Postprint of LWT - Food Science and Technology Volume 88, February 2018, Pages 103-108
DOI: 10.1016/j.lwt.2017.09.047

Nondestructive measurement of squalene in olive oil by near infrared spectroscopy

José A. Cayuela[1], Juan F. García[2]
[1]Instituto de la Grasa, CSIC
Campus of the University 'Pablo de Olavide', Ed. 46. 41013 Seville, Spain
[2]Department of Chemical Engineering, University of Seville, C/ Profesor García González, 1, 41012 Seville, Spain

[1]Corresponding author: jacayuela@ig.csic.es

ABSTRACT

This study sets the basis for developing a rapid technique for measuring olive oil squalene, which is a healthy compound. This technique, based on near infrared spectroscopy, is environmentally friendly. The most suitable wavelength ranges were defined, studying the possible contribution from the visible spectra. For this purpose, Partial Least Squares analysis was independently set up using two optical arrangements, with wavelengths 350-2500 nm and 1100-2300 nm. Models from only near infrared wavelengths gave the best outcomes. The external validation exercise for estimating olive oil squalene was satisfactory, with r^2 0.83 and residual predictive deviation 2.31. The results suggest the proposed technique is useful for estimating olive oil squalene content. A sorting test of olive oil in two classes according to its squalene content was carried out, with threshold in 5.0 $g.kg^{-1}$, using the model built. The success of this classification was 90%.

Keywords: classification; environment friendly; NIR; PLS model; ; threshold.

1. Introduction

Squalene (Figure 1) is a triterpene aliphatic hydrocarbon, and it was named because of its profusion in shark liver oil, its richest source, where it reaches 900 $g \cdot kg^{-1}$. Shark liver oil has long been used as a traditional health food in Japan, with a particular benefit for vascular health (Hamadate et al., 2015). Squalene is widely distributed in nature, especially in vegetable oils such as olive oil, palm oil, wheat-germ oil, amaranth oil, or rice bran oil (Huang, Lin, & Fang, 2009). Therefore, by using different extraction

methods, vegetables or marine animals can be suitable squalene sources (Vázquez, Torres, Fornari, Senorans, & Reglero, 2007).

The main part of virgin olive oil is the saponifiable fraction, a lipid matrix of triglycerides, diglycerides and monoglycerides accounting for 985-995 $g \cdot kg^{-1}$ (Civantos, 1999). Squalene is in relatively high quantities within the olive oil minor fraction. Eisner, Iverson, Mozingo, & Firestone (1965) stated that squalene makes up around 85-90% of the hydrocarbon fraction of olive oils. Besides, it makes up 60–75% of the olive oil unsaponifiable fraction, in concentrations between 0.2 and 7.5 $g \cdot kg^{-1}$ (Tiscornia, & Evangelisti, 1982).

Figure 1

One of the most important differences between olive oil and plant seed oils is squalene. Compared to seed oils, olive oil is an important source of squalene. In other edible vegetable oils, squalene makes up only 0.02–0.3 $g \cdot kg^{-1}$ (Rao, Newmark, & Reddy, 1998). Thus, olive oil contains 7 to 300 fold more squalene than other vegetable oils and up to 5000 fold more than some vegetable foods (Liu, Ahrens, Schreibman, & Crouse, 1976). Therefore, virgin olive oil may be a part of the human diet especially rich in squalene.

Besides, the squalene content varies widely depending on the olive oil product with a range of 2 to 7 $g \cdot kg^{-1}$ (Rao, Newmark, & Reddy, 1998). A significant difference between the extra virgin class (EVOO) and the virgin class (VOO) has been reported, with the latter having more squalene than the refined olive oils (Owen, Mier, Giacosa, Hull, Spiegelhalder, & Bartsh, 2000). Nergiz & Çelikkale (2011) showed that refining reduces the squalene content. Furthermore, they pointed out that the major decrease in squalene in vegetable oils within the refining steps occurs during oils' deodorization. Olive growing techniques (Psomiadou & Tsimidou, 1999), olive fruit variety (Nergiz, & Ünal, 1990) and extraction (Nergiz, & Ünal, 1990; Samaniego-Sánchez, Quesada-Granados, López-García de la Serrana, & López-Martínez, 2010) influence the level of squalene. Squalene acts as a weak antioxidant in olive oil (Owen, Mier, Giacosa, Hull, Spiegelhalder, & Bartsh, 2000). Thus, Psomiadou and Tsimidou (1999) proposed that squalene contributes to olive oil stability in a small quantity, even at low temperatures.

There is multiple scientific evidence on the beneficial effects that the intake of squalene from food has on health (Newmark, 1997; Lasekan, Clayton, Gendron, & Ney, 1990; Smith, 2000; Ostlund, Racette, & Stenson, 2002; Strandberg, Tilvis, & Miettinen, 1990; Smith, 2000; He, & Corke, 2003). However, this feature of olive oil has received little attention in the market so far, since most consumers are unaware on this fact.

The olive oil industry has great interest on determining the quality of olive oil, using fast and reliable techniques. Besides, developing non-destructive techniques to reduce the expense of solvents and reagents is increasingly important in an international context of convergence towards environmental sustainability. Among the various non-destructive solutions to these needs, near-infrared spectroscopy (NIRS) has made major achievements. NIRS is based on multivariate models in which the spectral data correlate with the analyzed feature. It provides several important advantages, as NIRS needs no solvents or reagents, thus avoiding a major expense, while being environmentally friendly. Additionally, NIRS is a rapid, non-destructive, and potentially multi-parameter method.

Several articles on the use of NIRS and chemometrics for the analysis of different olive oil features have been published in the recent years (Nenadis & Tsimidou, 2017). Stand out studies directed to characterizing intact olives and olive paste for optimizing the milling process (Giovenzana et al., 2017), to control the quality of olive pomace oil blended with palm oil used for deep-frying (Ben Hammouda, Zribi, Ben Mansour, Matthaus, & Bouaziz, 2017), as well as for the authentication and detection of fraud (Karunathilaka, Kia, Srigley, Chung, & Mossoba, 2016). Sorting olive oil based on alpha-tocopherol and total tocopherol content using NIRS has been recently reported (Cayuela & García, 2017). The NIRS ability to analyze the major olive oil quality features has been the subject of several studies (Armenta, Garrigues, & De la Guardia, 2007; Bendini, Cerretani, Di Virgilio, Belloni, Lercker, & Gallina-Toschi, 2007; Conte, Brussolo, Pizzale, Carazzolo, Meurens, & Pavan, 2003; Costa, Coelho, Gambarra, Bezerra, Harrop, & Ugulino, 2008; Cayuela, Moreda & García, 2013). In fact, NIRS techniques are methods for these routine analyses in a growing number of laboratories. However, the possibility of measuring squalene in olive oil by NIRS has never been reported up to date. Besides, squalene NIR absorption bands have not yet been described, to the best of our knowledge.

Since the concentration of squalene varies widely among different olive oils, there is an interest in the development of rapid techniques to distinguish olive oils according to its content. In fact, the industry might have an interest in separating olive oils according to different squalene contents. The traditional method for the analysis of squalene in olive oil is GC. However, it is not usually performed in the olive oil industry, since squalene is not considered in the regulation to characterize the quality or purity of olive oil (European Commission, 1991). Therefore, there is a challenge on characterizing olive oil regarding squalene. This work sets up the basis for developing new rapid NIRS techniques for measuring olive oil squalene content. It was convenient to clarify if there are any regions from the olive oil's visible spectrum contributing to model performance, since pure squalene is a pale yellow liquid. The wavelengths that contribute to predictive models have been defined.

2. Material and Methods

2.1. Olive Oils

A total set of 180 olive oil samples was made up from different origins. High quality Extra Virgin Olive Oils (EVOO) were bought at olive oil specialized shops; this group contributed with 32 samples, of which 27 were varietal and the remaining 5 were mixtures from different varieties. These EVOO were used to elaborate 17 additional coupage samples. Olive oils normally found in the market were also used; this group was composed of 10 EVOOs, 40 Current Olive Oils and 25 Pomace Olive Oils. Olive oil samples were provided also from a collaborator industry, contributing with 14 EVOOs, 25 Virgin Olive Oils and 14 Lampante Olive Oils. The characteristics of the olive oil samples are shown in Table S1.

2.2. Reference Analysis

Squalene analysis were carried out by Gas Chromatography (GC) according to Lanzón, Albi, Cert, & Gracián (1994), modified according to Moreda, Pérez-Camino, & Cert (2004), at the Instituto de la Grasa (CSIC). Briefly, 0.1 mg of olive oil sample was disposed in a 4 mL screw vial, adding 1 mL of squalane 5 $mg \cdot mL^{-1}$ as the internal pattern. This was dissolved in heptane to complete a volume of 3 mL and shaked gently by hand. Then 200 μL of methanolic 2 $mol \cdot L^{-1}$ KOH was added, separating the aqueous and lipid phases. The upper phase was collected into a 2 mL chromatography vial and

then injected into the GC instrument. A GC HP-5890 (Hewlet Packard Enterprise, Palo Alto, USA) equipped with a split/splitless injection system was used with a SP-5 capillary column 5% phenylmethylsilicone fused silica, 30 m long, 0.25 mm internal diameter and 0.25 μm phase thickness, (Merck, Darmstadt, Germany). Flame ionization detector (FID) and software Chem Station for the recording and processing of data were used. The analyses were conducted with two replicates. The results were given with one significant digit.

2.3. Near infra-red spectroscopy

Optical arrangements NIRS and VIS/NIRS were used for defining the wavelengths contributing to the predictive models, especially for clarifying the contribution from visible spectra. Besides, using two different instruments allowed checking their results, beyond their comparison.

The samples' spectra were recorded directly from olive oils without any other treatment. The temperature of a body has an important influence on NIR radiation. Therefore, the samples were taken from 4 °C storage and placed at room temperature in the laboratory 18h before processing. A thermostatic bath fixed at 33 °C for 30 min held the 20 mL sample containers to ensure temperature stability. The averaged spectrum from two measurements of 50 spectra each was registered, with each sample. The same procedure was used with both optical configurations.

For NIRS, the measuring mode was post dispersive transflectance. A Luminar (Brimrose Inc., Maryland, USA) spectrometer was used. This instrument consists of an acousto-optic tunable filter (AOTF) with InGaAs detector (1100-2300 nm). The reference is automatically taken, the scanning speed is 60 ms. The spectrometer is composed of a hand-held unit, equipped with a base for laboratory use. A transflectance probe accessory was used. The probe is in stainless steel, with threaded interchangeable optical path. The spectra were registered as a whole, the spectral variables matching at 2 nm intervals. The repeatability of the instrument, expressed as the standard deviation of the average absorbance of five measurements of a white tile, is $6.76\ 10^{-4}$. The signals were captured using Acquire software (Brimrose Corp., Maryland).

The VIS/NIRS was carried out using a Labspec (Analytical Spectral Devices Inc., Boulder, USA) spectrometer, with transmittance optical mode consisting of a liquid

accessory (Ocean Optics, Largo, USA). A quartz spectrophotometric cuvette (Hellma Analytics, Müllheim, Germany) with 10 mm path length held the samples. The whole VIS/NIRS spectrum (350−2500 nm) was registered, with each spectral variable corresponding to a 2 nm interval. The configuration for 50 spectra in continuous acquisition was used. Indico Pro software (Analytical Spectral Devices Inc., Boulder, USA) was used for this purpose. The spectrometer was equipped with three detectors. The detector for the visible range (350-1000 nm) was a fixed reflective holographic diode array with a sensitivity of 512 pixels. A holographic fast scanner InGaAs detector cooled at -25ºC covered the wavelength range of 1000-1800 nm. The same device coupled with a high order blocking filter was used for the interval 1800-2500 nm. The scanning speed was 100 ms, and the acquisition processing time is less than a minute for each sample, all steps included. The repeatability, expressed as the standard deviation of the average absorbance of five measurements of a white tile between 350 and 2500 nm, is $6.00\ 10^{-4}$.

2.4. Squalene spectrum

Squalene of 98% purity (Merck, Darmstadt, Germany) was used as pattern to characterize the squalene VIS/NIR absorption bands. The pattern spectrum was registered only once using the Labspec spectrometer, by averaging four replicates. The rest of the procedure was the same as previously described for olive oil samples.

2.5. Chemometry

Possible olive oil groups were analyzed by Principal Component Analysis (PCA), which was also used for detecting possible spectral outliers. It was carried out from the olive oils spectra of both optical configurations using The Unscrumbler 9.7 (CAMO Software AS, Oslo, Norway).

Multivariate Partial Least Squares (PLS) analysis was performed from the spectral variables of near infrared (NIRS) and visible-near-infrared (VIS/NIRS), using the squalene reference analysis as a dependent variable. Transmittance spectral data were averaged to 8 nm intervals and transformed into absorbance, then, mean normalization (MN), standard normal variate normalization (SNV), and first (D_1SG) and second (D_2SG) Savitzsky−Golay derivatives treatments were carried out. The PLS models for squalene were built from the averaged and treated spectrum using The Unscrumbler 9.7.

The full cross internal validation (FCV) procedure was used. The outliers were identified as samples showing significant high residuals, according to The Unscrumbler 9.7. Scores plot were displayed from the regression overview plot, then selecting the warning list option. The outlier list was displayed by clicking the outliers button. The residuals are the differences between the predicted and the analyzed values.

Two independent multivariate calibration models for squalene prediction (M_1 and M_2) were established from the spectra obtained with NIRS and VIS/NIRS. The models' principal components (PCs) were fixed after the tests using 10 PCs at first.

The calibration set for PLS models excluded the external validation set. It was defined as one third from the 180 olive oil samples available, counting from the first. Sixty samples were taken by including one of each three, from the data base of The Unscrambler 9.7., in the same order as the samples were included. This validation set was randomly formed, since each sample was randomly registered from a completely independent olive oil batch, even when the mechanical selection by the software it was not random. The squalene concentrations range in the validation set were similar to that range of the calibration, as it is shown forward. Those wavelengths whose correlation with squalene content was closer to zero were removed in successive PLS cycles, using The Unscrambler 9.7. The variable selection ended in the last cycle that improved the squared coefficient of cross validation of the calibration (R^2_{CV}). This procedure provided the spectral variables selected for the PLS models. To assess model's fitness, the standard error of calibration (SEC) and the closeness between their squared coefficient of calibration (R^2) and R^2_{CV} were considered.

2.6. Model Performance Assessment

Calibration models were assessed by external validation exercises. For this purpose, squalene content was predicted in a previously reserved set formed by 60 olive oil samples which did not participate in the multivariate models. The model performance was assessed according to the r^2 from the external validation exercises, which corresponds to the simple linear regression between the analyzed and predicted values. At the same time, the residual predictive deviation (RPD) from the external validation exercise was considered. The RPD was defined (Fearn, 2002) as the ratio between the σ from the reference data of the validation set and the standard error of performance

(SEP). Also, the separate analysis of the calibrations of NIRS and VIS/NIRS allowed confirming their predictive ability.

2.7. Classification Tests

Classification tests of olive oil according to their squalene content were conducted by using the PLS model as a qualitative discrimination technique. The spectral data from only the configuration providing the best yields were used for this purpose. Two classes of olive oils, High Squalene (HS), with squalene concentration above 5.0 $g \cdot kg^{-1}$, and the other one Low Squalene (LS), with squalene concentration below or equal to 5.0 $g \cdot kg^{-1}$, were fixed for classification tests into two squalene levels. These two classes were defined according to the squalene mean of the total sample set analyzed, whose value was this threshold. The technique performance assessment was by its success grade. This was expressed as the percentage of samples in which the predicted and actual classifications coincided.

3. Results and Discussion

3.1. Olive Oil Spectrum

Near-infrared spectra show various overlapping bands, due to the first and second overtones and a combination of the fundamental vibrations, mainly carbon–hydrogen (Shenk, Workman, & Westerhaus, 2001). Assigning the major visible absorption bands of olive oil was done by Moyano, Meléndez, Alba, & Heredia (2008). Olive oil spectra from the samples analyzed in this work, shown in Figure 2, are consistent with the previously indicated reports. A first minor peak occured near 415 nm. This area suits the wavelengths of oil absorption for dark blue colored light. It could be due mainly to carotenoids, as well as to pheophytin A, pheophorbide A and pyropheophytin A. A second peak was near 450 nm, which corresponds to blue light absorption, a characteristic of carotenoids. A third peak appeared around 670 nm, which coincides with chlorophylls absorption (Moyano, Meléndez, Alba, & Heredia, 2008). A broad absorbance band showed around 1220 nm, probably due to second overtones of C–H and CH=CH– stretching vibrations of oil. A high intensity absorbance peak occured around 2300 nm, caused by a combination of fundamental vibrations of the C-H groups (Hourant, Baeten, Morales, Meurens, & Aparicio, 2000). The squalene VIS/NIR spectra registered in this study are shown in Figure 3. The major differences in the spectrum of

248 squalene with respect to olive oil correspond to the visible zone, since squalene is
249 almost colorless. On the contrary, in the NIR region the peaks with squalene were
250 practically the same as in olive oil, without remarkable differences. As a hydrocarbon
251 molecule, a part of the squalene C–H and CH=CH– stretching vibrations may be
252 overlapping those of triglycerides, whose structure is also carbon–hydrogen. However,
253 the spatial configuration of the squalene molecule is clearly different from that of
254 triglycerides due to its specific bonds. Moreover, the latter have oxygen, while squalene
255 lacks this chemical element. Therefore, NIR absorption intensities may reflect such
256 differences among different olive oils, despite the fact that they maybe not explicit in
257 the spectrum shape.

258 Figure 2 –Figure 3

259 *3.2. Population Characterization*

260 The values of the squalene reference analysis of the calibration and external validation
261 sets are gathered in Table 1. As can be seen, a wide squalene variation integrates into
262 the calibrations, ranging from 1.01 $g \cdot kg^{-1}$ to 10.15 $g \cdot kg^{-1}$. The statistical mean of the
263 calibration sets analyzed was 5.02 $g \cdot kg^{-1}$ of squalene. The squalene content in the olive
264 oils analyzed in this work showed a range even wider than those described in the
265 literature (Tiscornia & Evangelisti, 1982; Rao, Newmark & Reddy, 1998; Owen, Mier,
266 Giacosa, Hull, Spiegelhalder, & Bartsh, 2000). This range reflects the extent of
267 variation of squalene in this product, which accounts for the interest of distinguishing
268 olive oils that are helpful to health due to their high squalene content.

269 Table 1

270 *3.3. Principal Component Analysis*

271 The olive oils spectra PCA analysis of both spectrometers showed the absence of
272 sample groups. The PCA of NIRS, shown in Figure 4, stands out for showing three
273 samples widely separated from the major group. Two more samples appeared separated
274 from the major group, both consistent with the HS class. The remaining samples did not
275 show any consistent grouping trend. In the PCA of the VIS/NIRS spectra, shown in
276 Figure 5, only two samples appear clearly separated. One of these samples matches the

high squalene class and the other one matches the low squalene class. Therefore, the PCA analysis of spectra did not show sample groups.

Figure 4

Figure 5

3.4. Squalene PLS Models

The absorbance data treated by MN and D_1SG provided the best performance for squalene content calibrations with both configurations, whose statistics are gathered in Table 2. The same three samples which stood out in the PCA of the NIR spectra due to their wide separation from the major group, they were identified as outliers in each calibration and removed. The procedure for selecting the spectral variables provided a wide range of contributing wavelengths. However, many of them did not contribute positively to the PLS models, since their removal improved the calibration. The model M_1 provided R^2 0.86, R^2_{CV} 0.83 and SEC 0.88, while M_2 calibration statistics were R^2 0.76, R^2_{CV} 0.72 and SEC 1.19. The spectral windows and single wavelengths shared in the M_1 model are shown in Figure 6. It is worth mentioning the matching among several major wavelengths of M_1 with the major absorption area of the squalene spectra (Figure 3), corresponding to the 1700-1850 nm spectral window.

Figure 7 represents the regression for M_1. In these graphs, the horizontal line within a certain wavelength shows zero contribution to the model.

Table 2

Figure 6

Figure 7

3.5. External Validation Exercises Using M_1

One of the two samples separated from the major group in the PCA (Figure 4) belonged to the validation set. This was separated, since it appeared in the validation as an aberrant point, according to the criterion for outliers.. The statistics of the validation

exercise are shown in Table 2. The dispersion plot of V_1 is shown in Figure 8. The squalene models performance is shown by the r_1^2 and r_2^2 values, 0.83 and 0.74, from V_1 and V_2, respectively. The RPD values were 2.31 and 1.94 respectively for the same. According to Fearn (2002), predictive models with RPD values between 2 and 10 are suitable depending on the use they must carry out. Considering this, the V_1 prediction exercise using M_1 suggests that this model is suitable for estimating the squalene content in olive oils. Routine analysis of hydrocarbons in the olive oil industry is not frequent, as it is not compulsory, thus the prediction technique here proposed is useful for a preliminary characterization of olive oil on its squalene content.

Figure 8

3.6. Classification Tests

The M_1 PLS model was tested to distinguish two olive oil classes, HS and LS, according to the threshold 5.0 $g \cdot kg^{-1}$. This technique provided 89.8% success in distinguishing both classes. The results of this sorting test are shown in Table 3.

Table 3

4. Conclusions

The wavelengths that contribute to a PLS model for squalene prediction have been defined. The NIRS technique based on this PLS model has been proved useful for the rapid estimation of olive oil squalene content. This model was used successfully to separate olive oil into two classes according to squalene content. The new technique proposed here provides an opportunity for characterizing olive oils based on their squalene content. This information has nutritional interest for the consumers. Besides, it constitutes a diversification opportunity for the olive oil industry. The wavelengths defined may allow choosing a suitable instrument for this purpose.

According to the test results, an error of about 10% can be expected in the separation of olive oil into two classes. The use of the proposed technique for estimating the squalene content of olive oil requires the calibration of a specific spectrophotometer, as well as establishing a periodic validation protocol. The proposed technique is to be used with

olive oil directly, without neither solvents nor reagents, which makes it environmentally friendly.

Acknowledgments

This research was made within the project Recupera 2020 1.4.4. We are thankful to the European Regional Development Fund, as well to the Ministry of Economy and Competitiveness of Spain and the Spanish Council for Scientific Research, for funding this project.

Bibliography

Armenta, S., Garrigues, S., & De la Guardia, M. (2007). Determination of edible oil parameters by near infrared spectrometry. *Analitical Chemical Acta, 596,* 330−337.

Ben Hammouda, I., Zribi, A., Ben Mansour, A., Matthaus, B. & Bouaziz, M. (2017). Effect of deep-frying on 3-MCPD esters and glycidyl esters contents and quality control of refined olive pomace oil blended with refined palm oil. *European Food Research and Technology 243*, 7, 1219-1227.

Bendini, A., Cerretani, L., Di Virgilio, F., Belloni, P., Lercker, G., & Gallina-Toschi, T. (2007). In-process monitoring in industrial olive mill by means of FT-NIR. *European Journal of Lipid Science and Technology*, *109,* 498−504.

Cayuela, J. A. Moreda, W., & García, J. M. (2013). Rapid determination of olive oil oxidative stability and its major quality parameters using Vis/NIR transmittance spectroscopy. *Journal of Agricultural and Food Chemistry*, *61*, 8056−8062.

Cayuela, J. A., & García, J. F. (2017). Sorting olive oil based on alpha-tocopherol and total tocopherol content using Near-Infra-Red Spectroscopy (NIRS) analysis. *Journal of Food Engineering, 202,* 79-88.

Civantos, L. (1999). Obtención del aceite de oliva. (2nd ed.). Madrid: Editorial Agrícola Española, S.A. (Chapter 2).

European Commission (1991). Commission Regulation No 2568/91 of 11 July 1991 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis. *Official Journal of the European Communities*, No L 248/1.

Conte, L. S., Brussolo, G., Pizzale, L., Carazzolo, A., Meurens, M., & Pavan, O. (2003). Application of near infrared reflectance analysis to olive oil production quality control. *Rivista Italiana delle Sostanze Grasse, 80,* 213−217.

Costa, A. F., Coelho, M. J., Gambarra, F. F., Bezerra, S. R., Harrop, R. K., & Ugulino, M. C. (2008). NIR spectrometric determination of quality parameters in vegetable oils using PLS and variable selection. *Food Research International, 41,* 341−348.

Eisner, J., Iverson, J. L., Mozingo, A. K., & Firestone, D. (1965). Gas Chromatography of unsaponifiable matter. III. Identification of hydrocarbons, aliphatic alcohols, tocopherols, triterpenoid alcohols, and sterols present in olive oils. *Journal of the Association of Official Analitycal Chemists 48,* 417-433.

Fearn, T. (2002). Assesing calibration. *NIR News, 13,* 12-14.

Giovenzana, V., Beghi, R., Civelli, R., Trapani, S., Migliorini, M., Cini, E., ... & Guidetti, R. (2017). Rapid determination of crucial parameters for the optimization of milling process by using visible/near infrared spectroscopy in intact olives and olive paste. *Italian Journal of Food Science, 29,* 357-369.

Hamadate N., Matsumoto, Y., Seto, K., Yamamoto, T., Yamaguchi. H., Nakagawa, T., … & Yazawa, K. (2015). Vascular effects and safety of supplementation with shark liver oil in middle-aged and elderly males. *Experimental and Therapeutic Medicine, 10,* 641-646.

Huang, S. R., Lin, Y. K., & Fang, J. Y. (2009). Biological and Pharmacological Activities of Squalene and Related Compounds: Potential Uses in Cosmetic Dermatology. *Molecules, 14,* 540-554.

Hourant, P., Baeten, V., Morales, M. T., Meurens, M., & Aparicio, R. (2000). Oil and fat classification by selected bands of near-infrared spectroscopy. *Applyed Spectroscopy, 54,* 1168−1174.

Karunathilaka, S. R., Kia, A. R. F., Srigley, C., Chung, J. K., Mossoba, M. M. (2016). Nontargeted, Rapid Screening of Extra Virgin Olive Oil Products for Authenticity

Using Near-Infrared Spectroscopy in Combination with Conformity Index and Multivariate Statistical Analyses. *Journal of Food Science, 81*(10), C2390-C2397.

Lanzón, A., Albi, T., Cert, A., & Gracián, J. (1994). The hydrocarbon fraction of virgin olive oil and changes resulting from refining. *Journal of the American Oil Chemists Society, 71,* 285-291.

Lasekan, J. B., Clayton, M. K., Gendron, A., & Ney, D. M. (1990). Dietary olive and safflower oil in promotion of DMBA-induced mammary tumorigenesis in rats. *Nutrition and Cancer, 13,* 153–163.

Liu, G. C. K., Ahrens, E. H., Schreibman, P. H. & Crouse, J. R. (1976). Measurement of squalene in human tissues and plasma: validation and application. *Journal of Lipid Research 17,* 38 –45.

Moreda, W., Pérez-Camino, M. C., & Cert, A. (2004). Analysis of neutral lipids: Unsaponifiables. In: M. Leo and L. Nollet (Eds.), Handbook of Food Analysis: Physical characterization and nutrient analysis, 2nd Edition (pp. 320-324). New York: Marcel Dekker.

Moyano, M. J., Meléndez, A. J.; Alba, J., & Heredia, F. J. (2008). A comprehensive study on the colour of virgin olive oils and its relationship with their chlorophylls and carotenoids indexes (I): CIEXYZ non-uniform colour space. *Food Research International, 41,* 505−512.

Nenadis, N., Tsimidou, M. (2017). Perspective of vibrational spectroscopy analytical methods in on-field/official control of olives and virgin olive oil. *European Joural of Lipid Science and Technology, 119,* 1600148. http\\doi.org.10.1002/ejlt.201600148.

Nergiz, C., & Ünal, K. (1990). The effect of extraction systems on triterpene alcohols and squalene content of virgin olive oil. *Grasas y Aceites, 41,* 117–121.

Nergiz, C., & Çelikkale, D. (2011). The effect of consecutive steps of refining on squalene content of vegetable oils. *Journal of Food Science and Technology, 48,* 382–385.

Newmark, H. L. (1997). Squalene, olive oil, and cancer risk: A review and hypothesis. *Cancer Epidemiology, Biomarkers & Prevention, 6,* 1101-1103.

Owen, R. W., Mier, W., Giacosa, A., Hull, W. E., Spiegelhalder, B., & Bartsh, H. (2000). Phenolic compounds and squalene in olive oils: the concentration and antioxidant potential of total phenols, simple phenols, secoiridoids, lignans and squalene. *Food and Chemical Toxicology, 38,* 647-659.

Ostlund, R., Jr, M. D., Racette, S. B., & Stenson, W. F. (2002). Effects of trace components of dietary fat on cholesterol metabolism: phytosterols, oxysterols, and squalene. *Nutrition Reviews, 60,* 349-359.

Psomiadou, E., & Tsimidou, M., (1999). On the Role of Squalene in Olive Oil Stability. *Journal of Agricultural and Food Chemistry, 47,* 4025-4032.

Rao, C. V., Newmark, K., & Reddy, B. S. (1998). Chemopreventive effect of squalene on colon cancer. *Carcinogenesis, 19,* 287–290.

Samaniego-Sánchez, C., Quesada-Granados, J. J., Lopez-Garcıa de la Serrana, H., & Lopez-Martınez, M. C. (2010). β-Carotene, squalene and waxes determined by chromatographic method in picual extra virgin olive oil obtained by a new cold extraction system. *Journal of Food Composition and Analysis, 23,* 671–676.

Shenk, J. S.; Workman, J. J.; Westerhaus, M. O. (2001). Application of NIR spectroscopy to agricultural products. In: D. A. Burns, and C. W. Ciurcak (Eds.), Handbook of Near Infrared Analysis, 2nd Edition (pp. 419−474). New York: Marcel Dekker.

Smith, T. J. (2000). Squalene: potential chemopreventive agent. *Expert Opinion on Investigational Drugs, 9,* 1841-1848

Strandberg, T. E., Tilvis, R. S., & Miettinen, T. A. (1990). Metabolic variables of cholesterol during squalene feeding in humans: comparison with cholestyramine treatment. *Journal of Lipid Research 31,* 1637-1643.

Tiscornia, E. F. M., & Evangelisti, F. (1982). Chemical compsqualenetion of olive oil and its variations induced by refining. *Rivista Italiana delle Sostanze Grasse*, 59, 519–56.

Rao, C. V., & Reddy, B. S. (1993). Modulating effect of amount and types of dietary fat on ornithine decarboxylase, tyrsqualenene protein kinase and prostaglandins production during colon carcinogenesis in male F344 rats. *Carcinognesis, 14,* 1327–1333.

Vazquez, L., Torres, C. F., Fornari, T., Senorans, F. J., & Reglero, G. (2007). Recovery of squalene from vegetable oil sources using countercurrent supercritical carbon dioxide extraction. *Journal of Supercritical Fluids, 40,* 59–66.

Figure captions

Figure 1. Squalene structure.
Figure 2. Olive oil visible and near infrared spectra from the samples analyzed.
Figure 3. Visible and near infrared spectra of the squalene pettern (purity 98.0%).
Figure 4. Principal Component Analysis of the near infrared spectra (1100-2300 nm) from the analyzed olive oils.
Figure 5. Principal Component Analysis of the visible and near infrared spectra (350-2500 nm) from the analyzed olive oils.
Figure 6. Spectral variables contributing to the squalene near infrared model (M_1).
Figure 7. Partial Least Squares quantitative squalene model (M_1).
Figure 8. Prediction exercise of the olive oil squalene content (V_1).

Figure 1

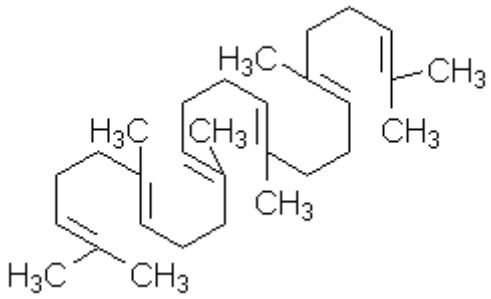

Figure 2

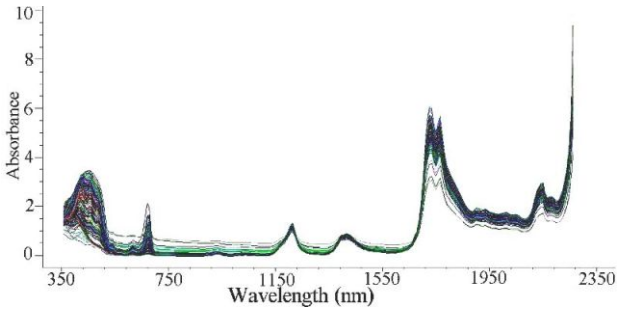
Absorbance
10
8
6
4
2
0
350
750
1150
1550
1950
2350
Wavelength (nm)

Figure 3

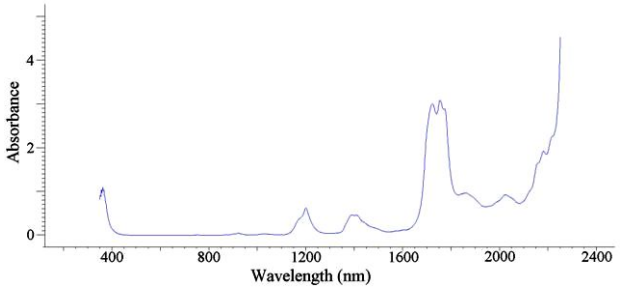

Figure 4

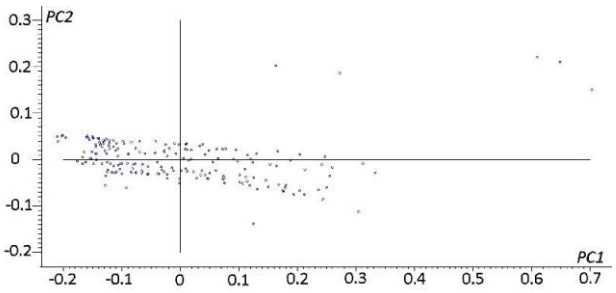

Figure 5

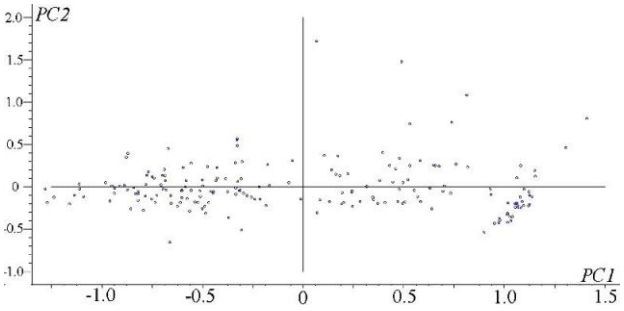

Figure 6

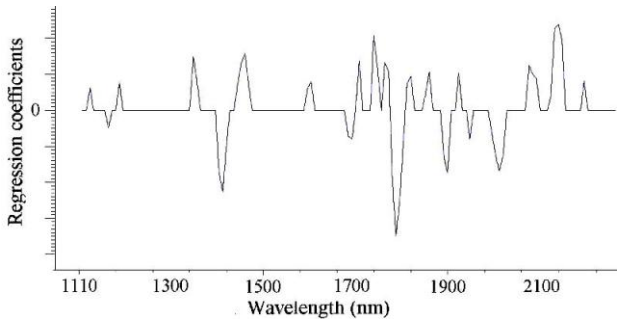
Regression coefficients
0
1110
1300
1500
1700
1900
2100
Wavelength (nm)

Figure 7

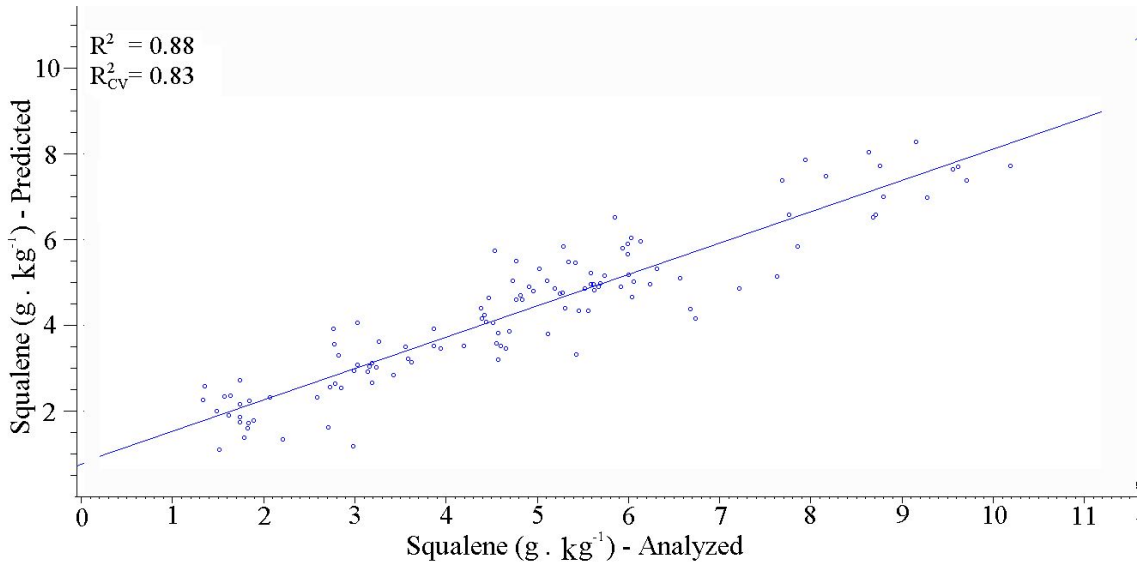
R² = 0.88
R²CV= 0.83
Squalene (g . kg-1) - Predicted
Squalene (g . kg-1) - Analyzed
0
1
2
3
4
5
6
7
8
9
10
11
2
4
6
8
10

Figure 8

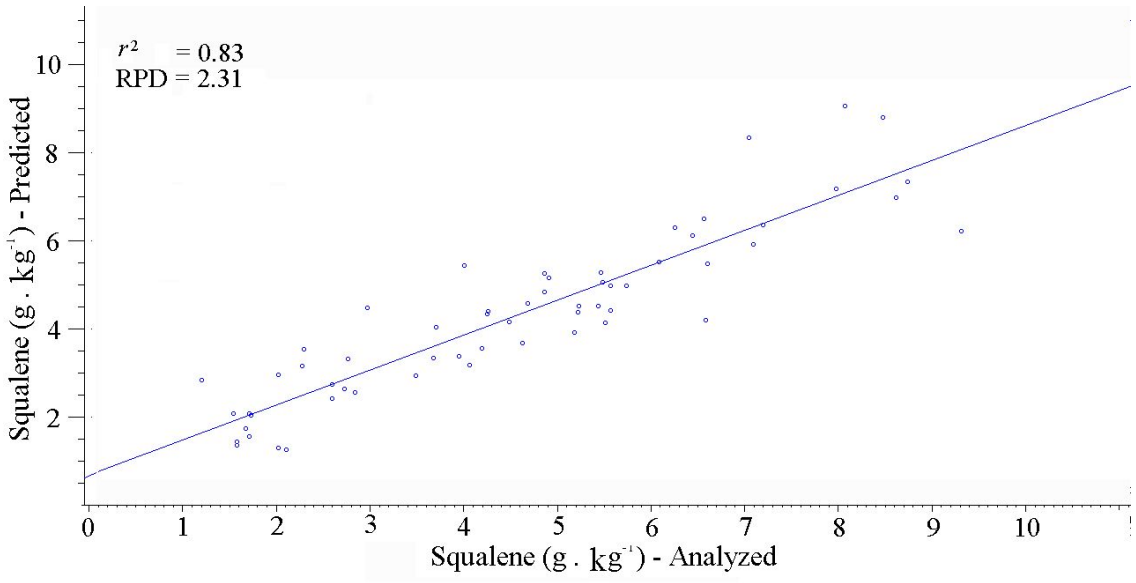
r^2 = 0.83
RPD = 2.31
Squalene (g . kg^{-1}) - Predicted
Squalene (g . kg^{-1}) - Analyzed
0
1
2
3
4
5
6
7
8
9
10

Table 1

N	$\bar{x}$	σ	Range
Calibration			
118	5.10	2.19	1.01-10.15
Validation			
59	4.88	2.33	1.22-10.02

Table 1. Statistics of the olive oil squalene content ($g \cdot kg^{-1}$) of the calibration and external validation sets. N, sample set size; $\bar{x}$, mean; σ, standard deviation.

Table 2

Calibrations			*External validations*		
	NIRS - M_1	VIS/NIRS - M_2		NIRS - V_1	VIS/NIRS - V_2
N	118	118	N	59	59
PC	6	4	SEP	1.01	1.20
R^2	0.86	0.76	r^2	0.83	0.74
R^2_{CV}	0.83	0.72	RPD	2.31	1.94
SEC	0.88	1.19			

Table 2. Statistics of PLS models of the olive oil squalene content

NIRS, near infrared (1100-2300 nm); VIS/NIRS, visible and near infrared (350-2500 nm); M_1, NIRS model; M_2, VIS/NIRS model; V_1, NIRS model validation; V_2, VIS/NIRS model validation; N, size; PC, principal components; R^2, squared calibration regression coefficient; R_{CV}, squared calibration coefficient of cross validation; SEC, standard error of calibration; SEP, standard error of performance; r^2, squared validation regression coefficient; RPD, residual predictive deviation.

Table 3

Class	Actual	Predicted
HS	30	31
LS	29	28
Total	59	59

Table 3. Olive oil sorting test according to squalene content using M_1 model (number of samples actual and predicted of each olive oil class).

HS, olive oil 'High squalene' (squalene > 5 $g.kg^{-1}$); LS, olive oil 'Low squalene' (squalene ≤ 5 $g.kg^{-1}$).

Table S1

Nº	*Olive oil class*	*Commercial names*	*Variety, blend, coupage, or acidity*
1		Acanto	
2		Castillo de Canena	
3		Ñ Organic	
4		Castillo de Tabernas	
5		Melgarejo Picual	Picual
6		Spirito Santo	
7		Rincón de la Subbética	
8		La Solana 2	
9		Marqués de Griñón	
10		Románico	
11		Alma Oliva	
12		L'Estornell	Arbequina
13		Basilipo	
14		Melgarejo Arbequina	
15		La Torre	
16		La Cultivada	Hojiblanca
17		Melgarejo Hojiblanca	
18		La Española	
19		Vega Oliva	Manzanillla
20		El Lagar del Soto	Manzanilla Cacereña
21		Changlot Real	Changlot Real
22		Acrópolis	Koroneiki
23	EVOO	Supremo	Arbosana
24		Morellana	Picudo
25	*Gourmet trade*	De Ortegas	Cornicabra
26		Melgarejo Frantoio	Frantoio
27		Oleo Aureo	Pico Limón
28		Duque de Baena	Undefined
29		Almenara Premium	Undefined
30		Molino de Gines	Undefined
31		Acanto + Supremo	
32		Spirito Santo + Rincón de la Subbética	
33		Spirito Santo + La Solana 2	Picual blend (½+½)
34		Castillo de Canena + La Torre	
35		Castillo de Tabernas + Castillo Canena	
36		L'Estornell +Alma Oliva	
37		L'Estornell + Basilipo	
38		L'Estornell + Melgarejo Arbequina	Arbequina blend (½+½)
39		L'Estornell + Románico	
40		La Torre + La Cultivada	
41		La Torre + Melgarejo Hojiblanca	Hojiblanca blend (½+½)
42		La Cultivada + Melgarejo Hojiblanca	
43		La Española + Vega OLiva	Manzanilla blend (½+½)
44		El Lagar del Soto + La Española	Manzanilla blend (½+½)
45		Núñez de Prado + Acrópolis	Picual+Picuda+Hojiblanca+Manzanilla

46		Changlot Real + Acrópolis	Changlot Real + Koroneiki (½+½)
47		Melgarejo Frantoio + Oleo Aureo	Frantoio + Pico Limón(½+½)
48		Molino de Gines + Almenara	Undefined (½+½)
49		Morellana + Melgarejo	Picudo + undefined (½+½)
50		De Ortegas + Oleo Aureo	Cornicabra + Pico Limón (½+½)
51		Núñez de Prado	Picual, Picuda, and Hojiblanca
52		Melgarejo Coupage	Coupage (undefined)
53	EVOO	La Masía	Undefined
54		Carrefour	
55		La Española	
56		Carrefour Picual	
57		5 Aceitunas	
58-62		La Masía, Carrefour, La Española, Carrefour Picual and 5 Aceitunas in different proportions	Blends
63-76	EVOO *Industry (*In bulk)	Undefined	Undefined
77-101	VOO *Industry (*In bulk)	Undefined	Undefined
102-142	OO	Alcampo	Acidity ≤ 0.4
		Alcampo	Acidity ≤ 1.0
		Abaco	Acidity ≤ 0.4
		Abaco	Acidity ≤ 1.0
		Ybarra	Acidity ≤ 0.4
		Alcampo, Abaco and Ybarra in different proportions	Blends
143-156	LOO *Industry (*In bulk)	Undefined	Undefined
157-180	POO	5 Aceitunas	Undefined
		Capicua (1 s)	
		Capicua (2 s)	
		La Masia (1 s)	
		La Masia (2 s)	
		5 Aceitunas, Capicua (1 s), Capicua (2 s), La Masia (1 s) and La Masia (2 s). In different proportions.	Blends

Table S1. Characteristics of the olive oil samples.

EVOO, extra virgin olive oil; VOO, virgin olive oil; OO, olive oil; LOO; lampante olive oil; POO, pomace olive oil; s, samples.