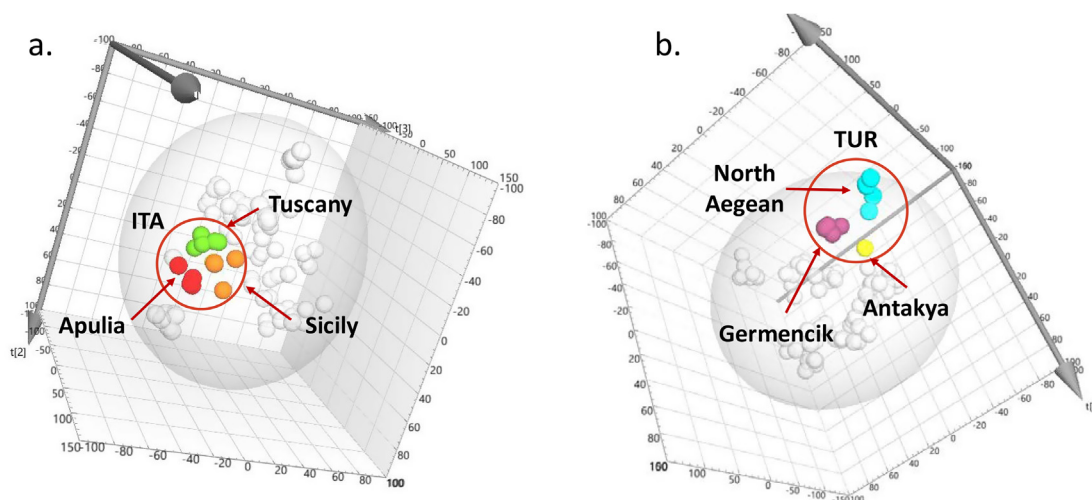


Fig. 5. Score scatter plot (first 3 latent variables) of sesquiterpene fingerprint classification model (PLS-DA) supervised by country of origin, exemplifying the grouping of extra virgin olive oils into sub-regions of origin: a) samples from Italy (ITA); b) samples from Turkey (TUR).

and set the basis for the most efficient approach for the treatment of SH analytical data with this purpose up to date.

#### 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

This work was developed in the context of the project OLEUM “Advanced solutions for assuring authenticity and quality of olive oil at global scale”, funded by the European Commission within the Horizon 2020 Program (2014–2020, grant agreement no. 635690). The information and views set out in this article are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein. The study was also supported by the Ministerio de Ciencia, Innovación y Universidades (MICINN) from Spain through the *Juan de la Cierva* and *Ramon y Cajal* programs (JCI-2012\_13412 and RYC-2017-23601), and by the Ministerio de Educación, Cultura y Deporte (MECD) from Spain through the *FPU* pre-doctoral program (FPU16/01744).

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodchem.2019.125556>.

#### References

- Alonso-Salces, R. M., Segebarth, N., Garmon-Lobato, S., Holland, M. V., Moreno-Rojas, J. M., Fernandez-Pierna, J. A., ... Heberger, K. (2015). <sup>1</sup>H-NMR and isotopic fingerprinting of olive oil and its unsaponifiable fraction: Geographical origin of virgin olive oils by pattern recognition. *European Journal of Lipid Science and Technology*, 117, 1991–2006. <https://doi.org/10.1002/ejlt.201400243>.
- 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. <https://doi.org/10.1080/10408398.2016.1225666>.
- Ballin, N. Z., & Laursen, K. H. (2019). To target or not to target? Definitions and nomenclature for targeted versus non-targeted analytical food authentication. *Trends in Food Science & Technology*, 86, 537–543. <https://doi.org/10.1016/j.tifs.2018.09.025>.
- Ben Temime, S., Campeol, E., Cioni, P. L., Daoud, D., & Zarrouk, M. (2006). Volatile compounds from Chétoui olive oil and variations induced by growing area. *Food Chemistry*, 99(2), 3015–3325. <https://doi.org/10.1016/j.foodchem.2005.07.046>.
- Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A*, 1158, 196–214. <https://doi.org/10.1016/j.chroma.2007.05.024>.
- Bortolomeazzi, R., Berno, P., Pizzale, L., & Conte, L. (2001). Sesquiterpene, alkene, and alkane hydrocarbons in virgin olive oils of different varieties and geographical origins. *Journal of Agricultural and Food Chemistry*, 49, 3278–3283. <https://doi.org/10.1021/jf001271w>.
- Bosque-Sendra, J. M., Cuadros-Rodríguez, L., Ruiz-Samblás, C., & de la Mata, A. P. (2012). Combining chromatography and chemometrics for the characterization and authentication of fats and oils from triacylglycerol compositional data—A review. *Analytica Chimica Acta*, 724, 1–11. <https://doi.org/10.1016/j.aca.2012.02.041>.
- Camin, F., Pavone, A., Bontempo, L., Wehrens, R., Paolini, M., Faberi, A., ... Mannina, L. (2016). The use of IRMS, <sup>1</sup>H NMR and chemical analysis to characterise Italian and imported Tunisian olive oils. *Food Chemistry*, 196, 98–105. <https://doi.org/10.1016/j.foodchem.2015.08.132>.
- Commission Implementing Regulation (EU) No 29/2012 of 13 January 2012 on marketing standards for olive oil. [http://data.europa.eu/eli/reg\\_impl/2012/29/oj](http://data.europa.eu/eli/reg_impl/2012/29/oj).
- Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., ... Gallina Toschi, T. (2019). Olive oil quality and authenticity: A review of current EU legislation, standards, relevant methods of analyses, their drawbacks and recommendations for the future. *Trends in Food Science & Technology*. <https://doi.org/10.1016/j.tifs.2019.02.025> In Press.
- Damascelli, A., & Palmisano, F. (2013). Sesquiterpene fingerprinting by headspace SPME-GC-MS: Preliminary study for a simple and powerful analytical tool for traceability of olive oils. *Food Analytical Methods*, 6, 900–905. <https://doi.org/10.1007/s12161-012-9500-9>.
- García-González, D. L., & Aparicio, R. (2010). Research in olive oil: Challenges for the near future. *Journal of Agricultural and Food Chemistry*, 24, 12569–12577. <https://doi.org/10.1021/jf102735n>.
- Guinda, Á., Lanzón, A., & Albi, T. (1996). Differences in hydrocarbons of virgin olive oils obtained from several olive varieties. *Journal of Agricultural and Food Chemistry*, 44, 1723–1726. <https://doi.org/10.1021/jf950571o>.
- Melucci, D., Bendini, A., Tesini, F., Barbieri, S., Zappi, A., Vichi, S., ... Gallina Toschi, T. (2016). Rapid direct analysis to discriminate geographic origin of extra virgin olive oils by flash gas chromatography electronic nose and chemometrics. *Food Chemistry*, 204, 263–273. <https://doi.org/10.1016/j.foodchem.2016.02.131>.
- Osorio-Bueno, E., Sánchez-Casas, J., Montaña García, A., & Gallardo González, L. (2005). Discriminating power of the hydrocarbon content from virgin olive oil of Extremadura cultivars. *Journal of the American Oil Chemists' Society*, 82, 1–6. <https://doi.org/10.1007/s11746-005-1034-0>.
- Parliament Resolution EU No 2013/2091 (INI) of 14 January 2014 on food crisis, fraud in the food chain and control thereof. Official Journal of the European Union, C 482, 23 December 2016.
- Regulation (EU) No 1151/2012 of the European Parliament and of the Council of 21 November 2012 on quality schemes for agricultural products and foodstuffs. <http://data.europa.eu/eli/reg/2012/1151/oj>.
- Tomas, G., Savorani, F., & Engelsen, S. B. (2011). Icoshift: An effective tool for the alignment of chromatographic data. *Journal of Chromatography A*, 1218, 7832–7840. <https://doi.org/10.1016/j.chroma.2011.08.086>.
- Tres, A., van der Veer, G., & van Ruth, S. (2013). Vegetable oils. In M. de la Guardia, & A. Gonzalez (Eds.). *Food protected designation of origin: Methodologies and applications* (pp. 543–572). Oxford, UK: Elsevier. <https://doi.org/10.1016/B978-0-444-59562-1.00021-9>.
- Van den Dool, H., & Kratz, P. D. (1963). A generalization of the retention index system including linear temperature programmed gas—liquid partition chromatography. *Journal of Chromatography A*, 11, 463–471.

- Vichi, S., Guadayol, J. M., Caixach, J., López-Tamames, E., & Buxaderas, S. (2006). Monoterpene and sesquiterpene hydrocarbons of virgin olive oil by headspace solid-phase microextraction coupled to gas chromatography/mass spectrometry. *Journal of Chromatography A*, 1125, 117–123. <https://doi.org/10.1016/j.chroma.2006.05.029>.
- Vichi, S., Lazzez, A., Grati-Kamoun, N., López-Tamames, E., & Buxaderas, S. (2010). Evolution of sesquiterpene hydrocarbons in virgin olive oil during fruit ripening. *Journal of Agricultural and Food Chemistry*, 58, 6972–6976. <https://doi.org/10.1021/jf100497c>.
- Vichi, S., Lazzez, A., Méndez, I., & Caixach, J. (2015). “Olive oil sesquiterpene hydrocarbons as geographical markers.” Paper presented at the 13th Euro Fed Lipid Congress: Fats, Oils and Lipids: New Challenges in Technology, Quality Control and Health. Florence, Italy, September 27–30.
- Vichi, S., Tres, A., Quintanilla-Casas, B., Bustamante, J., & Guardiola, F. (2018). Sesquiterpene hydrocarbons, a promising tool for virgin olive oil geographical authentication. In M. Kontominas (Ed.). *Authentication and detection of adulteration of olive oil (chapter 11)* New York, NY: Nova Science Publishers, Inc ISBN: 978-1-53614-596-0.
- Youssef, O., Guido, F., Manel, I., Ben Youssef, N., Luigi, C. P., Mohamed, H., ... Mokhtar, Z. (2011). Volatile compounds and compositional quality of virgin olive oil from Oueslati variety: Influence of geographical origin. *Food Chemistry*, 124(4), 1770–1776. <https://doi.org/10.1016/j.foodchem.2010.08.023>.
- Zunin, P., Boggia, R., Salvadeo, P., & Evangelisti, F. (2005). Geographical traceability of West Liguria extra virgin olive oils by the analysis of volatile terpenoid hydrocarbons. *Journal of Chromatography A*, 1089, 243–249. <https://doi.org/10.1016/j.chroma.2005.07.005>.