Study of the minor fraction of virgin olive oil by a multi-class GC-MS approach: comprehensive quantitative characterization and varietal discrimination potential

Lucía Olmo-García[a], Juan J. Polari[b], Xueqi Li[c], Aadil Bajoub[d], Alberto Fernández-Gutiérrez[a], Selina C. Wang[b,c], Alegría Carrasco-Pancorbo[a*]

[a]*Department of Analytical Chemistry, Faculty of Sciences, University of Granada, Ave. Fuentenueva s/n, E-18071 Granada, Spain*

[b]*Department of Food Science and Technology, University of California Davis, One Shields Avenue, Davis, CA 95616, USA*

[c]*Olive Center, University of California Davis, One Shields Avenue, Davis, CA 95616, USA.*

[d] *Department of Basic Sciences, National School of Agriculture, km 10, Haj Kaddour Road, B.P. S/40, Meknès, Morocco.*

[*]Corresponding author:

Dr. A. Carrasco-Pancorbo, Research Group FQM-297, Department of Analytical Chemistry, Faculty of Sciences, University of Granada, Ave. Fuentenueva s/n, E-18071 Granada, Spain.

E-mail: alegriac@ugr.es

Telephone: +34 958 242785

Abstract

For the first time, a multi-class GC-MS method was applied to perform the quantitative-profiling of the minor fraction of VOOs (considering more than 40 compounds) in a single run. This comprehensive methodology has demonstrated a remarkable profiling ability on five groups of compounds (phenolic and triterpenic compounds, tocopherols, sterols and free fatty acids) with wide range of polarities/volatilities and chemical entities. After the complete analytical validation of the method, 32 VOO samples from eight different cultivars (some of them very scarcely studied before) were analyzed and the quantitative results were subjected to both non-supervised and supervised multivariate statistics for testing the capability of the determined VOO minor compounds to discriminate the varietal origin of the samples. Typical compositional profiles were defined for each cultivar and promising potential varietal markers were pointed out. The models built to discriminate Cayon and Maurino samples from the rest exhibited the best quality parameters. The relative levels of tocopherols together with characteristic concentration of luteolin, *β*-sitosterol and tyrosol were, for instance, the most specific features of Cayon VOOs.

Keywords: virgin olive oil; minor fraction; gas chromatography; mass spectrometry; multi-class methodology; chemometrics; olive cultivar; varietal marker.

Abbreviations: VOO, virgin olive oil; GC, gas chromatography; MS, mass spectrometry; EtOH, ethanol; BSTFA+1% TMCS, N,O-bis(trimethylsilyl)trifluoroacetamide plus 1% of trimethylchlorosilane; C16:1, palmitoleic acid; C18:1, oleic acid; C18:2, linoleic acid; *α*-, *β*-, *γ*- and *δ*-Toc, *α*-, *β*-, *γ*- and *δ*-tocopherols; OA, oleanolic acid, UA, ursolic acid; BA, betulinic acid; MA, maslinic acid; ER, erythrodiol; UV, uvaol; Sti, stigmasterol; Cam, campesterol; *β*-Sit, *β*-sitosterol; Lut, luteolin; Api, apigenin; Pin, pinoresinol; Van, vanillin; *p*-Cou, *p*-

coumaric acid; Qui, quinic acid; Fer, ferulic acid; TY, tyrosol; HTY, hydroxytyrosol; QC, quality control sample; LLE, liquid-liquid extraction; IT, Ion Trap; S/N, Signal to noise ratio; LOD, limit of detection; LOQ, limit of quantification; Rt, retention time; %RSD, relative standard deviation; AcHTY acetylated HTY; OleAgly, oleuropein aglycone; DOA, decarboxymethyl oleuropein aglycone; LigAgly, ligstroside aglycone; DLA, decarboxymethyl ligstroside aglycone; AcPin, acetoxypinoresinol; C16:0, palmitic acid; C18:0, stearic acid; Δ^5-Ave, Δ^5-avenasterol; CyArten, cycloartenol; MeCyArtan, methylencycloartanol; Cit, citrostadienol; EA, elenolic acid; ANOVA, analysis of variance; PCA, Principal Component Analysis; PLS-DA, Partial Least Squares-Discriminant Analysis; SD, standard deviation; R^2, regression coefficient; PC, principal component; R, correlation; RMSEP, Root Mean Square Error of Prediction; SEP, Standard Error of Performance.

1. Introduction

The health benefits associated with virgin olive oil (VOO) intake and its unique sensory properties are the main reasons of increased olive oil consumption and production during the last decades (Türkekul, Günden, Abay, & Miran, 2010; Vossen, 2007).

Different olive oil categories can be found in the markets, but only VOO -obtained exclusively by mechanical means without any refining steps- preserves its minor compounds that are responsible for the taste and nutritional value. The VOO minor fraction comprises a heterogeneous mix of molecules, including phenolic compounds (simple phenols, phenolic acids, secoiridoids, flavonoids and lignans), triterpenic compounds (acids and dialcohols), tocopherols and sterols (Ghanbari, Anwar, Alkharfy, Gilani, & Saari, 2012). In any VOO the concentration of these minor compounds is highly influenced by agro-technological parameters such as cultivar, pedoclimatic conditions, irrigation methods, extraction procedures and storage practices (Servili et al., 2014). Acceptable concentration ranges for some of these compounds have been included in several legal frames settled to protect consumers from product mislabeling and adulteration (Codex Alimentarius, 2015; European Commission, 2016; United States Department of Agriculture, 2010) as well as to promote health claims related to VOO biomolecules (European Commission, 2012). Therefore, the determination of these compounds is of great interest to both VOO producers and regulatory bodies, who are continuously challenging the analytical community to offer rapid and accurate testing methods (Bajoub, Bendini, Fernández-Gutiérrez, & Carrasco-Pancorbo, 2017; Tena, Wang, Aparicio-Ruiz, García-González, & Aparicio, 2015; Tsimidou et al., 2019).

Gas chromatography (GC) is a very common separative technique used by routine laboratories. Indeed, it is the technique of choice in several official methods for the analysis of different VOO components such as waxes, fatty acid methyl esters, fatty acid ethyl esters, aliphatic alcohols, sterols and triterpenic dialcohols, among others (European Commission,

2016). The flame ionization detector has traditionally been the most used detection system for GC because of its acceptable sensitivity, broad linear dynamic ranges and affordable prices. It is still widely employed as seen in recent contributions (Gorassini, Verardo, & Bortolomeazzi, 2019; Li, Flynn, & Wang, 2019). However, mass spectrometry (MS) is ousting the just mentioned detector due to its capability for identity confirmation and quantification of overlapped peaks. Thus, the use of MS detectors coupled to the unbeatable resolving power of GC seems to be a promising analytical approach for characterizing the complex VOO minor fraction, as demonstrated by some interesting applications published over the last years. For example, different GC-MS methods were successfully applied to investigate the impact of the decanter type on the fatty acids, sterols, triterpenic acids and phenolic compounds profiles of the obtained oils (Kalogeropoulos, Kaliora, Artemiou, & Giogios, 2014). Later on, vegetable oil minor (apolar and mid-polar) components fingerprinting was achieved by applying a bidimensional GC×GC-MS method (Purcaro, Barp, Beccaria, & Conte, 2015). More recently, a simple methodology based on solid phase micro extraction and GC-MS was described for the characterization of edible oils minor components (including alcohols, aldehydes, epoxides, hydrocarbons, ketones, sterols and tocols, among others) in one chromatographic run (Alberdi-Cedeño, Ibargoitia, Cristillo, Sopelana, & Guillén, 2017).

The last two contributions represent a couple of meaningful examples of multi-class methodologies that exhibit the ability to monitor analytes belonging to diverse chemical classes in one single analysis. As a consequence, they bring out a remarkable progression of the traditional single-class methods in terms of throughput and cost. At the same time, they enlarge the information achievable by the analyst and provide enhanced possibilities to take advantage of the results. In other words, comprehensive profiling methods allow comparing samples from a deeper perspective, providing quantitative data on a great number of

substances and facilitating the extraction of relevant information through the use of chemometrics. When applied to VOO, multi-class methodologies can be used, for instance, to correlate the healthy properties of an oil with its minor fraction composition (Vazquez et al., 2019), to authenticate the commercial category of the oils (Kalogiouri, Alygizakis, Aalizadeh, & Thomaidis, 2016), or to look for distinctive features to classify samples depending on their botanical (Kalogiouri, Aalizadeh, & Thomaidis, 2018) or geographical origin (Olmo-García, Wendt, et al., 2019).

The main goal of this study was to demonstrate the suitability of a GC-MS multi-class methodology for the quantitative assessment of VOO minor compounds of different chemical nature (phenolic and triterpenic compounds, tocopherols, sterols and free fatty acids) in a single run. We also had the intention of checking the capability of the determined compounds to trace the varietal origin of VOO samples, in the same way as single-class approaches have previously demonstrated for compounds such as triacylglycerols (Blasi, Pollini, & Cossignani, 2019), phenolic compounds (Miho et al., 2018), sterols (Abdallah et al., 2018) and volatile aroma compounds (Lukić, Carlin, Horvat, & Vrhovsek, 2019). The analytical performance of the method was assessed and then, it was applied to the analysis of VOO samples from different cultivars grown under the same conditions in an orchard in California. The quantitative characterization of the selected oils is considered to be very relevant, as the information about the chemical composition of some of the cultivars included in this study is quite scarce in literature. Apart from that, the use of chemometrics expedited the study of the results and made possible the establishment of statistical models to differentiate samples with distinctive botanical provenance.

2. Materials and methods

2.1. Chemicals and reagents

Deionized water generated by a MilliQ system (Millipore, Bedford, MA) and ethanol (EtOH) 95% from Koptec (King of Prussia, PA, USA) were used for the extraction of the VOO minor components. The derivatization reagent (N,O-bis(trimethylsilyl)trifluoroacetamide plus 1% of trimethylchlorosilane (BSTFA+1% TMCS)) as well as commercial standards of palmitoleic (C16:1), oleic (C18:1) and linoleic (C18:2) acids, *α*-, *β*-, *γ*- and *δ*-tocopherols (*α*-, *β*-, *γ*- and *δ*-Toc), oleanolic (OA), ursolic (UA), betulinic (BA) and maslinic (MA) acids, erythrodiol (ER), uvaol (UV), stigmasterol (Sti), campesterol (Cam), *β*-sitosterol (*β*-Sit), luteolin (Lut), apigenin (Api), pinoresinol (Pin), vanillin (Van), *p*-coumaric (*p*-Cou), quinic (Qui) and ferulic (Fer) acids, tyrosol (TY) and hydroxytyrosol (HTY) were purchased from Sigma Aldrich (St. Louis, MO, USA). All the stock solutions and extracts were filtered through 0.4 µm nylon syringe filters (Thermo Scientific, Santa Clara, CA, USA) and stored in dark glass vials at −20 ºC.

2.2. Samples

Olive fruit sampling was performed in October 2016 on eight different cultivars (*cv.* Carolea, Casaliva, Cayon, Frantoio, Kalamon, Maurino, Moraiolo and Taggiasca) grown in an experimental orchard of the UC Davis Olive Center (Davis, CA, USA) under controlled agronomical conditions. The geographical coordinates of olive grove are 38°32'10"N 121°47'42"W and the altitude is around 16 m. The area has a *Csa* climate type according to the Köppen-Geiger climate classification (Kottek, Grieser, Beck, Rudolf, & Rubel, 2006), with average temperatures of 16.2ºC (8.7 and 23.7°C, for minimum and maximum means, respectively), and annual rainfalls of around 500 mm for the year of 2016. Four batches of fruit samples (4 kg) from each cultivar (32 samples in total), with ripening indices between 2.3 and 2.9 (according to the Jaén method (Uceda & Hermoso, 1998)), were hand-picked from randomly selected olive trees. Those relatively low ripening index values with regard to the European standards are quite common taking into account the location of the olive grove

and the Californian practices. VOOs from each sample were obtained within the next 3 h by means of an Abencor® laboratory oil mill (MC2 Ingeniería y Sistemas, Seville, Spain) and stored at -20ºC until further analysis. A quality control (QC) sample was prepared by mixing equivalent amounts of individual VOO samples; it was used to assess the analytical performance of the method in a first stage of the study, as it will be explained in section 2.4.

2.3. Extraction and GC-MS determination of minor compounds

The extraction of the minor components from VOO samples was performed by applying a previously published liquid-liquid extraction (LLE) protocol (Olmo-García et al., 2018b). In short, 1.00 ± 0.01 g of VOO was successively extracted (by using vortex followed by centrifugation and collection of the supernatants) with three 6 mL portions of EtOH/H_2O (80:20, v/v) and one portion of EtOH/H_2O (60:40, v/v). After solvent evaporation, the residue was reconstituted in 1 mL of EtOH/H_2O (80:20, v/v). An aliquot of 200 μL of the extract (or the appropriate amount of standards mixture) was evaporated to dryness under nitrogen. Then, the residue was derivatized by adding 50 μL BSTFA+1% TMCS and kept at room temperature for 1 h before injecting into the GC. When necessary, more diluted extracts were also injected to assure the quantification of the analytes under study within the linear dynamic ranges.

The analysis of the prepared extracts was performed as described in a previous work (Olmo-García et al., 2018b) on a Varian 450 gas chromatograph coupled with a Varian 220 mass spectrometer equipped with an Ion Trap (IT) analyzer (Agilent Technologies, Santa Clara, CA, USA). The separation of the analytes was carried out using a (5%-phenyl)-methyl polysiloxane (HP-5MS) capillary column (30 m x 0.25 mm i.d., 0.25 μm) (Agilent Technologies) with He as carrier gas at 1 mL/min. A 52.5 min temperature gradient was used: the temperature was initially held at 140ºC for 5 min, followed by a 4°C/min ramp to 310°C (held for 5 min). A sample volume of 1 μL was injected at a split ratio of 1:25. The injector

and transfer line temperatures were 240ºC and 290ºC, respectively. Spectra (in a range from 50 to 600 *m/z*) were recorded with the electron impact ion source operating in positive mode at 70 eV and 200ºC.

2.4. Method validation

Both the QC sample and a standards mixture containing 25 VOO minor compounds were used for method validation.

Method linearity was checked by establishing external calibration curves for each individual compound, considering the concentration values (X-axis) and MS area (Y-axis). For this purpose, a stock solution of the standards mixture was prepared in EtOH/H_2O (80:20, v/v) and serially diluted to eight different concentration levels (ranging from 0.1 mg/L to a maximum level that depended on the compound (and varied from 50 to 500 mg/L)); each level was analyzed in triplicate. For most of the substances the maximum assayed level was 50 mg/L (approx.); for C18:2, β-Toc, γ-Toc, Cam, OA and MA, the highest concentration level of the calibration range was 200 mg/L; in the case of TY, HTY, C16:1, C18:1 and α-Toc, 500 mg/L was the maximum concentration value which was tested.

The concentration levels of some of the compounds that were found in the VOO samples, such as fatty acids, secoiridoid derivatives and α-Toc, fluctuated a lot. Thus, we had to cover a broad concentration range, which made us to establish several quantification linear ranges for each analyte. In the table including the analytical parameters of the method (Table 1, which will appear in section 3.1), we just give the equation of the intermediate linear range.

Method accuracy was assessed in terms of precision (under repeatability conditions) and trueness. *Intra* and *inter*-day repeatability for peak area and retention time (Rt) were expressed as the relative standard deviation (%RSD) obtained from four injections of four different extracts of the QC, which were carried out within the same day and over four different days, respectively. Trueness was estimated by analyzing the QC extracted before and

after fortification with the mixture of standards at three distinct concentration levels (low, intermediate and high); the recovery for each single pure standard was estimated afterwards, by applying the following equation:

% Recovery = (Concentration in the fortified QC – Concentration in the QC) / Spiked concentration × *100*

Additionally, matrix effect was evaluated for some standards which have not been previously studied by our team (Olmo-García et al., 2018b), such as C16:1, C18:1, *δ*-Toc, Cam, Sti, *β*-Sit and UA. For that purpose, the slope of the external calibration curve (prepared in solvent) was compared with the slope resulting from the standard addition (at three concentration levels) to the QC extract. A matrix effect coefficient was calculated (in percentage) for each analyte: the more similarity between the values of two slopes, the less significant the magnitude of the matrix effect. In the same way, theoretical limits of detection (LOD) and quantification (LOQ) were estimated for those previously unevaluated standards by calculating the concentration that generates a signal to noise ratio (S/N) equals to 3 and 10, respectively.

2.5. Data treatment

Instrument control and data processing were performed with MS Workstation v. 6.9.3 (Agilent Technologies). All the samples were analyzed in triplicate. External calibration curves were used to convert automatically integrated peak areas into concentrations. Good linearity was achieved for all the standards based on least-squares regression. Analytes lacking pure standards were quantified in terms of a structure-related compound (commercially available): HTY calibration curve was used for quantification of acetylated HTY (AcHTY), oleuropein aglycone isomers (OleAgly) and decarboxymethyl oleuropein aglycone (DOA); TY calibration curve was applied for ligstroside aglycone isomers (LigAgly) and decarboxymethyl ligstroside aglycone (DLA); Pin calibration curve was used for acetoxypinoresinol (AcPin); the relative response of C16:1 standard was used to quantify

palmitic acid (C16:0); C18:1 for stearic acid (C18:0); and β-Sit for Δ^5-avenasterol (Δ^5-Ave), cycloartenol (CyArten), methylencycloartanol (MeCyArtan) and citrostadienol (Cit). Elenolic acid (EA) does not have a commercially available standard, but since it is considered as a highly related compound to secoiridoids, it has been frequently quantified in terms of oleuropein in LC-MS (Bajoub et al., 2016). In this study, the *m/z* of the derivatized oleuropein pure standard was out of the selected working mass range and, therefore, it could not be properly detected. Thus, in the absence of a suitable standard to accurately carry out EA isomers quantification, their area was directly used for statistics after the required normalization.

Statistical analysis was performed with Microsoft Excel 2013 (Microsoft Corporation, Redmond, WA, USA) and The Unscrambler v9.7 (CAMO Software, Inc., Woodbridge, NJ, USA). In a first stage, one-way analysis of variance (ANOVA) was carried out to determine the significant difference(s) regarding the concentration of the targeted analytes among different cultivars. Afterwards, the natural clustering of the samples was studied by conducting Principal Component Analysis (PCA). The PCA matrix was composed by 39 variables (determined VOO minor compounds) and 32 samples (average value of triplicate measurements). Apart from it, Partial Least Squares-Discriminant Analysis (PLS-DA) was performed to build two-class models by confronting the samples of each cultivar against the rest of the samples (which composed one global group in each case). Data normalization was carried out (for both PCA and PLS-DA) to reduce experimental variance and all variables were weighted by 1/standard deviation (1/SD) for PLS-DA to allow all of them to contribute to the model, regardless of whether the quantitative value had a small or large standard deviation (SD) from the outset. Full cross-validation was applied to evaluate the prediction power of the obtained models.

3. Results and discussion

3.1. Analytical parameters of the method

In a first stage of the study, full method validation was conducted in order to evaluate the adequacy of the quantification methodology to study VOO samples. Thus, analytical parameters of the method were tested considering 25 minor VOO compounds belonging to 5 different chemical classes.

Table 1a summarizes the main analytical parameters of the method, which give a profound idea of its linearity, sensitivity and accuracy. The table also contains information about the *m/z* signals used for the identification and quantification of each substance. The selected quantifier signals were chosen considering specificity, relative response and S/N (seeking the achievement of the most favorable LOD and LOQ). All the external calibration curves showed good linearity within the evaluated concentration ranges, with regression coefficients (R^2) higher than 0.988.

As far as precision is concerned, %RSD values for *intra* and *inter*-day repeatability, in terms of area, were lower than 5.9% (Pin) and 9.2% (*β*-Sit), respectively. In general, the *intra* and *inter*-day repeatability in terms of Rt were excellent, exhibiting very low values; Pin was the compound which presented the highest *inter*-day %RSD value (0.05%). In addition, good recoveries were found for most of the analytes with values ranging from 80.7 to 105.7%, which are within the limits proposed by the AOAC for a truthful method (AOAC, 2002). Only two sterols (Cam and *β*-Sit) presented recoveries slightly lower than 80% (78.6 and 75.1%, respectively); in spite of it, those values were reasonably good and, most importantly, the repeatability of the overall process including both the sample extraction and instrumental analysis was outstanding (<10% RSD measured in terms of area).

Table 1b shows matrix effect, LOD and LOQ values of the standards that had not been appraised before, which were assessed as described in Section 2.4. Having in mind the previously reported results (Olmo-García et al., 2018b), for 23 out of 25 standards, the

calculated matrix effect coefficients were in the range between -14.2% and +16.3%, indicating a mild signal suppression or enhancement effect (from −20% to +20%). Nevertheless, Lut presented a slight enhancement effect (21.3 %) and Cam was suppressed to some extent (-36.2%). External calibration equations were used for targeted analytes quantification based on the following assumptions: i) standard addition calibration implies the construction of a calibration curve for each sample; and ii) the matrix effect was firstly evaluated using a QC sample which was a mixture of equivalent volumes of all the VOOs under study. Afterwards, the matrix effect of each cultivar was checked individually and the slight enhancement/suppression observed for Lut and Camp was not found as noticeable as within the QC sample. Thus, reliable quantitative results could be obtained for the 32 VOO samples by employing the external calibration approach.

3.2. Application of the method to the analysis of the selected samples

Extracts of 32 VOO samples from eight different cultivars were analyzed in the current study by using the described GC-MS methodology. Fig. 1 shows the complexity of the chromatograms obtained from three oils (Frantoio, Kalamon, and Cayon) which were selected as example of very distinct minor compounds profiles. The applied methodology was useful to get information about five different chemical classes of VOO minor compounds. In particular, 41 compounds were determined in the studied samples: 19 phenolic compounds (seven simple phenols, eight secoiridoid derivatives, two flavonoids and two lignans), four tocopherols, six triterpenic compounds (two triterpenic alcohols and four acids), seven sterols and five free fatty acids. Moreover, two EA peaks were identified and integrated in the chromatograms; and their reported area (normalized) was also included in the statistical analysis.

Table 1 SM (Supplementary material) shows Rt, name of the analyte, *m/z* signals sorted by relative abundance within the mass spectrum and the formula of the detected feature that allowed the identification of each compound (in bold letters).

The quantitative data obtained for the 41 selected analytes are presented in Table 2. The given number is the average value coming from four independent VOO samples from each cultivar, which were obtained from olives harvested from different olive trees and processed separately. For each sample, we calculated the mean of three extraction and injection replicates. Later on, further calculations were made to achieve the global value (on average) for each cultivar, combining all the results from the different samples belonging to each variety. Some compounds presented high variability within samples from the same cultivar, whilst the levels of some others remained constant. For example, ER levels were very similar in all the VOOs coming from the same olive variety (less than 11% RSD among samples belonging to the same cultivar) conversely to Qui, whose *intra*-cultivar fluctuation was substantially higher in Casaliva samples. Some analytes belonging to phenolic compounds-chemical class (in particular, secoiridoids) usually show several isomeric forms in the profiles, as it has been extensively discussed in literature (Olmo-García, Bajoub, Monasterio, Fernández-Gutiérrez, & Carrasco-Pancorbo, 2017; Olmo-García, Fernández-Fernández, et al., 2019). In Table 2, we have denoted the different isomers by adding a Roman numeral after the name of the compound. *Intra*-cultivar variations were also found, for instance, for LigAgly I and OleAgly III concentrations; they varied more than 30% in the samples of five (Taggiasca, Moraiolo, Frantoio, Cayon and Carolea) and four cultivars (Taggiasca, Moraiolo, Frantoio, Cayon), respectively. However, when the total concentration of secoiridoid aglycones was considered, the overall *intra*-cultivar variability remarkably decreased, suggesting that the distribution of the isomers varied more than their global levels in the samples. It is also worthwhile to highlight the high *intra*-cultivar %RSD of free fatty acids in

Moraiolo samples (ranging from 45 to 50%). This fact, together with the high concentration levels of C16:0 and C18:1 found in these VOOs, generated the highest SD values shown in Table 2 (241 and 197 mg/kg, respectively).

The level of different minor compounds greatly differed from one cultivar to the other (as illustrated in Fig. 1), which is in agreement with previous reports which state that the cultivar is one of the most influential factors affecting VOO composition (Bajoub et al., 2016, 2015; Olmo-García, Bajoub, Monasterio, Fernández-Gutiérrez, & Carrasco-Pancorbo, 2018; Sánchez de Medina, Priego-Capote, & Luque de Castro, 2015). Regarding *inter*-cultivar variability, *p*-Cou and DLA were the compounds presenting the greatest variances, with concentration ranges from around 0.2 to 7.6 mg/kg and 1.6 to 356 mg/kg, respectively. Consequently, it might be expected that they will have a significant impact on the statistical modeling to discriminate the varietal origin of the samples, as it will be detailed in section 3.3.

Even though the *inter*-cultivar differences were not as drastic as for the just two mentioned compounds, significant disparities were found in other substances. For instance, the concentration of LigAgly I and OleAgly III considerably varied among the cultivars; indeed, LigAgly I levels varied from 2.3 to 62 mg/kg in Moraiolo and Casaliva, respectively, and the found amounts of OleAgly III ranged from 3 to 73 mg/kg for Cayon and Maurino, apiece. Something similar was observed for other isomeric forms of these secoiridoid aglycons. Regarding fatty acids, C16:0 levels varied substantially in the studied oils, showing averaged values of 28 mg/kg for Cayon *cv*. and of 520 mg/kg for Frantoio oils. Another example to be cited is α-Toc, which showed concentrations within the range 96-460 mg/kg, defined by the mean values of Moraiolo and Cayon, respectively.

Fig. 2 illustrates the differences in the VOO minor compounds content in the tested oils, grouped by chemical class. A normalized scale is used to facilitate the proper evaluation of

the results. Kalamon was considered the richest cultivar in terms of phenolic and triterpenic compounds. This cultivar is characterized by very high levels of DLA as well as by a remarkable content of OA and MA. Cayon samples presented the highest levels of tocopherols and sterols, whereas Frantoio VOOs had the major content of free fatty acids. Sterols turned out to be the chemical class presenting the lowest variation among the eight tested cultivars.

A reasonable comparison with previously published results regarding concentration values could just be made when the same pure standard is used for the quantification of a given compound. Moreover, information about the chemical composition of VOO from some of the cultivars studied herein is barely reported. Nevertheless, a global comparison of the concentration ranges found in VOO minor compounds reveals that the applied methodology gives comparable results to previous studies. Levels found in phenolic compounds are in agreement with results obtained by applying LC-FLD and LC-MS methodologies (Bajoub et al., 2016; Monasterio, Olmo-García, Bajoub, Fernández-Gutiérrez, & Carrasco-Pancorbo, 2016), except for lignans, which are closer to the levels reported by Fuentes et al. (Fuentes et al., 2018). As far as tocopherols are concerned, concentrations found in this study are similar to those described by different authors for VOOs of different varietal origins (Beltrán et al., 2010; Sayago, González-Domínguez, Urbano, & Fernández-Recamales, 2019). Triterpenic compounds content is also within the range previously found in other VOOs (Olmo-García, Bajoub, et al., 2018). Sterols are generally quantified using an internal standard (Kyçyk, Aguilera, Gaforio, Jiménez, & Beltrán, 2016) instead of the corresponding response of each pure standard (which would make possible to achieve an absolute quantification). When comparing sterol content in mg/kg, our results, particularly regarding β-Sit, seem to be lower than the concentrations obtained by Gu and coworkers (Gu et al., 2016), although in that study the number of samples was very limited; in a more recent publication, the β-Sit mean

value found for 16 olive oils was even lower than the levels found in our samples (Zhang et al., 2019). Regarding free fatty acids, C18:2 content is generally below the concentration range reported by Wabaidur and coworkers (Wabaidur et al., 2016); the same is observed for the C18:1 content in Cayon, Kalamon and Maurino VOOs.

3.3. Statistical analysis

The visual inspection of chromatograms and quantitative data showed the significant compositional heterogeneity among the samples obtained from different cultivars. This was confirmed by ANOVA, finding significant differences ($p < 0.05$) for all the tested variables (determined chemical compounds) except for Qui and logically for the three triterpenic compounds that were found under the LOQs or LOQs (UV, BA and UA). Multivariate analyses were consequently carried out on the data matrix composed by a total of 39 variables, including the 37 quantified compounds from Table 2 (without Qui) plus two isomers of EA (considering their area value), in order to evaluate the whole data set-structure and test the discriminant power of the determined VOO minor compounds to distinguish the varietal origin of the oils under evaluation.

Firstly, non-supervised PCA was performed as an exploratory approach to study data structure over a reduced dimension. Among the 20 identified principal components (PCs), the first five components explained 99.05 % of the total variance. The obtained PCA score plots and loadings plots for the first four PCs are shown in Fig. 1 SM. The first two PCs, which covered 76.46 % and 16.99 % of the variance, respectively, exhibited good discrimination capability among Kalamon, Cayon, Casaliva and Maurino (to a less extent, as one of the Moraiolo samples was quite close to Maurino´s) VOOs (Fig. 1 SM A1). Another grouping could be perceived in the PC1 *vs*. PC2 scores plot which encompasses Carolea, Moraiolo, Taggiasca and Frantoio samples together. The third and fourth PCs simply covered 2.98 % and 1.54 % of the variance, respectively. Improved separation of Taggiasca, Frantoio and

Carolea samples could be barely found in the PC3 *vs*. PC4 scores plot (Fig. 1 SM B1). Both PC1 *vs*. PC2 (Fig. 1 SM A2) and PC3 *vs*. PC4 (Fig. 1 SM B2) loadings plots, revealed the importance of the following variables for the clustering of the samples: DLA, α-Toc, β-Sit, C16:0, C18:1 and OleAgly III. An extra cluster containing QC samples was found in both scores plots (Fig. 1 SM A1 and B1). Those samples were interspersed in the analytical sequence at regular intervals (every 20 injections) to evaluate system stability. Thus, the fact that they were grouped in the center of the graphs demonstrates the satisfactory repeatability of the applied analytical method.

Next, supervised chemometrics were applied to build two-class discrimination models through PLS-DA; the resulting PC1 *vs*. PC2 scores plots are presented in Fig. 3. The worst class separation was found for the models built to discriminate Frantoio (d), Moraiolo (g) or Taggiasca (h) samples from the rest. That can be corroborated looking at Table 3, which presents the key parameters used to assess the quality of the models, such as R (or correlation), which measures the linear relationship between the predicted and measured values; R-Square; Root Mean Square Error of Prediction (RMSEP), which can be interpreted as the average prediction error; Standard Error of Performance (SEP), which is the standard deviation of the prediction residuals; and Bias, which is calculated as the average value of the residuals. Table 3 also includes the possible varietal markers that are useful to distinguish the VOO samples belonging to each cultivar from the rest.

Moraiolo and Taggiasca models had the lowest correlation and R-square parameters and the model for discriminating Frantoio from the rest needed the highest number of components to achieve reasonable quality parameters. The best quality parameters were found for Cayon and Maurino models. In Table 3, distinctive features are presented with their estimated regression coefficients (between brackets), value which points out the cumulative importance of each variable (chemical compound) to identify the varietal origin. Some compounds were common

possible markers for different cultivars, such as MA, γ-Toc and Δ^5-Ave, which were significant variables for three models, or δ-Toc which was influential for four. Negative coefficients imply a negative contribution; for example, ER and AcHTY were two of the most distinctive features for Moraiolo and Taggiasca, but it was due to their low levels in the first case and because of their high content in the second one. As revealed in Table 3, and trying to underline specific varietal-features, high levels of *p*-Cou, EA isomers, C18:2 and Api, together with low levels of MA were typical for Carolea oils. The model to distinguish Casaliva *cv*. from the rest was mainly defined by the influence of Δ^5-Ave, two tocopherols, LigAgly I and the low concentrations of C16:1 and MA. The relative levels of tocopherols (β-,α-, and δ-Toc) together with characteristic concentrations of Lut, β-Sit and TY were the most specific features of Cayon VOOs. Fer and *p*-Cou, two phenolic acids, seemed to be particularly relevant in Frantoio. The role of δ-Toc and AcPin, among other substances (two free fatty acids), was also remarkable in the same model. Moreover, as stated above, Kalamon VOOs showed the highest levels of triterpenic compounds (MA and OA in particular) and DLA and DOA, with contrasting low levels of Δ^5-Ave and γ-Toc. The three determined isomers of OleAgly, as well as Pin and Van were found at considerably high levels in Maurino oils, for which the comparatively moderate concentrations of DOA could be noted as a typical feature for this cultivar. Moraiolo was one of the cultivars presenting the highest amounts of γ-Toc; on the contrary, its overall profile was unusual considering its low concentrations of AcPin, Cit, Fer, ER and AcHTY. The variables with higher absolute values of regression coefficients to characterize Taggiasca VOOs were Fer, ER and Cit.

Nonetheless, the potential markers designated in the current study have to be further tested and validated with a more comprehensive sample-set (covering different seasons). However, we believe that the importance of the findings presented herewith is undeniable. The latter together with the evidence of the usefulness of the powerful GC-MS multi-class methodology

to perform the comprehensive quantitative-profiling of VOO minor compounds within a single run are, from our point of view, the most valuable achievements of this work.

4. Conclusions

The effectiveness of a multi-class GC-MS methodology to carry out the minor fraction profiling of VOOs has been evaluated using 32 samples coming from eight different cultivars. Promising results have been achieved as: 1) a satisfactory analytical performance has been exhibited by the proposed method; 2) a comprehensive quantitative characterization of eight cultivars has been accomplished, successfully determining more than 40 compounds (phenolic and triterpenic compounds, tocopherols, sterols and free fatty acids); and 3) PLS-DA models have been established to discriminate among the eight selected cultivars and, most importantly, to identify potential varietal markers. Innovative tools and methods providing extensive information in just one run are absolutely in great demand when demonstrating the typicity and genuineness of an olive oil. Future studies could apply the proposed analytical methodology and statistical models; indeed, the new methodology represents a very useful implement for the "tool-box" of a wide number of laboratories worldwide.

Acknowledgements

The authors would like to thank Ava Moin for the assistance with sample preparation. This study was partially supported by a grant from the Spanish Government (Ministerio de Educación, Cultura y Deporte) (FPU13/06438) and the Mobility Program for young researchers University of Granada and CEI BioTic 2015-2016, which made possible L.O-G's pre-doctoral stay at UC Davis. J. P. acknowledges financial support from Fulbright Foundation.

Conflict of interest statement

The authors declare no conflict of interest.

References

Abdallah, M., Vergara-Barberán, M., Lerma-García, M. J., Herrero-Martínez, J. M., Zarrouk, M., Guerfel, M., & Simó-Alfonso, E. F. (2018). Sterol profiles of Tunisian virgin olive oils: classification among different cultivars and maturity indexes. *European Food Research and Technology*, *244*(4), 675–684. https://doi.org/10.1007/s00217-017-2990-3

Alberdi-Cedeño, J., Ibargoitia, M. L., Cristillo, G., Sopelana, P., & Guillén, M. D. (2017). A new methodology capable of characterizing most volatile and less volatile minor edible oils components in a single chromatographic run without solvents or reagents. Detection of new components. *Food Chemistry*, *221*, 1135–1144. https://doi.org/10.1016/j.foodchem.2016.11.046

AOAC. (2002). Guidelines for single laboratory validation of chemical methods for dietary supplements and botanicals. *AOAC International*, 1–38.

Bajoub, A., Bendini, A., Fernández-Gutiérrez, A., & Carrasco-Pancorbo, A. (2017). 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 Food Sci. Nutr.*, (November), 1–116. https://doi.org/10.1016/j.jss.2009.11.004

Bajoub, A., Medina-Rodríguez, S., Olmo-García, L., Ajal, E. A., Monasterio, R., Hanine, H., … Carrasco-Pancorbo, A. (2016). In-Depth Two-Year Study of Phenolic Profile Variability among Olive Oils from Autochthonous and Mediterranean Varieties in Morocco, as Revealed by a LC-MS Chemometric Profiling Approach. *International Journal of Molecular Sciences*, *18*, 52–74. https://doi.org/10.3390/ijms18010052

Bajoub, A., Pacchiarotta, T., Hurtado-Fernández, E., Olmo-García, L., García-Villalba, R., Fernández-Gutiérrez, A., … Carrasco-Pancorbo, A. (2015). Comparing two metabolic profiling approaches (liquid chromatography and gas chromatography coupled to mass spectrometry) for extra-virgin olive oil phenolic compounds analysis: A botanical classification perspective. *Journal of Chromatography A*, *1428*, 267–279. https://doi.org/10.1016/j.chroma.2015.10.059

Beltrán, G., Jiménez, A., del Rio, C., Sánchez, S., Martínez, L., Uceda, M., & Aguilera, M. P. (2010). Variability of vitamin E in virgin olive oil by agronomical and genetic factors. *Journal of Food Composition and Analysis*, *23*(6), 633–639. https://doi.org/10.1016/j.jfca.2010.03.003

Blasi, F., Pollini, L., & Cossignani, L. (2019). Varietal Authentication of Extra Virgin Olive Oils by Triacylglycerols and Volatiles Analysis. *Foods*, *8*(2), 58. https://doi.org/10.3390/foods8020058

Codex Alimentarius. (2015). Standard for olive oils and olive pomace oils CODEX STAN 33-1981. Retrieved from www.fao.org/input/download/standards/88/CXS_033e_2015.pdf

European Commission. (2012). Commission Regulation (EU) No 432/2012 establishing a list of permitted health claims made on foods other than those referring to the reduction of disease risk and to children's development and health. *Official Journal of the European*

520 *Union*, *L136*, 1–40.

521 European Commission. (2016). Commission implementing regulation (EU) 2016/1095 of 6
522 July 2016 amending Regulation (EEC) No 2568/91 on the characteristics of olive oil and
523 olive-residue oil and on the relevant methods of analysis. *Official Journal of the*
524 *European Union*.

525 Fuentes, E., Paucar, F., Tapia, F., Ortiz, J., Jimenez, P., & Romero, N. (2018). Effect of the
526 composition of extra virgin olive oils on the differentiation and antioxidant capacities of
527 twelve monovarietals. *Food Chemistry*, *243*, 285–294.
528 https://doi.org/10.1016/j.foodchem.2017.09.130

529 Ghanbari, R., Anwar, F., Alkharfy, K. M., Gilani, A. H., & Saari, N. (2012). Valuable
530 nutrients and functional bioactives in different parts of olive (Olea europaea L.)-A
531 review. *International Journal of Molecular Sciences*, *13*(3), 3291–3340.
532 https://doi.org/10.3390/ijms13033291

533 Gorassini, A., Verardo, G., & Bortolomeazzi, R. (2019). Polymeric reversed phase and small
534 particle size silica gel solid phase extractions for rapid analysis of sterols and triterpene
535 dialcohols in olive oils by GC-FID. *Food Chemistry*, *283*, 177–182.
536 https://doi.org/10.1016/j.foodchem.2018.12.120

537 Gu, Q., Yi, X., Zhang, Z., Yan, H., Shi, J., Zhang, H., … Shao, J. (2016). A facile method for
538 simultaneous analysis of phytosterols, erythrodiol, uvaol, tocopherols and lutein in olive
539 oils by LC-MS. *Analytical Methods*, *8*(6), 1373–1380.
540 https://doi.org/10.1039/c5ay02193f

541 Kalogeropoulos, N., Kaliora, A. C., Artemiou, A., & Giogios, I. (2014). Composition, volatile
542 profiles and functional properties of virgin olive oils produced by two-phase vs three-
543 phase centrifugal decanters. *LWT - Food Science and Technology*, *58*(1), 272–279.
544 https://doi.org/10.1016/j.lwt.2014.02.052

545 Kalogiouri, N. P., Aalizadeh, R., & Thomaidis, N. S. (2018). Application of an advanced and
546 wide scope non-target screening workflow with LC-ESI-QTOF-MS and chemometrics
547 for the classification of the Greek olive oil varieties. *Food Chemistry*, *256*, 53–61.
548 https://doi.org/10.1016/j.foodchem.2018.02.101

549 Kalogiouri, N. P., Alygizakis, N. A., Aalizadeh, R., & Thomaidis, N. S. (2016). Olive oil
550 authenticity studies by target and nontarget LC-QTOF-MS combined with advanced
551 chemometric techniques. *Analytical and Bioanalytical Chemistry*, *408*(28), 7955–7970.
552 https://doi.org/10.1007/s00216-016-9891-3

553 Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-
554 Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3), 259–263.
555 https://doi.org/10.1127/0941-2948/2006/0130

556 Kyçyk, O., Aguilera, M. P., Gaforio, J. J., Jiménez, A., & Beltrán, G. (2016). Sterol
557 composition of virgin olive oil of forty-three olive cultivars from the World Collection
558 Olive Germplasm Bank of Cordoba. *Journal of the Science of Food and Agriculture*, *96*,
559 4143–4150. https://doi.org/10.1002/jsfa.7616

560 Li, X., Flynn, J. D., & Wang, S. C. (2019). The Effects of Variety, Growing Region, and

561 Drought Stress on Fatty Acid and Sterol Compositions of California Olive Oil. *JAOCS, Journal of the American Oil Chemists' Society*, *96*(3), 215–230.
563 https://doi.org/10.1002/aocs.12192

564 Lukić, I., Carlin, S., Horvat, I., & Vrhovsek, U. (2019). Combined targeted and untargeted
565 profiling of volatile aroma compounds with comprehensive two-dimensional gas
566 chromatography for differentiation of virgin olive oils according to variety and
567 geographical origin. *Food Chemistry*, *270*, 403–414.
568 https://doi.org/10.1016/j.foodchem.2018.07.133

569 Miho, H., Díez, C. M., Mena-Bravo, A., Sánchez de Medina, V., Moral, J., Melliou, E., …
570 Priego-Capote, F. (2018). Cultivar influence on variability in olive oil phenolic profiles
571 determined through an extensive germplasm survey. *Food Chemistry*, *266*, 192–199.
572 https://doi.org/10.1016/j.foodchem.2018.06.002

573 Monasterio, R. P., Olmo-García, L., Bajoub, A., Fernández-Gutiérrez, A., & Carrasco-
574 Pancorbo, A. (2016). Potential of LC Coupled to Fluorescence Detection in Food
575 Metabolomics: Determination of Phenolic Compounds in Virgin Olive Oil. *International*
576 *Journal of Molecular Sciences*, *17*(10), 1627–1644.
577 https://doi.org/10.3390/ijms17101627

578 Olmo-García, L., Bajoub, A., Monasterio, R. P., Fernández-Gutiérrez, A., & Carrasco-
579 Pancorbo, A. (2017). Metabolic profiling approach to determine phenolic compounds of
580 virgin olive oil by direct injection and liquid chromatography coupled to mass
581 spectrometry. *Food Chemistry*, *231*, 374–385.
582 https://doi.org/10.1016/j.foodchem.2017.03.139

583 Olmo-García, L., Bajoub, A., Monasterio, R. P., Fernández-Gutiérrez, A., & Carrasco-
584 Pancorbo, A. (2018). Development and validation of LC-MS-based alternative
585 methodologies to GC – MS for the simultaneous determination of triterpenic acids and
586 dialcohols in virgin olive oil. *Food Chemistry*, *239*, 631–639.
587 https://doi.org/10.1016/j.foodchem.2017.06.142

588 Olmo-García, L., Fernández-Fernández, C., Hidalgo, A., Vílchez, P., Fernández-Gutiérrez,
589 A., Marchal, R., & Carrasco-Pancorbo, A. (2019). Evaluating the reliability of specific
590 and global methods to assess the phenolic content of virgin olive oil: Do they drive to
591 equivalent results? *Journal of Chromatography A*, *1585*, 56–69.
592 https://doi.org/10.1016/j.chroma.2018.11.031

593 Olmo-García, L., Polari, J. J., Li, X., Bajoub, A., Fernández-Gutiérrez, A., Wang, S. C., &
594 Carrasco-Pancorbo, A. (2018). Deep insight into the minor fraction of virgin olive oil by
595 using LC-MS and GC-MS multi-class methodologies. *Food Chemistry*, *261*, 184–193.
596 https://doi.org/10.1016/j.foodchem.2018.04.006

597 Olmo-García, L., Wendt, K., Kessler, N., Bajoub, A., Fernández-Gutiérrez, A., Baessmann,
598 C., & Carrasco-Pancorbo, A. (2019). Exploring the capability of LC-MS and GC-MS
599 multi-class methods to discriminate virgin olive oils from different geographical
600 indications and to identify potential origin markers. *European Journal of Lipid Science*
601 *and Technology*, 1800336. https://doi.org/10.1002/ejlt.201800336

602 Purcaro, G., Barp, L., Beccaria, M., & Conte, L. S. (2015). Fingerprinting of vegetable oil

603 minor components by multidimensional comprehensive gas chromatography with dual
604 detection. *Analytical and Bioanalytical Chemistry*, *407*(1), 309–319.
605 https://doi.org/10.1007/s00216-014-8140-x

606 Sánchez de Medina, V., Priego-Capote, F., & Luque de Castro, M. D. (2015).
607 Characterization of monovarietal virgin olive oils by phenols profiling. *Talanta*, *132*,
608 424–432. https://doi.org/10.1016/j.talanta.2014.09.039

609 Sayago, A., González-Domínguez, R., Urbano, J., & Fernández-Recamales, Á. (2019).
610 Combination of vintage and new-fashioned analytical approaches for varietal and
611 geographical traceability of olive oils. *LWT-Food Science and Technology*, *111*, 99–104.
612 https://doi.org/10.1016/j.lwt.2019.05.009

613 Servili, M., Esposto, S., Taticchi, A., Urbani, S., Di Maio, I., Sordini, B., & Selvaggini, R.
614 (2014). The effect of diverse agricultural and technological factors on olive oil quality
615 and yield. *Acta Horticulturae*, *1057*, 603–618.
616 https://doi.org/10.17660/ActaHortic.2014.1057.77

617 Tena, N., Wang, S. C., Aparicio-Ruiz, R., García-González, D. L., & Aparicio, R. (2015). In-
618 Depth Assessment of Analytical Methods for Olive Oil Purity, Safety, and Quality
619 Characterization. *Journal of Agricultural and Food Chemistry*, *63*(18), 4509–4526.
620 https://doi.org/10.1021/jf5062265

621 Tsimidou, M., Sotiroglou, M., Mastralexi, A., Nenadis, N., García-González, D., & Gallina
622 Toschi, T. (2019). In House Validated UHPLC Protocol for the Determination of the
623 Total Hydroxytyrosol and Tyrosol Content in Virgin Olive Oil Fit for the Purpose of the
624 Health Claim Introduced by the EC Regulation 432/2012 for "Olive Oil Polyphenols."
625 *Molecules*, *24*(6), 1044. https://doi.org/10.3390/molecules24061044

626 Türkekul, B., Günden, C., Abay, C., & Miran, B. (2010). Competitiveness of mediterranean
627 countries in the olive oil market. *New Medit*, *9*(1), 41–46.

628 Uceda, M., & Hermoso, M. (1998). La calidad del aceite de oliva. In D. Barranco, R.
629 Fernández-Escobar, & L. Rallo (Eds.), *El Cultivo del Olivo* (pp. 547–572). Madrid,
630 Spain: Ediciones Mundi-Prensa.

631 United States Department of Agriculture. (2010). *United States Standards for Grades of Olive*
632 *Oil and Olive-Pomace Oil*. *Federal Register*. Retrieved from
633 https://www.ams.usda.gov/sites/default/files/media/Olive_Oil_and_Olive-
634 Pomace_Oil_Standard%5B1%5D.pdf

635 Vazquez, A., Sanchez-Rodriguez, E., Vargas, F., Montoro-Molina, S., Romero, M., Espejo-
636 Calvo, J. A., … Mesa, M. D. (2019). Cardioprotective Effect of a Virgin Olive Oil
637 Enriched with Bioactive Compounds in Spontaneously Hypertensive Rats. *Nutrients*, *11*,
638 1728.

639 Vossen, P. (2007). Olive Oil: History, Production, and Characteristics of the World's Classic
640 Oils. *HortScience*, *42*(5), 1093–1100.

641 Wabaidur, S. M., AlAmmari, A., Aqel, A., AL-Tamrah, S. A., Alothman, Z. A., & Ahmed, A.
642 Y. B. H. (2016). Determination of free fatty acids in olive oils by UPHLC–MS. *Journal*
643 *of Chromatography B*, *1031*, 109–115. https://doi.org/10.1016/j.jchromb.2016.07.040

644 Zhang, L., Wang, S., Yang, R., Mao, J., Jiang, J., Wang, X., … Li, P. (2019). Simultaneous
645 determination of tocopherols, carotenoids and phytosterols in edible vegetable oil by
646 ultrasound-assisted saponification, LLE and LC-MS/MS. *Food Chemistry*, *289*, 313–
647 319. https://doi.org/10.1016/j.foodchem.2019.03.067

648

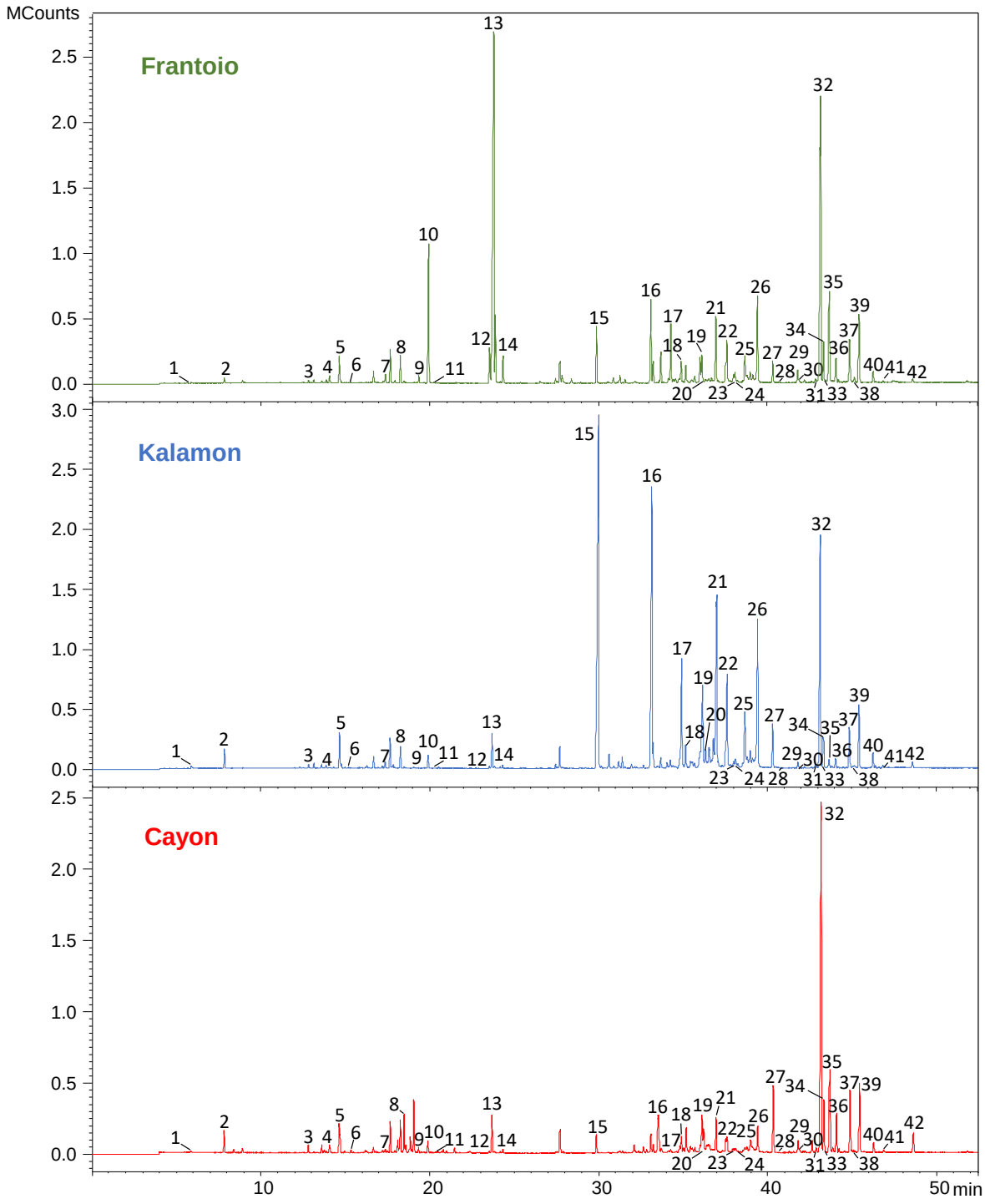

Fig. 1. Total ion chromatograms (TICs) of three extracts of monovarietal VOO showing great variability within the profiles. Peak identification numbers: 1, Van; 2,TY; 3, HTY; 4, AcHTY; 5, EA I; 6, Qui; 7, *p*-Cou; 8, EA II; 9, C16:1; 10, C16:0; 11, Fer; 12, C18:2; 13, C18:1; 14, C18:0; 15, DLA; 16, DOA; 17, LigAgly I; 18, squalene; 19, LigAgly II; 20, δ-Toc; 21, LigAgly III; 22, OleAgly I; 23, β-Toc; 24, γ-Toc; 25, OleAgly II; 26, OleAgly III; 27, α-Toc; 28, Api; 29, Cam; 30, Sti; 31, Lut; 32, β-Sit; 33, Pin; 34, Δ^5-Ave; 35, AcPin; 36, CyArten; 37, MeCyArtan; 38, ER; 39, Cit; 40, OA; 41, BA; 42, MA.

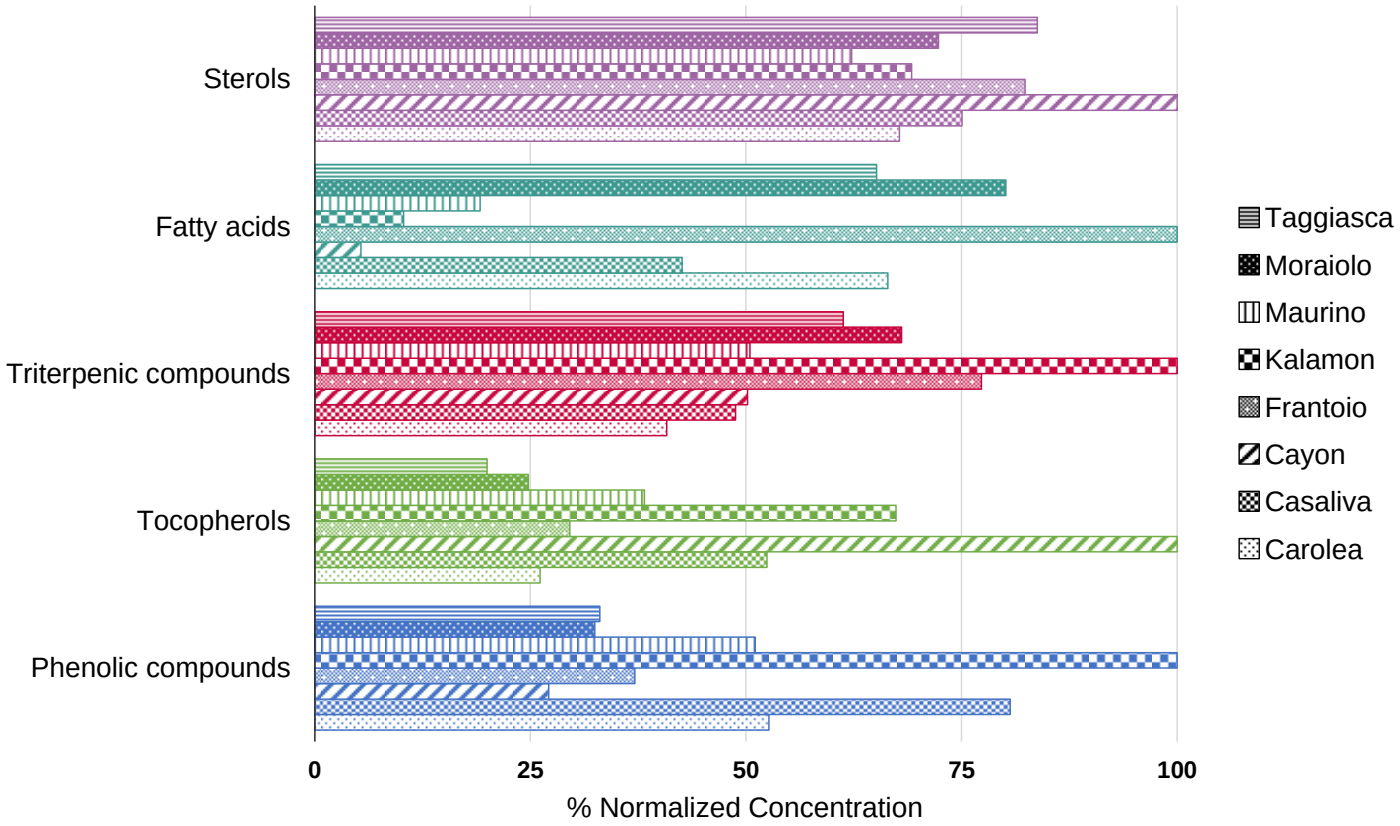

Fig. 2. Bars diagram comparing the average VOO minor compounds content in the eight different cultivars tested within this study, grouped by chemical class. Normalization was used considering the concentration level of the "richest" cultivar (for each chemical class) as 100% and referring the rest to that value. The compounds considered as members of each chemical class are detailed in Table 2.

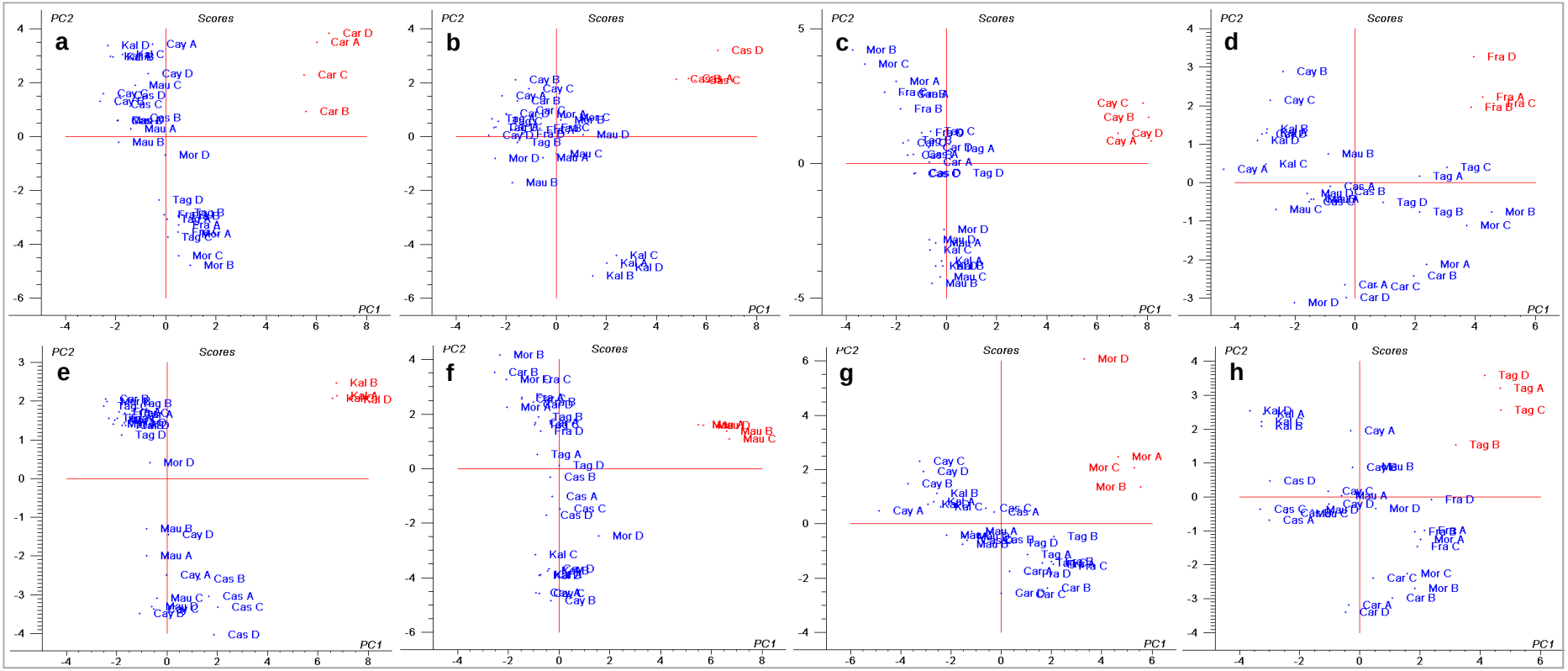

Fig. 3. Scores plot (PC1 *vs.* PC2) for the eight two-class models obtained applying PLS-DA to discriminate Carolea (a), Casaliva (b), Cayon (c), Frantoio (d), Kalamon (e), Maurino (f), Moraiolo (g) and Taggiasca (h) samples from the rest of the sample set. Abbreviations: Car (Carolea), Cas (Casaliva), Cay (Cayon), Fra (Frantoio), Kal (Kalamon), Mau (Maurino), Mor (Moraiolo) and Tag (Taggiasca).

Table 1a. Analytical parameters of the GC-MS method evaluated for the 25 available pure standards.

Rt (min)	Compound	quantifier *m/z*	qualifier *m/z*	Calibration function	R^2	*Intra-day* Repeatability (% RSD)[a]		*Inter-day* Repeatability (% RSD)[a]		Trueness (% recovery)[b]
						Area	*Rt*	*Area*	*Rt*	
6.9	Van	194	209, 224	$y = 2178\,x - 924$	0.998	2.73	<0.01	4.92	0.03	101.2
7.8	TY	179	193, 267, 282	$y = 8751\,x - 4173$	0.999	2.66	0.02	6.66	0.04	93.1
12.8	HTY	267	179, 193, 370	$y = 13{,}800\,x - 13{,}681$	0.998	2.11	0.02	4.01	0.02	91.7
15.3	Qui	346	256, 419	$y = 2605\,x - 1528$	0.997	4.15	0.03	4.56	0.03	97.3
17.2	*p*-Cou	294	308, 250, 219	$y = 6523\,x - 3211$	0.998	5.78	0.01	7.51	0.02	99.2
19.3	C16:1	312	129, 117, 326	$y = 6473\,x - 5741$	0.996	2.11	0.01	5.26	0.02	98.7
20.8	Fer	338	324, 294, 294	$y = 4250\,x - 1650$	0.992	2.70	0.01	4.85	0.02	100.3
23.5	C18:2	338	129, 340	$y = 4551\,x - 5846$	0.996	3.96	<0.01	5.54	0.02	93.8
23.8	C18:1	354	117, 129, 356	$y = 7990\,x - 3218$	0.995	5.36	<0.01	8.03	<0.01	90.2
36.6	*δ*-Toc	475	209, 249	$y = 7622\,x - 1002$	0.998	2.04	<0.01	5.70	0.01	90.5
37.9	*β*-Toc	489	223, 41	$y = 5633\,x - 4043$	1.000	4.49	0.01	7.63	0.01	86.6
38.1	*γ*-Toc	489	223, 43	$y = 7217\,x - 6405$	0.999	5.09	0.01	6.62	0.01	82.9
40.3	*α*-Toc	503	238, 43	$y = 7644\,x - 2011$	0.999	2.68	0.02	5.10	0.02	81.3
40.4	Api	472	399, 486	$y = 603\,x - 542$	0.988	4.88	0.03	5.86	0.03	98.9
41.8	Cam	503	472, 383	$y = 7026\,x - 16668$	0.996	5.22	0.01	7.82	0.01	78.6
42.2	Sti	395	485, 256	$y = 1094\,x - 2228$	0.996	3.01	0.03	5.43	0.03	80.7
42.6	Lut	560	472, 574	$y = 398\,x - 361$	0.989	1.25	<0.01	5.01	0.04	82.4
43.1	*β*-Sit	397	358, 486, 381	$y = 21{,}289\,x - 39{,}661$	0.998	5.40	0.01	9.19	0.01	75.1
43.2	Pin	502	223, 235, 488	$y = 1332\,x - 406$	1.000	5.86	0.01	6.44	0.05	102.3
45.4	ER	497	216, 203	$y = 2916\,x - 8770$	0.998	4.41	0.01	5.29	0.02	97.3
45.9	UV	497	216, 203	$y = 2410\,x - 5474$	0.993	3.66	0.01	7.33	0.01	98.5
46.3	OA	203	585, 483, 320	$y = 2648\,x - 4981$	0.994	5.76	0.02	7.49	0.02	99.2
46.6	BA	189	585, 320, 483	$y = 1308\,x - 2443$	0.993	2.11	0.01	6.75	0.01	101.3
47.0	UA	320	585, 203, 483	$y = 2595\,x - 5711$	0.988	3.66	0.03	7.33	0.03	102.5
48.6	MA	203	571, 320, 391	$y = 1198\,x - 3360$	0.991	4.67	0.02	8.40	0.03	99.8

[a] Repeatability is expressed as the RSD (%) of peak area values for four injections of four different extracts of the QC carried out within the same sequence (*intra*-day) or over four days (*inter*-day).

[b] Trueness, expressed as recovery (%), was estimated by analyzing the QC extracted before and after the standard addition and calculating the difference between the obtained results. The values included in this table are those achieved for the intermediate concentration level to contain the size of the table.

Table 1b. LOD, LOQ and matrix effect of the standards herewith reported for the first time.

Rt (min)	Compound	LOD (mg/L)	LOQ (mg/L)	Matrix Effect Coef. (%)[c]
19.3	C16:1	0.08	0.27	-6.6
23.8	C18:1	0.08	0.27	16.3
36.6	*δ*-Toc	0.04	0.15	10.3
41.8	Cam	0.05	0.17	-36.2
42.2	Sti	0.59	1.96	1.0
43.1	*β*-Sit	0.04	0.14	-3.7
47.0	UA	0.79	2.63	0.6

[c] Matrix effect coefficient (%) = [1−(slope matrix/slope solvent)]×100.

Table 2. Average concentration of the 41 determined compounds (mg/kg of VOO) in four samples of each cultivar. Results are given in mean value ± SD; SD expresses the *intra*-cultivar variability.

		Carolea	**Casaliva**	**Cayon**	**Frantoio**	**Kalamon**	**Maurino**	**Moraiolo**	**Taggiasca**
Phenolic compounds	Van	0.41 ± 0.03	0.42 ± 0.04	0.46 ± 0.09	0.89 ± 0.08	0.45 ± 0.01	0.69 ± 0.02	0.52 ± 0.03	0.69 ± 0.14
	TY	1.8 ± 0.2	3.8 ± 0.7	11 ± 5	2.7 ± 0.5	9 ± 1	1.1 ± 0.6	1.0 ± 0.1	2.1 ± 0.2
	HTY	1.28 ± 0.05	2.7 ± 0.5	1.8 ± 0.5	1.7 ± 0.3	2.0 ± 0.1	1.5 ± 0.5	1.7 ± 0.1	1.7 ± 0.2
	AcHTY	0.41 ± 0.02	0.63 ± 0.09	3 ± 1	2.7 ± 0.7	0.51 ± 0.03	0.16 ± 0.03	0.44 ± 0.07	3 ± 1
	Qui	0.8 ± 0.1	2 ± 2	1.5 ± 0.2	0.9 ± 0.1	0.89 ± 0.04	0.60 ± 0.03	0.44 ± 0.02	1.2 ± 0.4
	p-Cou	7.6 ± 0.5	0.49 ± 0.02	1.00 ± 0.04	0.54 ± 0.01	1.2 ± 0.1	0.28 ± 0.03	0.6 ± 0.1	0.17 ± 0.01
	Fer	0.10 ± <0.01	0.11 ± 0.01	nq	0.11 ± 0.01	0.12 ± 0.01	nq	nq	nd
	DLA	4 ± 1	126 ± 25	29 ± 4	22 ± 2	356 ± 27	1.6 ± 0.1	22 ± 9	22 ± 4
	DOA	3.9 ± 0.6	48 ± 7	16 ± 5	21 ± 3	90 ± 5	3.8 ± 0.4	38 ± 12	23 ± 3
	LigAgly I	15 ± 6	62 ± 6	6 ± 3	6 ± 2	55 ± 9	5.2 ± 0.8	2.3 ± 0.8	8 ± 3
	LigAgly II	11 ± 2	24 ± 2	1.6 ± 0.3	7.4 ± 0.9	33 ± 3	7 ± 3	2.5 ± 0.3	4 ± 2
	LigAgly III	27 ± 5	53 ± 2	5 ± 4	26 ± 10	74 ± 5	19 ± 5	10 ± 1	15 ± 6
	OleAgly I	28 ± 4	39 ± 4	6 ± 1	11 ± 2	41 ± 7	37 ± 6	14 ± 6	15 ± 4
	OleAgly II	9 ± 2	10 ± 1	0.8 ± 0.2	6 ± 3	11.7 ± 0.9	19 ± 4	2.9 ± 0.8	3 ± 2
	OleAgly III	41 ± 6	41 ± 3	3 ± 2	29 ± 12	39 ± 9	73 ± 17	20 ± 6	21 ± 7
	Api	12 ± 2	5.0 ± 0.2	9 ± 2	4.4 ± 0.2	4.6 ± 0.3	2.3 ± 0.6	4.2 ± 0.5	3.0 ± 0.4
	Lut	12 ± 2	5.5 ± 0.9	15 ± 4	6.5 ± 0.6	7.1 ± 0.7	5 ± 2	6 ± 1	5.7 ± 0.8
	Pin	9 ± 1	7.9 ± 0.5	14.3 ± 0.7	8.1 ± 0.2	9.2 ± 0.4	11 ± 2	4.3 ± 0.3	6.4 ± 0.4
	AcPin	4.1 ± 0.2	4.0 ± 0.2	6.5± 0.4	7.4 ± 0.3	2.62 ± 0.04	5.5 ± 0.9	1.8 ± 0.5	5.4 ± 0.9
Tocopherols	δ-Toc	nq	10.4 ± 0.4	11 ± 1	7.73 ± 0.02	8.0 ± 0.1	1.8 ± 0.2	1.9 ± 0.1	1.46 ± 0.06
	β-Toc	11.5 ± 0.4	15.2 ± 0.9	45 ± 8	12.7 ± 0.2	17.6 ± 0.7	6.7 ± 0.7	5.0 ± 0.5	6.5 ± 0.6
	γ-Toc	8.5 ± 0.4	42 ± 6	19 ± 5	9.6 ± 0.4	11.0 ± 0.6	15 ± 5	29 ± 4	5.9 ± 0.4
	α-Toc	112 ± 13	213 ± 32	460 ± 53	128 ± 19	324 ± 21	181 ± 49	96 ± 17	93 ± 21
Triterpenic compounds	ER	4.2 ± 0.1	3.5 ± 0.1	4.2 ± 0.4	5.11 ± 0.08	5.1 ± 0.2	2.9 ± 0.3	3.59 ± 0.08	5.1 ± 0.1
	UV	nd	nd	nd	nd	nd	nd	nd	nd
	OA	4.6 ± 0.7	7.2 ± 0.4	7.7 ± 0.8	14 ± 2	19 ± 4	9 ± 2	11 ± 1	11 ± 2
	BA	nd	nd	nd	nd	nq	nd	nq	nd
	UA	nd	nd	nd	nd	nd	nd	nd	nd
	MA	20 ± 2	24 ± 2	24 ± 6	35.9 ± 0.4	48 ± 1	23.8 ± 0.7	33.9 ± 0.9	28 ± 2
Sterols	Cam	12 ± 1	17 ± 2	22 ± 2	22 ± 3	11.0 ± 0.6	18 ± 2	18 ± 2	21 ± 2
	Sti	3.8 ± 0.5	2.8 ± 0.2	3.6 ± 0.5	4.6 ± 0.5	2.8 ± 0.2	2.17 ± 0.09	3.6 ± 0.1	4.4 ± 0.6
	β-Sit	280 ± 21	314 ± 40	452 ± 28	362 ± 66	287 ± 29	286 ± 48	330 ± 13	350 ± 22
	Δ^5-Ave	7.7 ± 0.5	37 ± 2	10 ± 1	15 ± 1	9 ± 1	20 ± 8	11 ± 4	15 ± 2
	CyArten	25 ± 4	20 ± 1	35 ± 7	25 ± 4	13 ± 2	9.9 ± 0.2	32 ± 10	32 ± 2
	MeCyArtan	60 ± 10	51 ± 6	61 ± 13	50 ± 5	69 ± 7	35 ± 4	38 ± 13	59 ± 7
	Cit	32 ± 1	25 ± 2	38 ± 8	34 ± 2	37 ± 3	17 ± 2	18 ± 3	39 ± 3
Free fatty acids	C16:1	3.1 ± 0.5	nq	1.9 ± 0.1	6 ± 1	nq	2.4 ± 0.3	4 ± 2	4.7 ± 0.7
	C16:0	355 ± 80	228 ± 35	28 ± 2	520 ± 108	49 ± 10	124 ± 22	498 ± 241	326 ± 52
	C18:2	40 ± 10	21 ± 3	1.9 ± 0.2	43 ± 6	3.1 ± 0.4	9 ± 2	28 ± 12	27 ± 4
	C18:1	364 ± 67	243 ± 33	26 ± 6	611 ± 67	64 ± 24	92 ± 20	411 ± 197	410 ± 57
	C18:0	63 ± 9	37 ± 4	8.4 ± 0.8	64 ± 7	11.2 ± 0.9	12 ± 2	55 ± 26	42 ± 7

nd: non detected / nq: non quantifiable

Table 3. Quality parameters of the two-class PLS-DA models (each cultivar compared with the rest) and most relevant distinctive features of each model.

Distinctive features (regression coefficients)	*p*-Cou (0.282) EA I (0.151) C18:2 (0.146) Api (0.144) EA II (0.109) MA (-0.103)	Δ^5-Ave (0.277) γ-Toc (0.237) LigAgly I (0.225) C16:1(-0.181) δ-Toc (0.130) MA (-0.130)	β-Toc (0.136) α-Toc (0.115) Lut (0.115) δ-Toc (0.094) β-Sit (0.090) TY (0.087)	Fer (0.784) δ-Toc (0.512) AcPin (0.354) C18:1 (0.337) *p*-Cou (-0.335) C16:1 (0.240)	DLA (0.147) MA (0.125) Δ^5-Ave (-0.118) DOA (0.117) γ-Toc (-0.117) OA (0.114)	OleAgly II (0.195) OleAgly III (0.191) Pin (0.170) OleAgly I (0.141) Van (0.114) DOA (-0.107)	γ-Toc (0.238) AcPin (-0.237) Cit (-0.220) Fer (-0.207) ER (-0.179) AcHTY (-0.172)	Fer (-0.350) ER (0.273) Cit (0.263) δ-Toc (-0.201) Δ^5-Ave (0.179) AcHTY (0.145)

- Offset: point where a regression line crosses the ordinate (Y-axis).
- R: covariance between the two variables divided by the square root of the product of their variances.
- R-Square: square of the correlation coefficient between predicted and measured values.
- RMSEP (Root Mean Square Error of Prediction): measurement of the average difference between predicted and measured response values, at the prediction or validation stage.
- SEP (Standard Error of Performance): standard deviation of the prediction residuals.
- Bias: average value of the residuals.