

1 **Cultivar discrimination, fatty acid profile and carotenoid characterization of monovarietal**
2 **olive oils by Raman spectroscopy at a single glance.**

3

4 **Silvia Portarena ^a, Chiara Anselmi ^a, Claudia Zadra ^b, Daniela Farinelli ^c, Franco Famiani ^c,**
5 **Chiara Baldacchini ^{a*}, Enrico Brugnoli ^a**

6 [a] Istituto di Biologia Agroambientale e Forestale, Consiglio Nazionale delle Ricerche, 05010,
7 Porano (TR), Italy

8 [b] Department of Pharmaceutical Sciences of the University of Perugia, 06121, Perugia, Italy.

9 [c] Department of Agricultural, Food and Environmental Sciences of the University of Perugia,
10 06121, Perugia, Italy.

11

12

13 *Corresponding author. E-mail address: chiara.baldacchini@ibaf.cnr.it. Full postal address: IBAF-
14 CNR, Via G. Marconi 2, 05010 Porano (TR), Italy.

15

16 **Abstract**

17 Monovarietal extra virgin olive oils (EVOOs) from *Olea europea* trees grown in the same
18 experimental orchard have been characterized by Raman spectroscopy combined with
19 chemometrics. Linear discriminant analysis of the EVOO Raman spectra correctly classified up to
20 94.4% of samples, as a function of the cultivar, at four different maturation stages. The sensitivity
21 of Raman spectroscopy to the oil cultivar is due to the genotypic dependence of the olive oil
22 composition in terms of both carotenoid and fatty acids. Thus, using gas chromatography as the
23 reference technique, models for predicting the oleic, palmitic and linoleic acid content are obtained
24 combining Raman spectra and partial least squares regression analysis. Evidences that a similar
25 approach could be implemented also for carotenoid characterization are provided.

26

27

28 Keywords: Olive oil, Raman spectroscopy, chemometrics, ripening, cultivar, fatty acids,
29 carotenoids.

1. Introduction

Extra virgin olive oil (EVOO) is a complex food that contains fatty acids (FAs), vitamins, pigments, volatile and anti-oxidative compounds, and it has received prominent attention during the last decades, due to the protective effect that its constituents exhibit against different types of cancer and cardiovascular diseases (Owen, Haubner, Wurtele, Spiegelhalder & Bartsch, 2004).

In particular, the oleic acid (OA), which is the most abundant FA in EVOOs, is a monounsaturated FA with 18 C atoms (C18:1) and it is known to low total cholesterol and low-density lipoprotein (LDL) levels and to raise high-density lipoprotein (HDL) (Sundram, Ismail, Hayes, Jeyamalar & Pathmanathan, 1997). It helps normalizing blood pressure (Teres et al., 2008), improving insulin levels and blood sugar regulation (Ryan et al., 2000). Positive effects on human health are also associated with linoleic acid (LA), a polyunsaturated FA with 18 C atoms (C18:2), which is known to reduce the levels of plasma total cholesterol and LDL (Risérus, Vessby, Arnlöv, & Basu, 2004). At the same time, carotenoids are well known as efficient quenchers of dangerous singlet oxygen and various reactive radicals by intercepting the chain of oxidative reactions. This function is directly related to the carotenoid conjugated double-bond (CDB) system that allows them to absorb the excess energy from other molecules (Britton, Liaaen-Jensen & Pfander, 2004). There is accumulating evidence that their antioxidative function is a key mechanism of protection against cancer, cardiovascular diseases, and macular degeneration in humans (Zhao & Sweet, 2008).

Thus, the EVOO quality in connection with healthy purposes mainly depends on its composition in terms of FAs and antioxidative compounds, whose determination is of primary importance. Such composition is generally the result of a complex combination of botanical, environmental and anthropogenic factors (Harwood & Ramón, 2000). Among these, the cultivar is emerging as one of the most relevant, since some olive cultivars are recognized to produce high-quality EVOOs as a consequence of the balance between mono- and polyunsaturated FAs and the bioactivity of minor components such as the antioxidant compounds, among which carotenoids (Giuffrida, Salvo, Salvo,

56 Cossignani & Dugo, 2011). As a consequence, monovarietal EVOO market is rapidly increasing
57 (Montealegre, Marina Alegre, & García-Ruiz, 2010), as well as the consequent adulteration and
58 market fraud (Breton, Claux, Metton, Skorski & Berville, 2004).

59

60 Current official EVOO quality control methods are based on the determination by chromatographic
61 means of the concentration of certain chemical components, e.g. FAs, sterols, alcohols or
62 stigmastadiene (Regulation 2568/91/EEC and amendments, in particular EU Reg. 299/2013), which
63 should belong to a specific range, between a maximum and a minimum limit. The chromatographic
64 results are supplemented with the measurement of quality parameters, such as peroxide value, free
65 acidity, and spectrophotometric values of K232, K270 and ΔK . However, these chemical indices are
66 unable to supply complete information on the quality of the product or to recognize frauds in EVOO
67 market. In particular, no official standard method is able to determine the olive cultivar, even if it is
68 the base of regional designation of origin covered by a protected designation of origin (PDO) or the
69 protected geographical indication (PGI). Therefore, there is a need to devise possible alternative,
70 rapid and reliable analytical methods verifying the authenticity of EVOO to assure the quality of
71 product and to protect consumers by both sanitary and commercial frauds.

72 In the last years, several spectroscopic techniques have been demonstrated to disclose the EVOO
73 biochemical fingerprint, even beyond the limit of commercial fraud, with high accuracy, such as
74 mid-infrared (MIR) (Bevilacqua, Bucci, Magrì, Magrì & Marini, 2012; Sinelli et al., 2010), near-
75 infrared (NIR) (Casale & Simonetti, 2014; Oliveri et al., 2011), nuclear magnetic resonance (NMR)
76 (Camin et al., 2016; Alonso-Salces et al., 2010), UV–vis spectroscopy (Domenici et al., 2014;
77 Casale et al., 2012; Kružlicová, Mocá, Katsoyannos & Lankmayr, 2008), and Raman spectroscopy
78 (Barthus & Poppi, 2001; Baeten, Dardenne & Aparicio, 2001; Muik, Lendl, Molina-Díaz & Ayora-
79 Cañada, 2003; El-Abassy, Donfack & Materny, 2009; Gouvinhas, Machado, Carvalho, De Almeida
80 & Barros, 2015; Portarena, Baldacchini & Brugnoli, 2017).

With respect to chromatographic techniques, spectroscopies have the advantages of requiring small sample size and no sample preparation (being thus low cost, fast, and non-destructive), and of being able to detect several characteristics within the same experimental output. In particular, Raman spectroscopy has emerged as a rapid tool to assess specific EVOO characteristics, such as the total unsaturation (Barthus & Poppi, 2001), free acidity (Baeten, Dardenne & Aparicio, 2001; Muik, Lendl, Molina-Díaz & Ayora-Cañada, 2003; El-Abassy, Donfack & Materny, 2009; Gouvinhas, Machado, Carvalho, De Almeida & Barros, 2015), peroxide value (Gouvinhas, Machado, Carvalho, De Almeida & Barros, 2015), and carotenoid content (El-Abassy, Donfack & Materny, 2009; Gouvinhas, Machado, Carvalho, De Almeida & Barros, 2015; Portarena, Baldacchini & Brugnoli, 2017).

In this work, Raman spectroscopy sensitivity to the EVOO FA and carotenoid content has been further investigated, in connection with the cultivar and the olive maturation stage, by acquiring Raman spectra from four monovarietal EVOOs similarly grown and harvested at four different dates. First of all, the total of the Raman data collected, analyzed within the framework of multivariate analyses, has allowed to obtain a clear distinction among the cultivars, with a correct classification of up to 94.4% of the samples. Moreover, using gas chromatography (GC) as the reference technique, models for predicting the FA profile have been developed from the same dataset, combining Raman spectra and partial least squares regression (PLS-R). Finally, preliminary evidences that similar predictive approaches can be applied also to the carotenoid composition have been obtained.

Thus, Raman spectroscopy is confirmed a powerful tool against the established standard methods of EVOO analysis, since it is rapid, low cost and requiring no sample preparation, and able to provide cultivar discrimination, FA profile and carotenoid characterization of EVOO from a single experimental output.

2. Materials and methods

107

108 **2.1 Sampling**

109 The *Olea europea* trees selected for the sampling campaign are grown in the experimental olive
110 orchard of the Department of Agricultural, Food and Environmental Sciences of the University of
111 Perugia, located nearby Perugia town (43°04'54.58"N, 12°22'53.41"E). The collection is placed on
112 a hill at 320 m a.s.l. The trees were 21 years old at the beginning of the trial, pruned according to
113 the vase training system, and spaced at 5 × 5 m. The orchard is rainfed. Fertilization is based on the
114 supply of nitrogen, potassium and phosphorous as chemical fertilizers. The soil is managed with a
115 spontaneous green cover mowed 2-3 times/year.

116 In the year of the experiment (2015), from July to October, main period of oil formation and
117 accumulation (Farinelli, Boco & Tombesi, 2002) the average maximum temperature was 28.4 °C,
118 the average minimum temperature was 16.4 °C
119 (<https://www.wunderground.com/history/airport/LIRZ/>) and the total rainfall was 208 mm
120 (Department of Agricultural, Food and Environmental Sciences of the University of Perugia).

121 Arbequina, Leccino, Maurino and Moraiolo cultivars were selected. The olives were randomly
122 hand-picked from three trees per cultivar, at four specific dates, corresponding to different stages of
123 the ripening process (SM1, Supplementary Material), resulting in total 48 different olive samples
124 (three per cultivar per harvest). EVOOs were then obtained from about 3.0 Kg of olives per sample.
125 Olives were crushed with a hammer mill and slowly mixed for 30 min at room temperature. Then,
126 the resulting paste was centrifuged at 3500 rpm for 3 min, without extra water.

127

128 **2.2 Chemicals**

129 Lutein (90% pure, Extrasynthese), β -carotene (99% pure, Alfa Aesar) and triolein (99%, Sigma
130 Aldrich) were provided by Carlo Erba. All the chemicals were used as purchased, without further
131 purification. Lutein and β -carotene solutions in triolein were prepared as reference standards for

132 carotenoid analysis in EVOOs (Domenici et al., 2014), at concentrations of $1.9 \times 10^{-5} \text{ M}$ and $2.9 \times 10^{-6} \text{ M}$, respectively.

134

135 **2.3 Raman spectroscopy data acquisition and treatment**

136 Raman spectra from the EVOO samples were obtained by Super Labram (Jobin Yvon) spectrometer
137 equipped with an Argon laser (514.5 nm wavelength), as previously described (Portarena,
138 Baldacchini & Brugnoli, 2017). In particular, same volume (100 μL) samples of selected EVOOs
139 were dripped on to glass microscope slides. The laser had an excitation power of less than 5 mW on
140 the sample stage and was focused within the drop using an optical microscope equipped with a 10x
141 objective. Optimal focusing within the oil drop was obtained by focusing first on the glass slide and
142 then moving the sample stage until the maximization of the typical EVOO Raman features and the
143 disappearing of the glass vibrational structures were simultaneously achieved. The spectra were
144 acquired in the $750\text{--}3050 \text{ cm}^{-1}$ range. Each spectrum was obtained by averaging five scans of 10 s
145 each. For each oil, three drops of about 1 ml were analyzed, by recording spectra from three
146 different points, therefore retrieving nine spectra (replicas) for each sample. The intensity variability
147 among spectra from the same sample was less than 5%.

148 Pre-treatment of the raw spectral data included fluorescence background subtraction (manually
149 performed by OriginPro 8.5, OriginLab, Northampton) and normalization of the nine spectra
150 representative of each sample at the intensity of the peak located at 1656 cm^{-1} , in order to correct
151 possible spectral differences due to oil film thickness, objective focusing or laser power.

152

153 **2.4 Chemical analysis**

154 The FA composition was determined as the corresponding methyl esters by following the procedure
155 described in the European Commission Regulation EEC no. 2568/91 (1991). The transmethylation
156 procedure, the GC with flame ionization detector (GC-FID) (Trace 2000, Thermo-Fisher), and the
157 analytical conditions have been previously described (Portarena et al., 2015). Relative amounts of

158 given FAs were calculated from their respective chromatographic peak areas, and the relative
159 percentage of each FA was related to the total peak areas of both saturated and unsaturated ones.

160

161 **2.5 Statistical treatment of data**

162 Raman spectral data can be considered as multidimensional sets of variables, where each input (i.e.,
163 the Raman spectrum) consists of hundreds of data, corresponding to the Raman intensities (values)
164 recorded at the different Raman shifts (variables). In order to reduce over-fitting due to a redundant
165 number of variables, the Raman spectra have been analyzed in the region between 900–1800 cm^{-1}
166 (El-Abassy, Donfack, & Materny, 2009).

167 Principal component analysis (PCA) was carried out to inspect the variability among samples.
168 Within spectral data, PCA summarizes the information contained in the Raman spectra variables in
169 a fewer independent principal components (PCs), obtained as linear combinations of the original
170 variables, lying in the direction of maximum variance (Lavine & Jerome, 2005).

171 The uncorrelated PCs mostly contributing to the variance of the data set were further subjected to
172 linear discriminant analysis (LDA), which is a supervised statistical method used to find a linear
173 combination of structures that characterize or separate classes of observations. Canonical Variates
174 (CVs) were obtained from LDA as the vectors minimizing the Wilks' lambda parameter, which is
175 calculated as the sum of squares of the distances between points belonging to the same category
176 divided by the total sum of squares. Values of Wilks' lambda approaching zero are obtained with
177 well resolved categories, whereas overlapped categories make Wilks' lambda to approach one
178 (Lerma-García, Ramis-Ramos, Herrero-Martínez, & Simó-Alfonso, 2010). To assess the quality of
179 the models retrieved from the LDA, cross-validation was applied. The dataset was divided into a
180 calibration (33 samples) and a cross validation (15 samples) sets 10 times, selecting each time
181 different sets of samples, with all the samples acting as the test items in turn. Then, the validation
182 results were averaged over the rounds to estimate a final predictive model. The LDA classification

183 matrixes are obtained by averaging over the 10 different calibration (CAL) and cross validation
184 (VAL) sets.

185 For quantitative analysis of EVOO composition in terms of FAs, Raman datasets were regressed to
186 the main FA contents using PLS-R method, applying cross-validation to the obtained models by
187 following the same procedure described for LDA. The resulted models were scrutinized by
188 regression coefficient (R^2 , where R is the correlation factor) and by the dimensionless root mean
189 square error of calibration (DRMSEC) and cross validation (DRMSECV), which permits to
190 compare the predictive capacity of models with the measured quantitative variables (FA relative
191 content), being better as DRMSEC(V) approaches zero (Sánchez-López et al., 2016).

192 PCA, LDA and PLS-R calculations were performed using Statistica v 8 (StatSoft Italia srl, Padua,
193 Italy).

194

195 **3. Results and discussion**

196

197 **3.1 Raman spectroscopy characterization of EVOOs**

198 Average Raman spectra of EVOOs are shown in Fig.1. The spectra present the same Raman
199 features, located at about 965 cm^{-1} , 1004 cm^{-1} , 1081 cm^{-1} , 1156 cm^{-1} , 1301 cm^{-1} , 1440 cm^{-1} , 1523
200 cm^{-1} , 1656 cm^{-1} , and 1748 cm^{-1} , in agreement with previous data as measured using a 514.5 nm
201 wavelength light (El-Abassy, Donfack, & Materny, 2009; Portarena, Baldacchini, & Brugnoli,
202 2017), *i.e.* in resonance with the vibrational modes of carotenoid molecules. The Raman spectrum
203 of triolein, the most abundant triglyceride in EVOO, is shown for comparison. Band assignment is
204 reported in Supplementary Material (SM2)

205 A strong variability is observed in the spectral intensity of the carotenoid related bands, which are
206 located at about 1004 cm^{-1} (C-CH₃ bending), 1156 cm^{-1} (C-C stretching) and 1523 cm^{-1} (C=C
207 stretching) (El-Abassy, Donfack & Materny, 2009; Portarena, Baldacchini & Brugnoli, 2017). The
208 spectral intensity of these bands strongly depends on the harvest date (Gimeno, Castellote,

209 Lamuela-Raventos, De la Torre & Lopez-Sabater, 2002), being more intense at the beginning and
 210 decreasing with the ripening, likely due to the lowering of olive biosynthetic activity (López-
 211 Sanchez, Ayora-Cañada & Molina-Díaz, 2010). Interestingly, the intensity decrease appears to be
 212 cultivar dependent (Fig.2a). Maurino displays the most intense carotenoid signals at the first and
 213 second harvests, and Leccino at the third and fourth ones. The less intense and less variable bands
 214 characterize Moraiolo at each ripening stage. This could be due to a different carotenoid
 215 quantitative and qualitative composition among the cultivars, since the two most abundant
 216 carotenoids in EVOO, *i.e.* lutein and β -carotene (Gandul-Rojas, Roca & Gallardo-Guerrero, 2016;
 217 Giuffrida, Salvo, Salvo, Cossignani & Dugo, 2011), are known to have rather different degradation
 218 pathways (Gandul-Rojas, Roca & Gallardo-Guerrero, 2016). In particular, β -carotene deteriorates
 219 faster than lutein (Czamara et al., 2015). As a consequence, the total carotenoid content of β -
 220 carotene-rich EVOOs is expected to decrease more rapidly with respect to that of lutein-rich ones.
 221 The bands at 965 cm^{-1} (C=C out of plane), 1081 cm^{-1} (C-C stretching), 1301 cm^{-1} (C-H twisting),
 222 1440 cm^{-1} (C-H scissoring), and 1656 cm^{-1} (*cis* C=C stretching) are attributed to FAs, both free and
 223 as triglycerides (El-Abassy, Donfack & Materny, 2009). These bands have similar spectral shape
 224 and intensity in the different monovarietal EVOOs, as well as the peak at 1748 cm^{-1} which is due to
 225 the C=O stretching of the ester linkage between FAs and glycerol (Weng, Weng & Tzeng, 2003).
 226 This suggest that the EVOO content of FAs, both free and as triglycerides, is less dependent from
 227 cultivar and maturation than the carotenoid one. However, a closer inspection to the FA related
 228 peak at 1656 cm^{-1} (Fig.2b) reveals differences in its intensity as a function of the cultivar. In
 229 particular, this peak is more intense in Arbequina and Maurino samples, and less intense in Leccino
 230 ones, almost disregarding to the harvest. The main FAs of EVOO (oleic, palmitic and linoleic acids)
 231 have similar vibrational modes, but with rather different relative spectral intensities (Czamara et al.,
 232 2015). For instance, the peak at about 1081 cm^{-1} is more intense in saturated FAs (such as palmitic
 233 acid; PA), while the peak at about 1656 cm^{-1} is peculiar of the unsaturated ones (OA and LA), with
 234 its intensity increasing with unsaturation. Consequently, differences in the Raman scattering relative

intensities could indicate variation of FA composition in EVOO. In particular, a high intensity of the 1656 cm^{-1} bands (as observed for Arbequina and Maurino samples) could indicate a high relative amount of LA.

3.2 Principal component and linear discriminant analyses for cultivar discrimination

PCA was performed on the complete spectral dataset (9 spectra per 48 samples), to investigate if the assessed sensitivity of Raman spectroscopy to carotenoid content and FA profile of EVOO could render the technique suitable for cultivar discrimination, beyond the limit of the intrasample variability.

Almost 90% of the total variance is explained by the first 9 PCs (SM3, Supplementary Material). The descriptors contributing the most in PC1 (45.4% of the total variance) are the spectral intensities of the bands at 1004 cm^{-1} , 1156 cm^{-1} and 1523 cm^{-1} , confirming that carotenoid content is the most discriminating factor (Portarena, Baldacchini & Brugnoli, 2017). Differentiation on PC2 axis (15.8% of the total variance) is mainly explained by the intensity of the bands at 1301 and 1656 cm^{-1} and, to a minor extent, from those at 1440 and 1081 cm^{-1} ; these vibrational modes being related with FAs.

The scatterplot of the EVOO samples within the bidimensional plane defined by PC1 and PC2 variables is shown in Fig.3. The spreading of the samples along PC1 is mainly driven by the harvesting date, with samples being distributed along a gradient from the first (PC1 most negative values) to the fourth (PC1 most positive values) harvest. Since PC1 is negatively correlated with the amount of carotenoids (SM3, Supplementary Materials), EVOOs from the first harvest are richer of carotenoid with respect to those from the fourth one, and that the carotenoid content is mainly dependent on the harvesting date. However, as previously stressed, the decreasing trend with maturation appears to be cultivar dependent, probably depending on the relative amount of lutein and β -carotene. In particular, according to PCA values, the first harvest of Maurino shows the

260 highest content of carotenoid and Moraiolo the lowest, while the fourth harvest of Leccino
261 preserves higher carotenoid content than any other cultivar.

262 Along PC2, EVOO samples cluster mainly in two groups, according to the cultivar: Leccino and
263 Moraiolo mostly gather at negative PC2 values, while Maurino and Arbequina at positive ones.
264 This clustering suggests that the composition in terms of FAs is driven more by the cultivar than by
265 the harvest date, and that Raman spectroscopy is highly sensitive to the EVOO FA composition.

266 In order to eliminate the dependence on the maturation stage, samples from the four harvests were
267 analyzed by PCA separately. For each harvesting date, about 100 PCs are obtained, with the first
268 three explaining almost the 70% of the total variance of the dataset (SM4, Supplementary Material).

269 At each harvest, the characteristics of the two first PCs recall what obtained by the total dataset:
270 PC1 discriminates the samples based on their carotenoid content, while the FA composition mainly
271 drives PC2. The scatterplots of the two first PC scores show a good clustering of the samples in
272 dependence of the cultivar, at each harvest (Fig. 4).

273 LDA analysis was further applied on these datasets, using the first 9 PCs from PCA, at each harvest,
274 as independent variables. About 100 spectra per harvest were used as calibration sets. Three
275 discriminant canonical variables (CVs) are extracted in all cases, accounting for 100.0% of the
276 variability, with the first two explaining more than 88.7% (SM5, Supplementary Material). The
277 calibration model derived by LDA correctly classifies from 83.3% to 94.4% of EVOO samples,
278 depending on the harvest (Table 1, upper panel). In particular, total correct predictions are obtained
279 for Moraiolo at harvest 1, 3 and 4, for Leccino at harvest 1 and 4, and for Maurino at harvest 2.

280 The validation procedure was undertaken with resort to the whole dataset used for the calibration,
281 from which one third of the replicas have been removed to be further classified as external set. In
282 this case, the percentage of right classification ranges from 78.8% to 93.9%, and total correct
283 predictions are obtained for the same cultivars at the same harvests, with respect to the calibration
284 models (Table 1, lower panel) confirming the stability of the model.

285 This result indicates that Raman spectroscopy can be used as a rapid tool for efficient cultivar
286 discrimination of EVOOs, likely due to its sensitivity to carotenoid content and FA composition.

287

288 **3.3 Fatty acid profile**

289 The EVOO FA profile was further determined by GC-FID (Table 2), which is the reference
290 technique for EVOO characterization, according to the International Olive Oil Council
291 (<http://www.internationaloliveoil.org>). In all samples, OA relative content is higher than the 70% of
292 the total FA content, while LA shows the lowest relative concentrations, being less than 10% of
293 total FAs. The relative content of PA ranges between 12.7% and 15.6%.

294 The four cultivars exhibit different FA compositions, probably reflecting the metabolic
295 characteristics of each genotype (Farinelli & Tombesi, 2015; Issaoui et al., 2010). Leccino has the
296 highest concentration of OA (with a moderate increasing trend from 78.8% at green stage to 80.7%
297 at full ripening) and the lowest amount of LA (with a moderate decreasing trend from 5.1% at green
298 stage to 4.4% at full ripening). OA and LA concentration in Moraiolo ranges from 75.7% to 76.2%
299 and from 7.1% to 7.8% respectively, as a function of maturation stage. Arbequina and Maurino
300 have the lowest concentration of OA and the highest percentage of PA and LA. In both cultivars,
301 OA and LA remain quite constant during ripening, while PA shows a moderate increasing of about
302 3%.

303 PLS-R models were carried out in order to determine a relationship between Raman scattering
304 intensities (predictor variables) and EVOO FA profile, in terms of OA, PA and LA percentages as
305 obtained by GC-FID (measured variables). The plots of the regressions showing the predicted
306 values of FA relative percentages based on Raman scattering intensities as a function of those
307 measured by GC are shown in Fig.5, and they clearly demonstrate that the Raman spectroscopy
308 sensitivity to the FA relative content of EVOOs is comparable with that of the reference technique.
309 The results achieved for the PLS-R models in terms of R^2 , DRMSEC and DRMSECV are presented
310 in the Supplementary Material (SM3). The high value of R^2 (ranging from 0.80 to 0.92 for the

311 calibration, and being about 0.68 for the cross validation) with DRMSEC lower than 0.08 and
312 DRMSECV lower than 0.16, indicate the good performance and precision of the models.

313

314 **3.4 Carotenoid composition**

315 The chemical structures of the two most abundant carotenoids in EVOO, namely lutein and β -
316 carotene, are rather similar, differing only for the terminal substituents. As a consequence, their
317 resonant Raman spectra show the same vibrational features, with the C=C stretching mode of lutein
318 being slightly more intense and shifted towards higher energy with respect to β -carotene (Baeten,
319 Dardenne & Aparicio, 2001; Andreeva & Velitchkova, 2005). On these bases, a further analysis of
320 the EVOO Raman spectra, beyond the strong variability due to the maturation stage, could provide
321 information on the carotenoid profile.

322 In Fig.6a, the average Raman spectra of the monovarietal EVOOs from the second harvest are
323 shown, after background subtraction and being normalized at the intensity of the C-C stretching
324 mode at 1156 cm^{-1} , which is not carotenoid relative content dependent (Andreeva, Apostolova, &
325 Velitchkova, 2011). The resonant Raman spectra of β -carotene and lutein reference solutions in
326 triolein are also shown, analogously background subtracted and normalized. The strong variability
327 of the FA related vibrational modes in the $1230\text{-}1500\text{ cm}^{-1}$ region of both EVOO and reference
328 samples is an artifact introduced by the normalization, due to the different total carotenoid content.

329 A zoomed view of the 1523 cm^{-1} region (C=C stretching mode) is shown in Fig.6b. According to
330 the literature, lutein C=C stretching mode has higher intensity and energy with respect to the
331 corresponding β -carotene vibration. The peak energy increases gradually from Leccino to Moraiolo,
332 suggesting that the relative lutein content increases as well. To put into evidence the differences
333 between the EVOO spectra, the peak positions have been plotted as a function of the intensity ratio
334 between the C=C and the C-C stretching modes (I_{1523}/I_{1156}) in Fig.6c. Here, a sort of calibration plot
335 is obtained, with Moraiolo being the lutein-richest cultivar (more high shifted peak, located at
336 1524.5 cm^{-1}), while the highest relative amount of β -carotene can be assigned to Leccino (less

intense peak, 1523 cm^{-1}). The lutein abundance of Moraiolo, together with its low variability during maturation (Fig.1), could descend from the early accumulation tendency of this cultivar (Portarena et al., 2015), being lutein the slowest degraded one (Gandul-Rojas, Roca & Gallardo-Guerrero, 2016). Analogously, the high relative amount of β -carotene shown by Maurino and Leccino, together with the high variability of the total carotenoid content of these cultivars with the ripening (Fig.1), could be consistent with a tardive oil accumulation process (Portarena et al., 2015). The peak position and intensity of the carotenoid C=C stretching mode in EVOO resonant Raman spectra could be then potentially used as quality markers, being able to assess the relative abundance of β -carotene and lutein.

346

347 **4. Conclusion**

A detailed Raman spectroscopy investigation of EVOOs obtained from olive fruits of monovarietal *Olea europea* trees grown in the same experimental orchard provided a reliable classification of oils according to the cultivar, based on PCA and LDA, with a correct classification up to 94.4%. Furthermore, the Raman intensity dependence of the FA bands on the EVOO FA profile allowed obtaining predictive models of these compounds, by using GC-FID as reference technique and PLS-R as statistical method. Finally, Raman markers to be used for the estimation of the lutein/ β -carotene content ratio, which is an emerging parameter for oil quality assessment, in connection with its healthy action, were identified. Then, information on cultivar, FA and carotenoid profiles of EVOOs can be obtained simultaneously, from the same Raman dataset. Thus, being rapid, low cost, non-destructive, user-friendly, requiring no sample preparation and small sample size, and it can be even performed on-site, Raman spectroscopy presents several practical advantages with respect to traditional techniques used for EVOO characterization (such as chromatographic techniques), and this will open interesting perspectives for its future application in quality control and authentication methods of olive oils and similar food products.

362

363 **Acknowledgements**

364 This work was supported by CURSA (Consorzio Universitario per la Ricerca Socioeconomica e per
365 l'Ambiente) [grant PALMO 2015/2017].

366 Thanks are due to Salvatore Cannistraro and Anna Rita Bizzarri from the Biophysics and
367 Nanoscience Centre of the University of Tuscia (Viterbo, Italy), for providing the Raman
368 spectroscopy facilities and support.

369

370 **References**

371

372 Alonso-Salces, R. M., Héberger, K., V. Holland, M., Moreno-Rjas, J. M., Mariani, C., Bellan, G.,
373 Reniero, F. & Guillou, C. (2005). Multivariate analysis of NMR fingerprint of the
374 unsaponifiable fraction of virgin olive oils for authentication purposes. *Food Chemistry*, 118,
375 956–965.

376 Andreeva, A., & Velitchkova, M. (2005). Resonance Raman spectroscopy of carotenoids in
377 Photosystem I particles. *Biophysical Chemistry*, 114(2–3), 129–135.
378 <https://doi.org/10.1016/j.bpc.2004.11.012>

379 Andreeva, A., Apostolova, I., & Velitchkova, M. (2011). Temperature dependence of resonance
380 Raman spectra of carotenoids. *Spectrochimica Acta - Part A: Molecular and Biomolecular*
381 *Spectroscopy*, 78(4), 1261–1265. <https://doi.org/10.1016/j.saa.2010.12.071>

382 Baeten, V., Dardenne, P., & Aparicio, R. (2001). Interpretation of fourier transform Raman spectra
383 of the unsaponifiable matter in a selection of edible oils. *Journal of Agricultural and Food*
384 *Chemistry*, 49(11), 5098–5107. <https://doi.org/10.1021/jf010146x>

385 Barthus, R. C., & Poppi, R. J. (2001). Determination of the total unsaturation in vegetable oils by
386 Fourier transform Raman spectroscopy and multivariate calibration. *Vibrational Spectroscopy*,
387 26(1), 99–105. Retrieved from <http://link.springer.com/article/10.1007/BF02540752>

388 Bevilacqua, M., Bucci, R., Magrì, A. D., Magrì, A. L., & Marini, F. (2012). Tracing the origin of
389 extra virgin olive oils by infrared spectroscopy and chemometrics: A case study. *Analytica*
390 *Chimica Acta*, 717, 39–51. <https://doi.org/10.1016/j.aca.2011.12.035>

391 Breton, C., Claux, D., Metton, I., Skorski, G., Berville, A. (2004). Comparative study of methods
392 for DNA preparation from olive oil samples to identify cultivars SSR alleles in commercial oil
393 samples: possible forensic applications. *Journal of Agricultural and Food Chemistry*, 52, 531–
394 537.

395 Britton, G., Liaaen-Jensen, S. & Pfander, H. (2004). *Carotenoids Handbook*, Birkhauser Verlag,

396 AG, Basel.

397 Camin, F., Pavone, A., Bontempo, L., Wehrens, R., Paolini, M., Faberi, A., Marianella, R.M.,
398 Capitani, D., Vista, S. & Mannina, L. (2016). The use of IRMS, ¹H NMR and chemical
399 analysis to characterise Italian and imported Tunisian olive oils. *Food Chemistry*, 196, 98-
400 105.

401 Casale, M., Oliveri, P., Casolino, C., Sinelli, N., Zunin, P., Armanino, C., Forina, M. & Lanteri, S.
402 (2012). Characterisation of PDO olive oil Chianti Classico by non-selective (UV-visible, NIR
403 and MIR spectroscopy) and selective (fatty acid composition) analytical techniques, *Analitica*.
404 *Chimica Acta*, 712, 56–63.

405 Casale, M. & Simonetti R. (2014). Review: Near infrared spectroscopy for analysing olive oils,
406 *Journal of Near Infrared Spectroscopy*, 22, 59-80.

407 Czamara, K., Majzner, K., Pacia, M. Z., Kochan, K., Kaczor, A., & Baranska, M. (2015). Raman
408 spectroscopy of lipids: A review. *Journal of Raman Spectroscopy*, 46(1), 4–20.
409 <https://doi.org/10.1002/jrs.4607>

410 Domenici, V., Ancora, D., Cifelli, M., Serani, A., Veracini, C.A., & Zandomeneghi M. (2014).
411 Extraction of pigment information from near-UV vis absorption spectra of extra virgin olive
412 oils, *Journal of Agricultural and Food Chemistry*, 62, 9317-9325.

413 El-Abassy, R. M., Donfack, P., & Materny, A. (2009). Rapid Determination of Free Fatty Acid in
414 Extra Virgin Olive Oil by Raman Spectroscopy and Multivariate Analysis. *Journal of the*
415 *American Oil Chemists Society*, 86(6), 507–511. <https://doi.org/10.1007/s11746-009-1389-0>

416 European Communities, Regulation 2658/01 of 11 July 1991 on the characteristics of olive oil and
417 olive-pomace oil and on the relevant methods of analysis.

418 Farinelli, D., Boco, M. & Tombesi, A. (2002). Intensity and growth period of fruit components of
419 olive varieties, *Acta Horticulturae*, 586, 607-610.

420 Farinelli, D. & Tombesi, S. (2015). Performance and oil quality of 'Arbequina' and four Italian olive
 421 cultivars under super high density hedgerow planting system cultivated in central Italy,
 422 *Scientia Horticulturae*, 192, 97-107 doi: 10.1016/j.scienta.2015.04.035

423 Fortes, C., García-Vilas, J. A., Quesada, A. R., & Medina, M. Á. (2012). Evaluation of the anti-
 424 angiogenic potential of hydroxytyrosol and tyrosol, two bio-active phenolic compounds of
 425 extra virgin olive oil, in endothelial cell cultures. *Food Chemistry*, 134(1), 134–140.
 426 <https://doi.org/10.1016/j.foodchem.2012.02.079>

427 Gandul-Rojas, B., Roca, M. & Gallardo-Guerrero, L. (2016). Chlorophylls and carotenoids in food
 428 products from olive tree, In *Products from Olive Tree*, 1st ed.; Boskou, D., Clodoveo, M.L.,
 429 Eds.; InTech: Rijeka, Croatia, 5, 67–98.

430 Gimeno, E., Castellote, A. I., Lamuela-Raventos, R. M., De la Torre, M. C., & Lopez-Sabater, M.
 431 C. (2002). The effects of harvest and extraction methods on the antioxidant content (phenolics,
 432 alpha-tocopherol, and beta-carotene) in virgin olive oil. *Food Chemistry*, 78(2), 207–211.
 433 [https://doi.org/10.1016/S0308-8146\(01\)00399-5](https://doi.org/10.1016/S0308-8146(01)00399-5)

434 Giuffrida, D., Salvo, F., Salvo, A., Cossignani, L., & Dugo, G. (2011). Pigments profile in
 435 monovarietal virgin olive oils from various Italian olive varieties. *Food Chemistry*, 124(3),
 436 1119–1123. <https://doi.org/10.1016/j.foodchem.2010.07.012>

437 Gouvínhas, I., Machado, N., Carvalho, T., De Almeida, J. M. M. M., & Barros, A. I. R. N. A.
 438 (2015). Short wavelength Raman spectroscopy applied to the discrimination and
 439 characterization of three cultivars of extra virgin olive oils in different maturation stages.
 440 *Talanta*, 132, 829–835. <https://doi.org/10.1016/j.talanta.2014.10.042>

441 Harwood, J., & Ramón, A., (2000). *Handbook of Olive Oil. Analysis and Properties*. Gaithersburg,
 442 MD, USA: Aspen Publishers.

443 Issaoui, M., Flamini, G., Brahmi, F., Dabbou, S., Ben Hassine, K., Taamali, A., Hechmi Chehab,
 444 H., Ellouza, M., Zarroukd, M. & Hammamia, M. (2010). Effect of the growing area

445 conditions on differentiation between Chemlali and Chétoui olive oils, *Food Chemistry*, 119,
 446 220–225.

447 Lavine, B.K. & Jerome, J. (2005). *Workman Chemometrics: Past, present, and future* B.K. Lavine
 448 (Ed.). *Chemometrics and chemoinformatics*. ACS symposium series, American Chemical
 449 Society, 894, 1-13.

450 Lerma-García, M. J., Ramis-Ramos, G., Herrero-Martínez, J. M., & Simó-Alfonso, E. F. (2010).
 451 Authentication of extra virgin olive oils by Fourier-transform infrared spectroscopy. *Food*
 452 *Chemistry*, 118(1), 78–83. <https://doi.org/10.1016/j.foodchem.2009.04.092>

453 López-Sanchez, M., Ayora-Cañada, M.J., Molina-Díaz, A. (2010). Olive Fruit Growth and
 454 Ripening as Seen by Vibrational Spectroscopy. *Journal of Agricultural and Food Chemistry*,
 455 58, 82–87.

456 Lozano-Sánchez, J., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2011). Characterisation of
 457 the phenolic compounds retained in different organic and inorganic filter aids used for
 458 filtration of extra virgin olive oil. *Food Chemistry*, 124(3), 1146–1150.
 459 <https://doi.org/10.1016/j.foodchem.2010.07.026>

460 Montealegre, C., Marina Alegre, M.L., García-Ruiz, C. (2010). Traceability markers to the
 461 botanical origin in olive oils. *Journal of Agricultural and Food Chemistry*, 58, 28–38.

462 Muik, B., Lendl, B., Molina-Díaz, A., & Ayora-Cañada, M. J. (2003). Direct, reagent-free
 463 determination of free fatty acid content in olive oil and olives by Fourier transform Raman
 464 spectrometry. *Analytica Chimica Acta*, 487(2), 211–220. [https://doi.org/10.1016/S0003-](https://doi.org/10.1016/S0003-2670(03)00560-9)
 465 [2670\(03\)00560-9](https://doi.org/10.1016/S0003-2670(03)00560-9)

466 Kružlicová, D., Mocá, K. J., Katsoyannos, E. & Lankmayr E. (2008). Classification and
 467 characterization of olive oils by UV-vis absorption spectrometry and sensorial analysis.
 468 *Journal of Food and Nutrition Research*, 47, 181–188.

469 Oliveri, P., Casale, M., Casolino, M. C., Baldo, M. A., Nizzi Grifi, F. & Forina, M. (2011)
 470 Comparison between classical and innovative class-modelling techniques for the

471 characterisation of a PDO olive oil. *Analytical and Bioanalytical Chemistry*, 399,
 472 2105–2113.

473 Owen, R.W., Haubner, R., Wurtele, G., Spiegelhalter, B., Bartsch, H. (2004). Olives and olive oil
 474 in cancer prevention. *European Journal of Cancer Prevention*, 13, 319-326.

475 Portarena, S., Baldacchini, C., & Brugnoli, E. (2017). Geographical discrimination of extra-virgin
 476 olive oils from the Italian coasts by combining stable isotope data and carotenoid content
 477 within a multivariate analysis. *Food Chemistry*, 215, 1–6.
 478 <https://doi.org/10.1016/j.foodchem.2016.07.135>

479 Portarena, S., Farinelli, D., Lauteri, M., Famiani, F., Esti, M., & Brugnoli, E. (2015). Stable isotope
 480 and fatty acid compositions of monovarietal olive oils: Implications of ripening stage and
 481 climate effects as determinants in traceability studies. *Food Control*, 57, 129–135.
 482 <https://doi.org/10.1016/j.foodcont.2015.03.052>

483 Risérus, U., Vessby, B., Arnlöv, J., & Basu, S. (2004). Effects of cis-9,trans-11 conjugated linoleic
 484 acid supplementation on insulin sensitivity, lipid peroxidation, and proinflammatory markers
 485 in obese men. *The American Journal of Clinical Nutrition*, 80(2), 279–83.
 486 <https://doi.org/80/2/279> [pii]

487 Ryan, M., McInerney, D., Owens, D., Collins, P., Johnson, A. & Tomkin, G.H. (2000). Diabetes
 488 and the Mediterranean diet: a beneficial effect of oleic acid on insulin sensitivity, adipocyte
 489 glucose transport and endothelium-dependent vasoreactivity, *QJM. An International Journal*
 490 *of Medicine* 93, 85-91.

491 Sánchez-López, E., Sánchez-Rodríguez, M. I., Marinas, A., Marinas, J. M., Urbano, F. J., Caridad,
 492 J. M., & Moalem, M. (2016). Chemometric study of Andalusian extra virgin olive oils Raman
 493 spectra: Qualitative and quantitative information. *Talanta*, 156–157, 180–190.
 494 <https://doi.org/10.1016/j.talanta.2016.05.014>

495 Sinelli, N., Casale, M., Di Egidio, V., Oliveri, P., Bassi, D., Tura, D., & Casiraghi, E. (2010).
 496 Varietal discrimination of extra virgin olive oils by near and mid infrared spectroscopy. *Food*

497 Research International, 43(8), 2126–2131. <https://doi.org/10.1016/j.foodres.2010.07.019>

498 Sundram, K., Ismail, A., Hayes, K. C., Jeyamalar, R. & Pathmanathan, R. (1997). Trans (Elaidic)
 499 fatty acids adversely affect the lipoprotein profile relative to specific saturated fatty acids in
 500 humans. *Journal of Nutrition*, 127, 514S–520S.

501 Teres, S., Barcelo-Coblijn, G., Benet, M., Alvarez, R., Bressani, R., Halver, J. E., & Escriba, P. V.
 502 (2008). Oleic acid content is responsible for the reduction in blood pressure induced by olive
 503 oil. *Proceedings of the National Academy of Sciences*, 105(37), 13811–13816.
 504 <https://doi.org/10.1073/pnas.0807500105>

505 Uceda, M. & Hermoso, M. (1998). La calidad del aceite de olive. In D. Barranco, R. Fernandez-
 506 Escobar, & L. Rallo (Eds.), *El Cultivo del Olivo*, Madrid, Spain: Junta de Andalucía Ediciones
 507 Mundi-Prensa, 547-572.

508 Weng, Y.M., Weng, R.H., Tzeng, C.Y. & Chen, W. (2003). Structural analysis of triacylglycerols
 509 and edible oils by near-infrared Fourier transform Raman spectroscopy. *Applied*
 510 *Spectroscopy*, 57, 413.

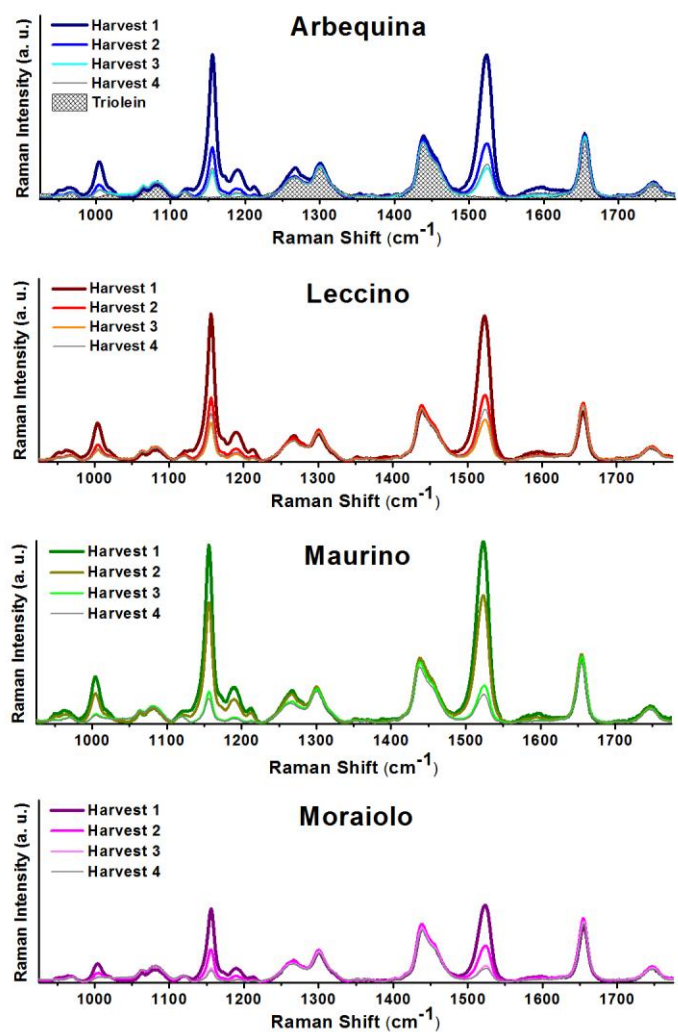
511 Zhao, L. & Sweet, B.V. (2008). Lutein and zeaxanthin for macular degeneration. *American Journal*
 512 *of Health-System Pharmacy*, 65, 1232–1238.

513

514

Highlights

- Monovarietal extra-virgin olive oils, at different maturation stages, are studied
- Correct cultivar discrimination of 94.4% of the samples is obtained
- Models for predicting fatty acid relative content are obtained
- Raman markers to assess the lutein/ β -carotene content ratio are identified



1

2

3

4

5

6

Figure 1. (Color online) Raman spectra of monovarietal Arbequina, Leccino, Maurino and Moraiolo from four harvest dates, averaged over three plants per cultivar. Average Raman spectra have been grouped by cultivar in the four panels. For comparison, the area corresponding to the Raman spectrum of the triolein standard sample taken in the same experimental conditions is patterned in the Arbequina panel.

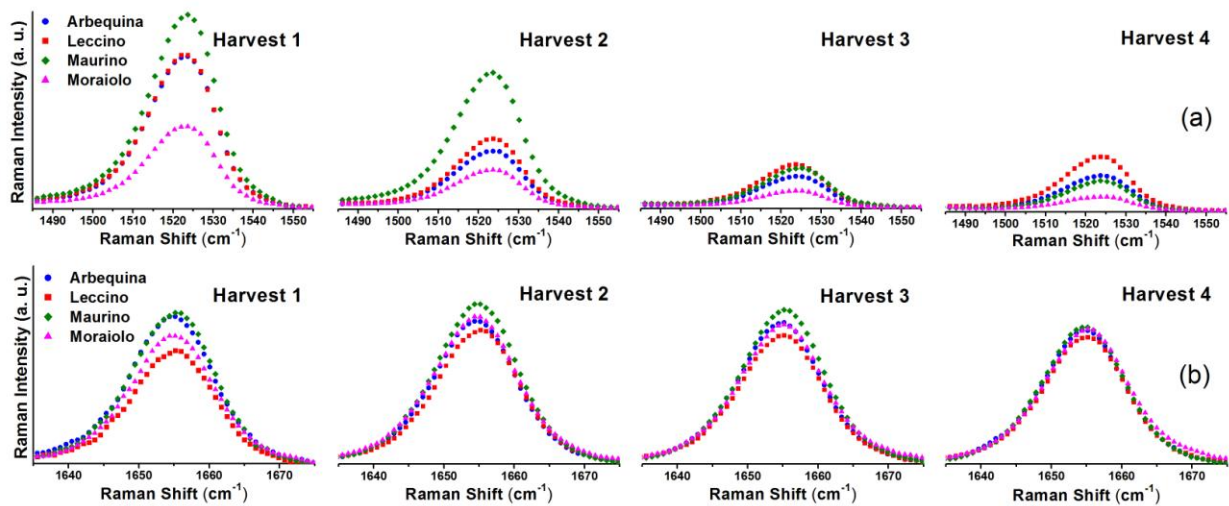


Figure 2. (Color online) Raman spectra of monovarietal Arbequina, Leccino, Maurino and Moraiolo from four harvest dates, averaged over three plants per cultivar. Average Raman spectra have been grouped by harvest in the four panels.

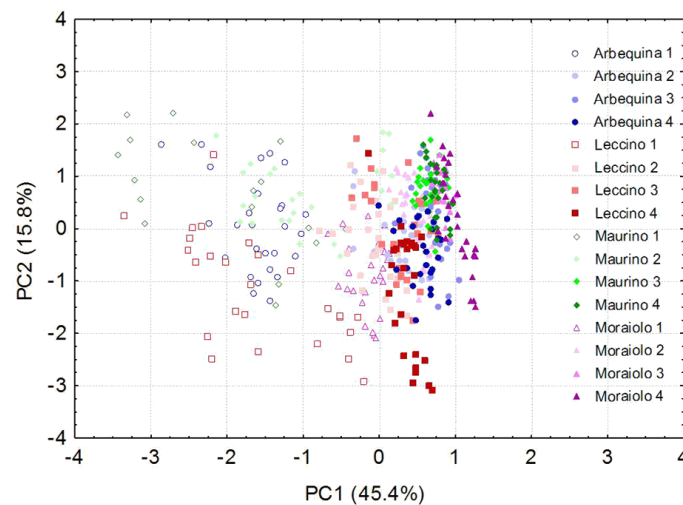
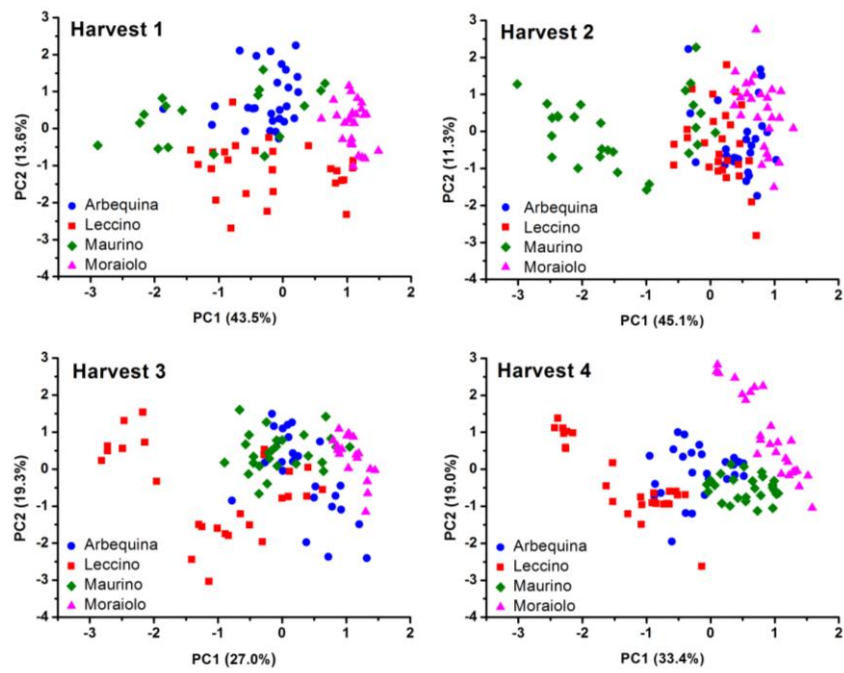


Figure 3. (Color online) Scatterplot of the scores of EVOO sample replicas (9 spectra for 48 samples) on the two-dimensional plane defined by the first two PCs as calculated by PCA of the Raman spectra. Each EVOO sample is represented by a symbol, grouped by cultivar and number of harvest, according to the legend.



17

18 **Figure 4.** (Color online) Scatterplot of the scores of the EVOO sample replicas on the two-
 19 dimensional plane defined by the first two PCs, at four harvests.

20

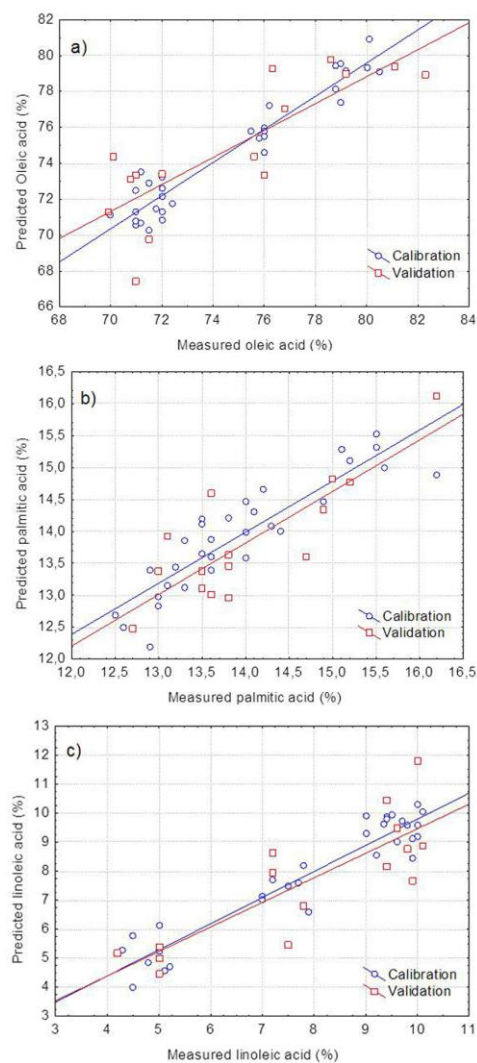
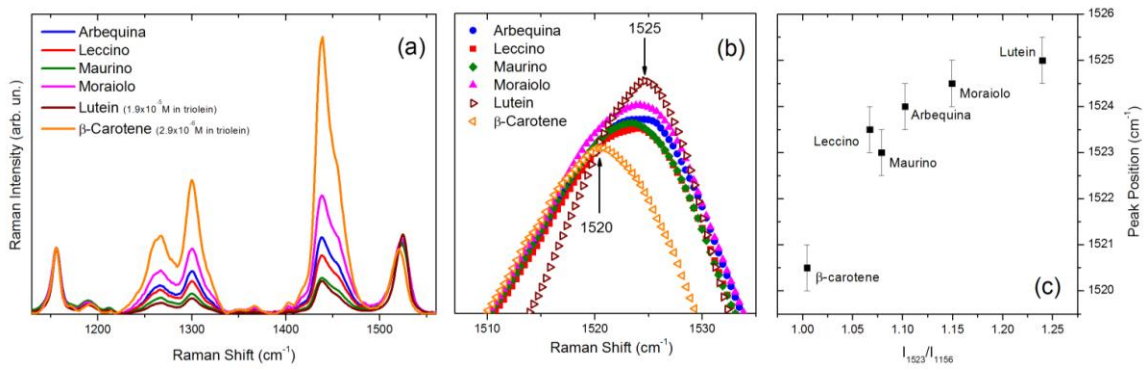


Figure 5. (Color online) PLS-R calibration models for the relationship between measured values and Raman predicted values of OA (a), PA (b) and LA (c) relative concentration in EVOOs.



25

26 **Figure 6.** (Color online) (a) Resonant Raman spectra from Arbequina, Leccino, Maurino and
 27 Moraiolo EVOOs at the second harvest, averaged over three plants per cultivar, compared with the
 28 resonant Raman spectra of β -carotene and lutein, normalized at the intensity of the C-C stretching
 29 vibration (1156 cm^{-1}). (b) Zoomed view of the carotenoid C=C stretching region. (c) Plot of the
 30 carotenoid C=C stretching peak position as a function of the intensity *ratio* between the C=C and
 31 the C-C stretching mode peaks (I_{1523}/I_{1156}).

1 **Table 1.** Classification matrixes as obtained from LDA, for the calibration (CAL) and cross-
 2 validation (VAL) sets, based on cultivar. Arb: Arbequina; Lec: Leccino; Mau: Maurino; Mor:
 3 Moraiolo.

CAL	Harvest 1					Harvest 2				
	Arb	Lec	Mau	Mor	% correct	Arb	Lec	Mau	Mor	% correct
Arb	22	0	4	1	81.5	19	3	0	5	70.4
Lec	0	27	0	0	100	1	24	0	2	88.9
Mau	8	0	9	0	52.9	0	0	27	0	100
Mor	0	0	0	27	100	5	2	0	20	74.1
Total					86.7					83.3
	Harvest 3					Harvest 4				
	Arb	Lec	Mau	Mor	% correct	Arb	Lec	Mau	Mor	% correct
Arb	26	0	1	0	96.3	22	1	3	0	84.6
Lec	3	23	0	0	88.5	0	27	0	0	100
Mau	1	0	25	1	92.6	2	0	25	0	92.6
Mor	0	0	0	18	100	0	0	0	27	100
Total					93.9					94.4

VAL	Harvest 1					Harvest 2				
	Arb	Lec	Mau	Mor	% correct	Arb	Lec	Mau	Mor	% correct
Arb	6	0	2	1	66.7	6	1	0	4	54.5
Lec	0	9	0	0	100	0	8	0	1	88.9
Mau	2	0	4	0	66.7	0	0	9	0	100
Mor	0	0	0	6	100	1	0	0	3	75
Total					83.3					78.8
	Harvest 3					Harvest 4				
	Arb	Lec	Mau	Mor	% correct	Arb	Lec	Mau	Mor	% correct
Arb	7	0	1	0	87.5	7	0	1	0	87.5
Lec	1	6	0	0	85.7	0	8	0	0	100
Mau	1	0	6	1	75.0	1	0	8	0	88.9
Mor	0	0	0	6	100	0	0	0	8	100
Total					86.2					93.9

11 **Table 2.** Averaged FA composition of the EVOO samples, as a function of cultivar and harvest
12 date. Standard deviations were calculated on the three samples belonging to the same cultivar.

Harvest	FA	FA composition %			
		Arbequina	Leccino	Maurino	Moraiolo
1	OA	70.3 ± 0.3	78.8 ± 0.1	71.3 ± 0.3	75.7 ± 0.2
	PA	14.9 ± 0.1	12.7 ± 0.1	13.7 ± 0.2	13.1 ± 0.1
	LA	9.9 ± 0.1	5.1 ± 0.1	9.4 ± 0.1	7.1 ± 0.1
2	OA	70.7 ± 0.4	79.4 ± 0.3	71.8 ± 0.2	76.1 ± 0.1
	PA	15.4 ± 0.4	12.8 ± 0.1	13.9 ± 0.1	13.6 ± 0.1
	LA	9.8 ± 0.2	4.8 ± 0.2	9.2 ± 0.1	7.5 ± 0.2
3	OA	71.1 ± 0.1	80.0 ± 0.7	71.9 ± 0.1	76.2 ± 0.2
	PA	15.6 ± 0.3	13.3 ± 0.1	14.2 ± 0.1	13.6 ± 0.1
	LA	9.9 ± 0.1	4.8 ± 0.1	9.3 ± 0.2	7.4 ± 0.2
4	OA	71.3 ± 0.4	80.7 ± 0.9	72.0 ± 0.3	76.2 ± 0.3
	PA	15.4 ± 0.1	13.3 ± 0.1	14.1 ± 0.1	13.7 ± 0.1
	LA	9.9 ± 0.1	4.4 ± 0.2	9.8 ± 0.1	7.8 ± 0.1

13

14

Supplementary material

[Click here to download Supplementary Interactive Plot Data \(CSV\): supplementary material_20180816.docx](#)