JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Study of the Cultivar−Composition Relationship in Sicilian Olive Oils by GC, NMR, and Statistical Methods

LUISA MANNINA,*,†,# GIACOMO DUGO,‡ FRANCESCO SALVO,‡ LUCIA CICERO,‡ GIULIANA ANSANELLI,§ CRISTINA CALCAGNI,# AND ANNALAURA SEGRE#

Dipartimento S.T.A.T., Facoltà di Scienze MM.FF.NN., Università del Molise, Via Mazzini 8, 86170 Isernia, Italy; Dipartimento di Chimica Organica e Biologica, Facoltà di Scienze MM.FF.NN., Università di Messina, Salita Sperone 31, 98166 Messina, Italy; C.R. ENEA-Trisaia, S.S. 106 Jonica, Km 419+500, 75025 Rotondella (MT), Italy; and Istituto di Metodologie Chimiche, CNR, 00016 Monterotondo Stazione, Roma, Italy

The aim of this research is to find if there is direct evidence relating the fatty acid composition of olive oils to specific cultivars grown within a well-limited geographical region. To group olive oils according to their own cultivars, ^{13}C high-field nuclear magnetic resonance (NMR) and gas chromatography (GC) were used to analyze 60 extra virgin olive oils from the same Italian region (southwestern Sicily) obtained from four monovarietal cultivars. The ^{13}C NMR spectrum provides information about glycerol triesters of olive oils, i.e., about the acyl composition of major components and about the fatty acids' positional distribution on the glycerol moiety. GC gives the complete fatty acid profile of olive oil samples. Selection of NMR and GC peaks on the basis of their sensitivity to the different cultivars was performed by using multivariate analysis of variance (MANOVA). Principal component analysis, tree clustering analysis, multidimensional scaling (MDS), and linear discriminant analysis (LDA) were then performed on the MANOVA-selected peaks. Results obtained from ^{13}C NMR and GC techniques combined with the multivariate statistical procedure are in good agreement and prove the usefulness of fatty acids analysis to group the monovarietal olive oils belonging to the same cultivars. Grouping of olive oils according to their cultivars occurs for particular ^{13}C resonances all belonging to fatty chains in the *sn* 1,3 position of the glycerol moiety.

KEYWORDS: Cultivar; extra virgin olive oil; gas chromatography; NMR; statistical analysis

INTRODUCTION

The determination of the geographical origin of extra virgin olive oils is a recent problem: the quality of an olive oil is the result of several factors including the cultivar, the pedoclimatic condition, and the production practice.

An important act of legislation, the "Protected Designation of Origin" (PDO), allows the labeling of some European extra virgin olive oils with the names of the areas where they are produced. This certification improves the commercial value of the product. Several attempts have been made to define the olive oil origin (*1*, *2*) by means of multivariate statistical analysis applied to fatty acids (*3*), triacylglycerols (*4*), sensory attributes (*5*), and sensory and chemical components (*6*).

In the past 10 years, we have been working on the systematic observation of components of extra virgin olive oils and on their relationship to the geographical origin and to the cultivar (*7*−*11*). It was shown that the choice of suitable components of olive oils allows the observation of two well-separable major effects, namely the pedoclimatic effect and the cultivar effect (*12*).

It was previously shown (*8*, *10*, *11*) that high-field ^{1}H NMR gives fair results in the characterization and geographical classification of extra virgin olive oils.

Moreover, in different vegetable oils, high-resolution ^{13}C NMR spectroscopy provides valuable information about the acyl composition and the *sn* (strictly numbered) 1,3 and *sn* 2 acyl positional distribution of glycerol triesters (*13*−*15*).

Until now, no direct evidence had been found relating the fatty acid composition of olive oils to the cultivar within a well-limited geographical region. Thus, we decided to study only the cultivar effect on the fatty acid composition, neglecting the pedoclimatic factor. This was accomplished by choosing four monovarietal cultivars with the precise restriction that all were grown in southwestern Sicily, i.e., in a very homogeneous microclimatic environment (*16*).

In this research, ^{13}C NMR and GC in conjunction with a multivariate statistical procedure have been used. The results obtained with both experimental techniques will be discussed, and finally all results and observations will be compared.

* Corresponding author (fax +39 06 906 72 477, E-mail mannina@mlib.cnr.it).
† Università del Molise.
‡ Università di Messina.
§ C.R. ENEA-Trisaia.
# Istituto di Metodologie Chimiche, CNR.

10.1021/jf025656l CCC: $25.00
Published on Web 11/26/2002

MATERIALS AND METHODS

Sampling. The following 60 olive oil samples, obtained within the project POM B02 Misura 2 ("Reduction of production costs, improving of the quality and protection of environment in the olive oil processing chain") have been analyzed: 16 olive oils derived from Nocellara del Belice cultivar (cv) were obtained from Castel Vetrano and Trapani areas, from picking days Oct 7 to Nov 18, 1999; 13 olive oils derived from Cerasuola cv were obtained from the Paceco area, from picking days Oct 10 to Dec 15, 1999; 22 olive oils from Biancolilla cv were obtained from Caltabellotta and Delia areas, from picking days Oct 11 to Nov 12, 1999; and 9 olive oils from Tonda Iblea cv were obtained from the Ragusa area, from picking days Nov 10 to Nov 15 1999.

^{13}C NMR Analysis. Olive oil samples (100 μL) were placed in 5-mm NMR tubes and dissolved in chloroform-*d* (600 μL) (*9*). ^{13}C spectra were recorded at 300 K on a Bruker AMX600 spectrometer operating at 150.9 MHz using the following acquisition parameters: acquired points, 256K; processed points, 128K; spectral width, 195 ppm; digital resolution, 0.22 Hz per point; relaxation delay, 8 s. The GARP sequence (*17*) was applied during the whole sequence for proton decoupling.

To perform the statistical analysis, the intensity of 77 ^{13}C signals has been measured (*13*).

Gas Chromatographic Analysis. The fatty acid composition was released as methyl ester by the official G.U. of the CEE methylation procedure (European Community Regulation 1991) and analyzed by gas chromatography (GLC). A Shimadzu GC 17A (Milano, Italy) instrument, equipped with a split/splitless injector (split ratio 70:1) and flame ionization detector, was used.

An SP 2380 fused silica capillary column, 30 m × 0.25 mm i.d., 0.20 μm film thickness (Supelco, Inc.,Bellefonte, PA) was employed.

The chromatographic conditions were: column temperature was programmed from 160 °C (kept for 10 min) to 190 °C at 4 °C/min (maintained for 13 min); injector and detector temperature, 250 °C; carrier gas, hydrogen at linear velocity of 50 cm/s.

Peak areas and percentages were calculated using Shimadzu CLASS-VP 4.3. The areas of normalized GC peaks have been submitted to the statistical procedure.

Statistical Methods. ^{13}C NMR and GC data were evaluated with the Statistica software package for Windows 1997 by Statsoft, Inc. A statistical procedure based on the following points was performed:

Multivariate Analysis of Variance (MANOVA). MANOVA allows us to select significantly discriminant variables with a high index of reliability. The parameter which indicates whether a variable discriminates between groups is the *F* value, which is computed as the ratio of the between-groups variance over the within-group variance. The larger this ratio, the larger is the discriminant power of the corresponding variable. The probability of error that is involved in accepting an observed result as valid is given by the *p*-level factor. According to conventions based on general research experience, results that yield *p*-level ≤ 0.05 (probability of error 5%) are considered borderline statistically significant, those with a *p*-level ≤ 0.01 are statistically significant, and finally, those with a *p*-level ≤ 0.005 or *p*-level ≤ 0.001 are often called "highly significant".

Principal Component Analysis (PCA). PCA allows us to obtain linear combinations of the selected variables which capture their "essence" and maximize the variability among groups.

Tree Clustering Analysis (TCA). TCA allows to classify samples without any a priori hypothesis. Different amalgamation rules can be used to determine when two clusters have to be joined together. In particular, in this work the complete-linkage (GC data) and the unweighted pair-group average (NMR data) methods have been used.

Multidimensional Scaling (MDS). MDS allows us to construct a configuration of samples which attempts to satisfy all the conditions imposed by the dissimilarity matrix obtained by TCA: a "stress value" indicates how well the ordination satisfies the dissimilarities between the samples. This coefficient indicates the degree to which the two-dimensional plot provides an acceptable summary of the multidimensional sample relationships. Stress values <0.05 indicate an excellent representation with no prospect of misinterpretation, whereas MDS plots with stress values >0.3 should be treated with caution, as the points are close to being arbitrarily placed in the two-dimensional ordination space (*18*).

Linear Discriminant Analysis (LDA). LDA allows us to classify samples with the a priori hypothesis, that is, the number of groups suggested by TCA, and to find the variables with the highest discriminant power. This analysis is used to determine whether the model (with all variables) leads to significant differences between the

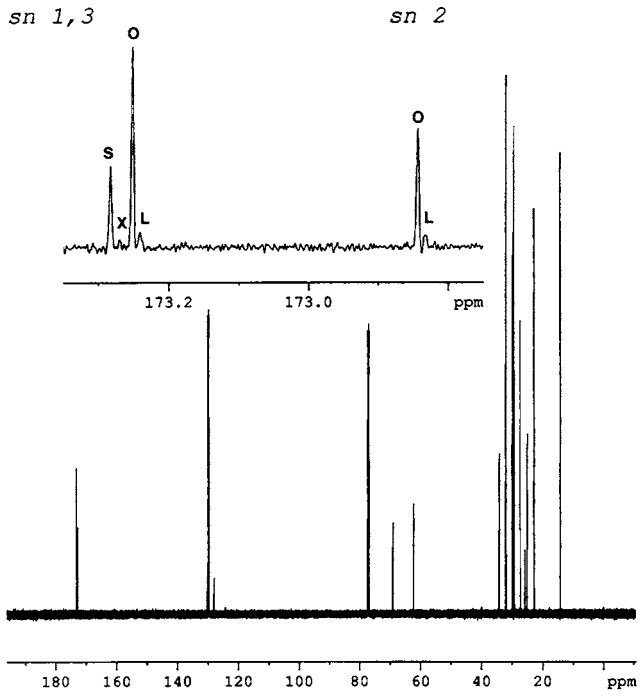

Figure 1. ^{13}C NMR spectrum of an extra virgin olive oil. In the inset, the expansion of the carbonyl region is reported: S, palmitic and stearic fatty chains; X, *cis*-vaccenic and eicosenoic fatty chains; O, oleic fatty chain; L, linoleic fatty chain.

Table 1. ^{13}C NMR Variables Selected by the Multivariate Analysis of Variance (MANOVA) Using a *p*-Level ≤ 0.00002[a]

variables	fatty acid of the resonance	*F*(3.60)	variables	fatty acid of the resonance	F(3.60)
V1	*sn* 1,3 S[a]	20.0955	V24	*sn* 1,3 O	68.4958
V2	*sn* 1,3 S	79.8307	V25	*sn* 1,3 O	24.1999
V3	*sn* 1,3 S	111.7780	V26	*sn* 1,3 O	26.6280
V4	*sn* 1,3 S	77.4107	V27	*sn* 1,3 X	37.6502
V5	*sn* 1,3 S	63.6667	V28	*sn* 1,3 X	39.2055
V6	*sn* 1,3 S	129.6028	V29	*sn* 1,3 X	67.1849
V7	*sn* 1,3 S	126.9832	V30	*sn* 1,3 X	86.1699
V8	*sn* 1,3 S	57.6859	V31	*sn* 1,3 X	36.1232
V9	*sn* 1,3 S	62.4963	V32	*sn* 1,3 X	30.0291
V10	*sn* 1,3S	20.7238	V33	*sn* 1,3 X	13.38033
V11	*sn* 1,3 O	15.8740	V34	*sn* 1,3 X	31.0989
V12	*sn* 1,3 O	9.1743	V35	*sn* 1,3 X	69.4424
V13	*sn* 1,3 O	18.2053	V36	*sn* 1,3 X	111.8663
V14	*sn* 1,3 O	49.2373	V37	*sn* 2 O	9.19101
V15	*sn* 1,3 O	47.8136	V38	*sn* 2 O	12.4186
V16	*sn* 1,3 O	49.5003	V39	*sn* 2 O	10.2915
V17	*sn* 1,3 O	27.2143	V40	*sn* 2 O	9.93652
V18	*sn* 1,3 O	29.1535	V41	*sn* 2 O	9.72731
V19	*sn* 1,3 O	34.7777	V42	*sn* 2 L	12.6875
V20	*sn* 1,3 O	61.4561	V43	*sn* 2 L	10.0832
V21	*sn* 1,3 O	30.0093	V44	*sn* 2 L	12.3162
V22	*sn* 1,3 O	21.37474	V45	*sn* 2 L	12.7497
V23	*sn* 1,3 O	38.4440	V46	CH2O	12.6749

[a] The *f*-value for ach variable is also reported. S, saturated fatty chains; X, *cis*-vaccenic and eicosenoic fatty chains; O, oleic fatty chain; L, linoleic fatty chain; CH2O glycerol moiety.

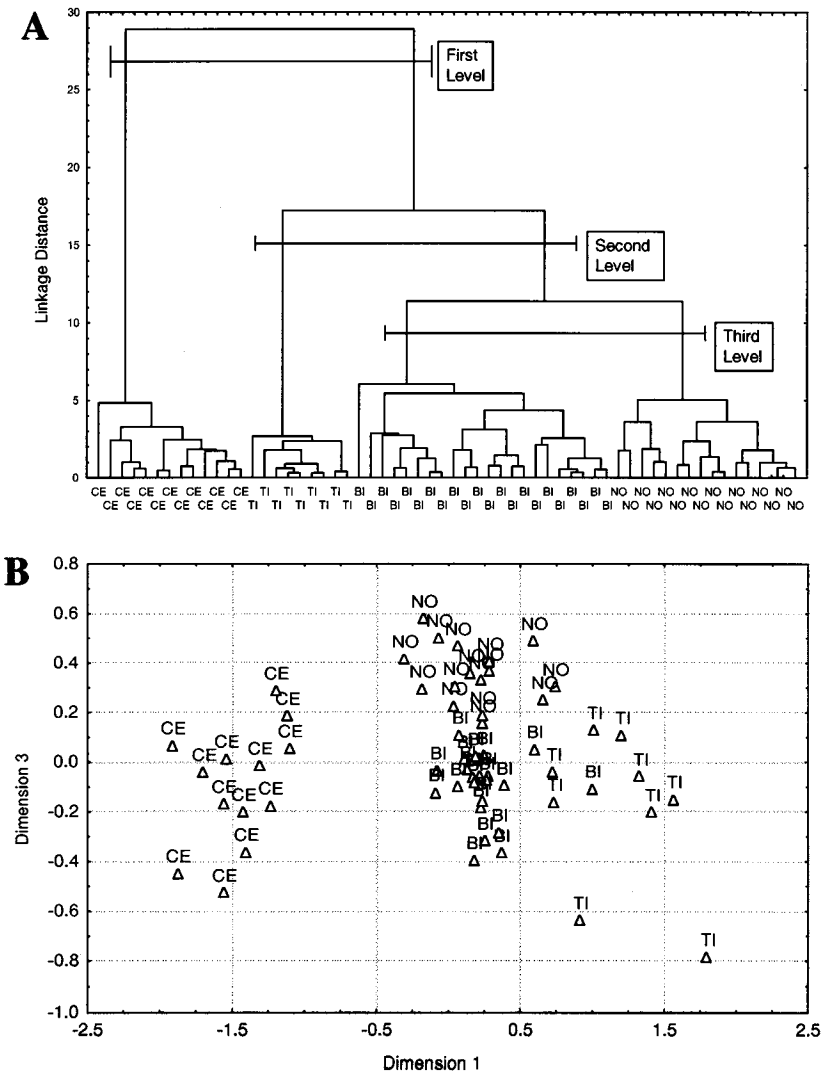

Figure 2. TCA (A) and MDS (B) analyses of 60 Sicilian extra virgin olive oil samples based on 46 ^{13}C NMR peaks. Samples labeled with the same letter come from the same cultivar: CE, Cerasuola; TI, Tonda Iblea; BI, Biancolilla; and NO, Nocellara. In panel A, three levels corresponding to a different number of groups are marked.

a priori defined groups and which variables have significantly different means across the groups. The selected variables are submitted to linear combinations to give rise to discriminant canonical functions, whose number is equal to the number of groups minus one: the first function provides the most overall discrimination between groups, the second provides the second-most, and so on. The discriminant power of the variables is evaluated using Wilks' λ, F, and p-level parameters. The Wilks' λ is computed as the ratio of the determinant of within-group variance/covariance matrix to the determinant of the total variance/covariance matrix: its value ranges from 1.0 (no discriminatory power) to 0.0 (perfect discriminatory power).

Reliability of the System (*19*, *20*). To prove the reliability of the system, some randomly selected olive oil samples are not included in the first statistical analysis and then are introduced as "unknown samples" in subsequent calculations. If these calculations classify correctly the "unknown samples", then the system is reliable and can be used for real samples.

RESULTS AND DISCUSSION

^{13}C NMR and Statistical Methods. In **Figure 1**, the ^{13}C spectrum of an olive oil is shown; the complete assignment is reported elsewhere (*13*). ^{13}C NMR technique (*14*, *15*, *21*) can provide valuable information about the *sn* 1,3 and *sn* 2 acyl positional distribution of major components of triacylglycerols, since some resonances are sensitive to the position on the glycerol moiety. For instance, in the carbonyl region of the ^{13}C spectrum (see the inset of **Figure 1**), signals due to *sn* 1,3 and *sn* 2 acyl chains show a strong chemical shift difference and a fair compositional sensitivity. In olive oils, the analysis of the positional distribution reveals (*13*) that unsaturated fatty chains are more abundant in position 2, with the noticeable exception of eicosen-11-oic acid and *cis*-vaccenic acid, which in position 2 are present only as a trace. As is well known, the saturated fatty chains are present only in position *sn* 1,3.

The intensity of all detectable ^{13}C resonances has been measured and then submitted to the following multivariate statistical methods:

MANOVA. This analysis was applied to the 77 ^{13}C resonances. The F value and the p-level allowed us to choose the more discriminant variables and, thus, to lower the number of variables. Using a p-level $\leq$ 0.00002, 46 variables have been selected; 36 variables are due to resonances on the *sn* 1,3 position, whereas only 10 variables are due to resonances on the *sn* 2 position and on the glycerol moiety. The 31 eliminated resonances belong to the linoleic acid in positions *sn* 1,3 and *sn* 2, and to the oleic acid in position *sn* 2. In detail, the 46 resonances meaningful for the statistical analysis are due to oleic

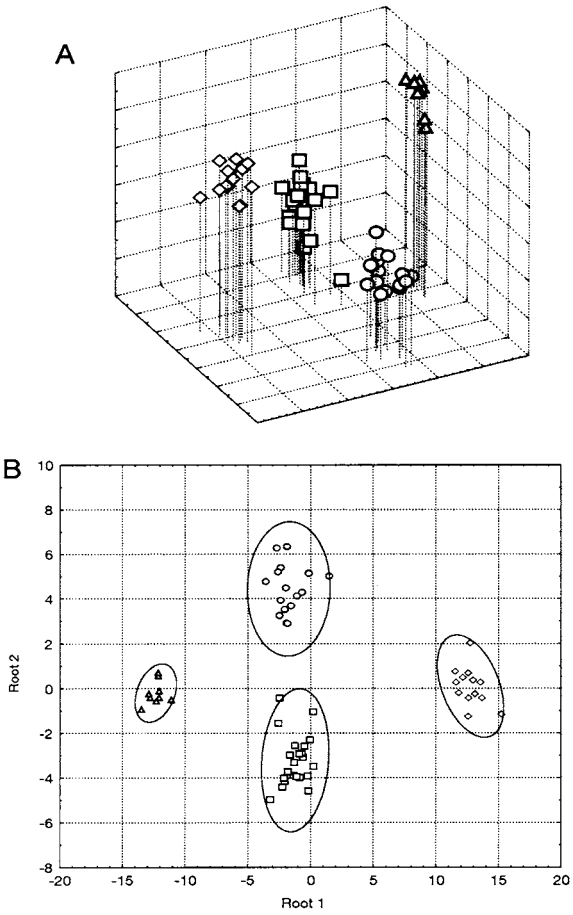

Figure 3. LDA of 60 Sicilian extra virgin olive oils based on 46 ^{13}C peaks. (A) Three-dimensional plot obtained with the canonical scores for the three discriminant equations (roots 1, 2, and 3). (B) Two-dimensional plot with the canonical scores for the two discriminant equations (roots 1 and 2). Ellipses represent the 95% confidence regions for each group. Samples labeled with the same symbol come from the same cultivar: ○, NO, Nocellara; □, BI, Biancolilla; ◇, CE, Cerasuola; △, TI, Tonda Iblea.

acid in position *sn* 1,3 (16 variables); saturated fatty chains in positions *sn* 1,3 (10 variables); *cis*-vaccenic and eicosenoic fatty chains in position *sn* 1,3 (10 variables); oleic and linoleic fatty chains in position *sn* 2 (9 variables); and glycerol moiety (1 variable) (see **Table 1**). It is worth noting that the last 10 variables have a low *F* value: this means that their discriminant power is lower than that of the other variables.

PCA. The 46 selected variables have been submitted to PCA to obtain a small number of linear combinations (called principal components) without significant loss of useful information (*22*). Three principal components have been obtained.

TCA, MDS, and LDA. These analyses have been applied to the 46 selected variables and to the three principal components derived from PCA. The distance matrix (Euclidean distances), obtained by TCA, has been used for the MDS analysis. The results obtained by all methods are fully consistent.

The result of TCA is reported as a dendrogram, which can be cut at different levels (see **Figure 2A**). The linkage distance is reported on the *y*-axis; this distance is proportional to the similarity among samples, whereas the *x*-axis has no particular meaning and is useful to indicate the grouping of the 60 Sicilian extra virgin olive oils. Cutting the dendrogram at a first level, two big groups are obtained: the first group consists of the 13 olive oils from Cerasuola cultivar, and the second one consists of all the other olive oil samples. Cutting the tree at a second level, the olive oils from Tonda Iblea cultivar are grouped together. Finally, a further level splits Biancolilla and Nocellara cultivars.

MDS analysis has been applied to get a distribution of samples respecting the constraints imposed by the dissimilarity matrix obtained by TCA. In **Figure 2B**, the MDS ordination plot shows Cerasuola and Tonda Iblea cultivars separated from each other and from the Biancolilla and Nocellara cultivars. Samples belonging to Biancolilla and Nocellara are positioned in close proximity to each other, and a better separation is observed along the third dimension. The stress value of 0.04 obtained for this analysis indicates that such ordination is a good representation of the dissimilarities among the samples.

LDA has been applied to the 46 variables using four groups, corresponding to the four cultivars, as an a priori input. In a three-dimensional plot with the three roots on the axes (see **Figure 3A**), a perfect separation among all the cultivars is observable. The proper selection of the roots allows the observation of four fully separated groups, in a root1/root2 bidimensional plot (see **Figure 3B**).

The LDA shows significant differences between groups; in fact, the Wilks' λ value is near zero (0.00018), indicating a very good discriminant power of the model. Moreover, the *F* ratio is greater than the *F* critic [$F(126.45) = 6.1175$], implying that the means of the groups differ more than would be expected by chance alone. Moreover, the very low *p*-level <0.00001 indicates a high probability of a correct classification.

To prove the reliability of the model, the method has been checked using known samples as unknown variables. Specifi-

Table 2. Mean Values and Standard Deviations of GC Areas of the 15 Selected Peaks for 60 Extra Virgin Olive Oils

fatty acids	Biancolilla	Cerasuola	Nocellara	Tonda Iblea	all cultivars
C14	0.013 ± 0.005	0.01 ± 0.00	0.01 ± 0.000	0.017 ± 0.009	0.012 ± 0.005
C16	12.690 ± 0.599	9.39 ± 1.16	12.830 ± 0.900	15.49 ± 0.665	12.43 ± 2.049
trans C*′16*	0.054 ± 0.007	0.077 ± 0.017	0.067 ± 0.015	0.103 ± 0.036	0.070 ± 0.024
C*′16*	0.691 ± 0.104	0.202 ± 0.041	0.742 ± 0.103	1.027 ± 0.146	0.649 ± 0.280
C17	0.244 ± 0.378	0.062 ± 0.047	0.038 ± 0.008	0.093 ± 0.111	0.127 ± 0.248
C*′17*	0.292 ± 0.055	0.108 ± 0.107	0.064 ± 0.010	0.107 ± 0.035	0.164 ± 0.117
C18	2.324 ± 0.169	2.508 ± 0.256	2.847 ± 0.206	2.09 ± 0.144	2.470 ± 0.326
C*′18*	71.808 ± 0.661	76.734 ± 1.670	72.244 ± 0.996	67.68 ± 1.650	72.373 ± 3.002
cis-vaccenic	1.628 ± 0.582	1.338 ± 0.863	1.919 ± 0.382	2.558 ± 0.350	1.782 ± 0.689
C*′′18*	9.031 ± 0.633	8.381 ± 1.152	7.911 ± 0.797	10.027 ± 1.017	8.741 ± 1.101
C*′′′18*	0.581 ± 0.070	0.465 ± 0.082	0.660 ± 0.047	0.282 ± 0.050	0.532 ± 0.141
C20	0.345 ± 0.073	0.427 ± 0.098	0.387 ± 0.071	0.407 ± 0.039	0.383 ± 0.080
C*′20*	0.276 ± 0.034	0.265 ± 0.087	0.249 ± 0.060	0.064 ± 0.044	0.235 ± 0.092
C22	0.01± 0.00	0.013 ± 0.005	0.012 ± 0.004	0.027 ± 0.013	0.014 ± 0.008
C24	0.01± 0.00	0.013 ± 0.005	0.0119 ± 0.004	0.022 ± 0.010	0.013 ± 0.006

Table 3. Multivariate Analysis of Variance (MANOVA) of the GC-Selected Data

GC selected fatty acids (f.a.)	*F*(3.60)	*p*-level
palmitic f.a	108.8545	0.000000
palmitoleic f.a	11.9779	0.000003
trans-palmitoleic f.a	144.7150	0.000000
heptadecanoic f.a	2.9751	0.038623
heptadecenoic f.a	43.0153	0.000000
stearic f.a	41.1743	0.000000
oleic f.a	58.3821	0.000000
cis-vaccenic f.a	10.1135	0.000017
linoleic f.a	13.4738	0.000001
linolenic f.a	18.3358	0.000000
arachic f.a	2.9226	0.041110
eicosenoic f.a	21.5222	0.000000

cally, three times a different and randomly selected set of 12 olive oils, composed of 3 olive oils from Cerasuola, 3 from Biancolilla, 3 from Tonda Iblea, and 3 from Nocellara, were removed from the data, and the model was calculated again. The excluded olive oils were introduced into the system as unknowns. Since in all runs, all samples were correctly classified, the system is stable and can be used for real samples.

Gas Chromatography and Statistical Methods. GC allows us to obtain a complete fatty acid profile of olive oil samples. The mean values and standard deviations of GC areas of the 15 selected peaks are reported in **Table 2**. Three fatty residues are practically constant in all samples and thus were omitted: these are the behenic, lignoceric, and miristic residues. The areas of the remaining 12 variables have been submitted to the following statistical analyses:

MANOVA. This analysis was applied to the 12 GC variables, each representing one fatty acid chain. Only 10 selected variables were highly discriminant, accordingly to the *p*-level ≤ 0.00002 (see **Table 3**). The heptadecanoic acid and the arachidic acid have been discarded.

PCA. The 10 selected variables have been submitted to PCA to obtain appropriate linear combinations (principal components) able to capture most of the "essence" of the variables and to maximize the variability among groups. Three principal components have been defined.

TCA, MDS, and LDA. TCA and LDA statistical methods have been applied directly both to the selected 10 variables and to the three principal components obtained by PCA, while MDS has been applied to the distance matrix (Euclidean distances)

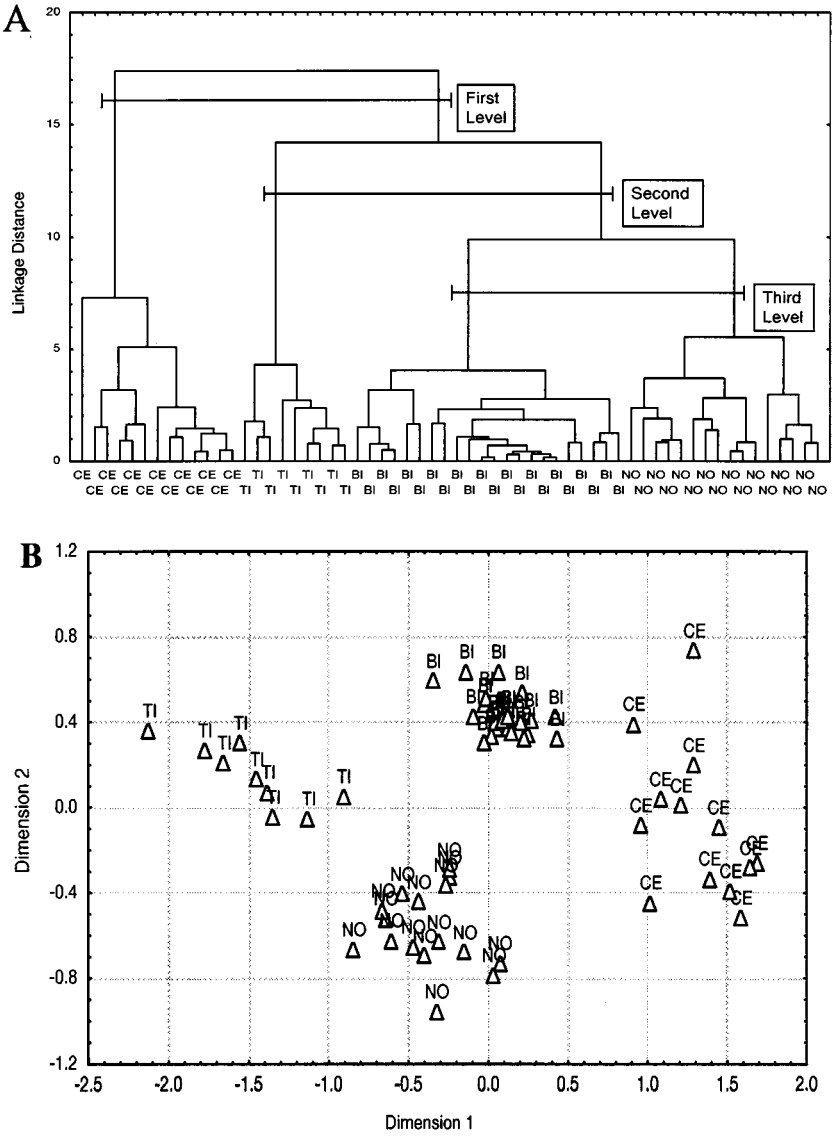

Figure 4. TCA (A) and MDS (B) of 60 Sicilian extra virgin olive oils based on 10 GC peaks. Samples labeled with the same letter come from the same cultivar: CE, Cerasuola; TI, Tonda Iblea; BI, Biancolilla; NO, Nocellara. Three levels corresponding to a different number of groups are marked.

obtained by TCA. The results obtained by all methods are fully consistent.

In **Figure 4A**, the dendrogram obtained by TCA is reported.

Cutting the dendrogram at a first level, two big groups are obtained: the first group consists of the 13 olive oils from Cerasuola cultivar, and the second one consists of all the other olive oil samples. Cutting the tree at a second level, the olive oils from Tonda Iblea cultivar are grouped together. Finally, a further level splits Biancolilla and Nocellara cultivars. Thus, the dendrogram suggests a strong separation between Tonda Iblea and Cerasuola.

In **Figure 4B**, the MDS ordination plot is reported. It shows a good separation among the four cultivars; the stress value is 0.026, indicating an excellent representation of the multidimensional sample relationships, with no prospect of misinterpretation.

LDA has been applied using four groups corresponding to the cultivars as an a priori input (see **Figure 5**). In a three-dimensional plot, with the three roots on the axes, a perfect separation among the cultivars appears (see **Figure 5A**). The observation of four groups fully separated is also clear both in a root2/root1 bidimensional plot (see **Figure 5B**) and in a root3/root1 bidimensional plot (see **Figure 5C**). Such discrimination is highly statistically significant; in fact, the Wilks' λvalue is extremely low (0.00093), according to a high discriminant power for the model, and the F ratio is high (45.197), suggesting significant differences between means across the groups. This result is highly reliable, as suggested by the low *p*-level (<0.00001).

Once again, to prove the stability of the model, the method has been checked using known samples as unknown variables. Specifically, three times a different and randomly selected set of 12 olive oils, composed of 3 olive oils from Cerasuola, 3 from Biancolilla, 3 from Nocellara, and 3 from Tonda Iblea**,** were removed from the data, and the model was calculated again. The excluded olive oils were introduced into the system as unknowns. Since in all runs, all samples were correctly classified, the system is stable and can be used for real samples.

Comparison of Analytical Methods. Results obtained from ^{13}C NMR and GC techniques, combined with a multivariate statistical procedure, are in good agreement and prove the usefulness of fatty acids to group the monovarietal olive oils. Nevertheless, as shown in **Figures 2** and **4**, a comparison of the two techniques shows that gas chromatographic analysis gives a better separation between cultivars. However, GC does not distinguish between fatty acids in position *sn* 2 or in position *sn* 1,3. Thus, we might inquire if the fatty acids selected by the GC method are evenly distributed or if a preference occurs for the *sn* 1,3 with respect to the *sn* 2 position.

The Wilks' λ factor for a forward stepwise LDA analysis, applied to the 10 GC variables, shows that palmitic, oleic, *cis*-vaccenic, linoleic, and eicosenoic acids have the highest discriminating power (see **Table 4**). Even if three out of five variables belong to the *sn* 1,3 position, i.e., palmitic, *cis*-vaccenic, and eicosenoic acids, we cannot conclude that this position is determinant for a cultivar selection.

To see if the *sn* 1,3 position is really able to group the cultivars, we chose to omit from the statistical analysis the values relative to the resonances due to the fatty acid distribution in position *sn* 2. Thus, TCA, MDS, and LDA have been applied on the 36 variables due to the resonances in position *sn* 1,3 (see **Figure 6**).

The TCA and MDS plots (see **Figure 6A**,**B**) show again Cerasuola and Tonda Iblea well separated from the other cultivars, and Biancolilla appears very near to Nocellara. The stress value of 0.045, obtained for the MDS ordination plot,

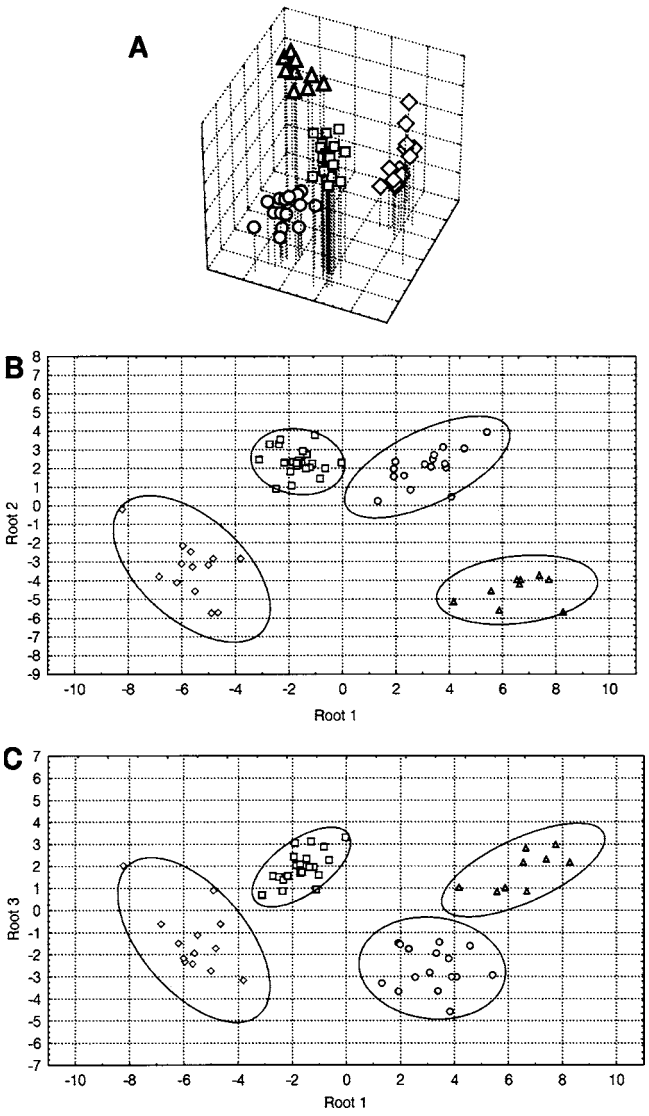

Figure 5. LDA of 60 Sicilian extra virgin olive oils based on 10 GC peaks. (A) Three-dimensional plot obtained with the canonical scores for the three discriminant equations (roots 1, 2, and 3). (B) Two-dimensional plot with the canonical scores for the two discriminant equations (roots 1 and 2). (C) Two-dimensional plot with the canonical scores for the two discriminant equations (roots 1 and 3). Ellipses represent the 95% confidence regions for each group. Samples labeled with the same symbol are from the same cultivar: ○, NO, Nocellara; □, BI, Biancolilla; ◇, CE, Cerasuola; △, TI, Tonda Iblea.

Table 4. Linear Discriminant Analysis of the 10 GC-Selected Peaks

GC-selected fatty acids	Wilks' λ	*F*-remove (3.47)	*p*-level	tolerance
palmitic	0.000982	0.85353	0.471786	0.061087
trans-palmitoleic	0.001187	4.31052	0.009130	0.662401
palmitoleic	0.001194	4.41708	0.008123	0.406463
heptadecenoic	0.002859	32.4489	0.000000	0.494066
stearic	0.001546	10.3393	0.000024	0.299536
oleic	0.000939	0.13117	0.941070	0.036870
cis-vaccenic	0.000934	0.04276	0.988083	0.168587
linoleic	0.000955	0.40335	0.751251	0.065224
linolenic	0.002440	25.3916	0.000000	0.668950
eicosenoic	0.000985	0.89958	0.448506	0.783813

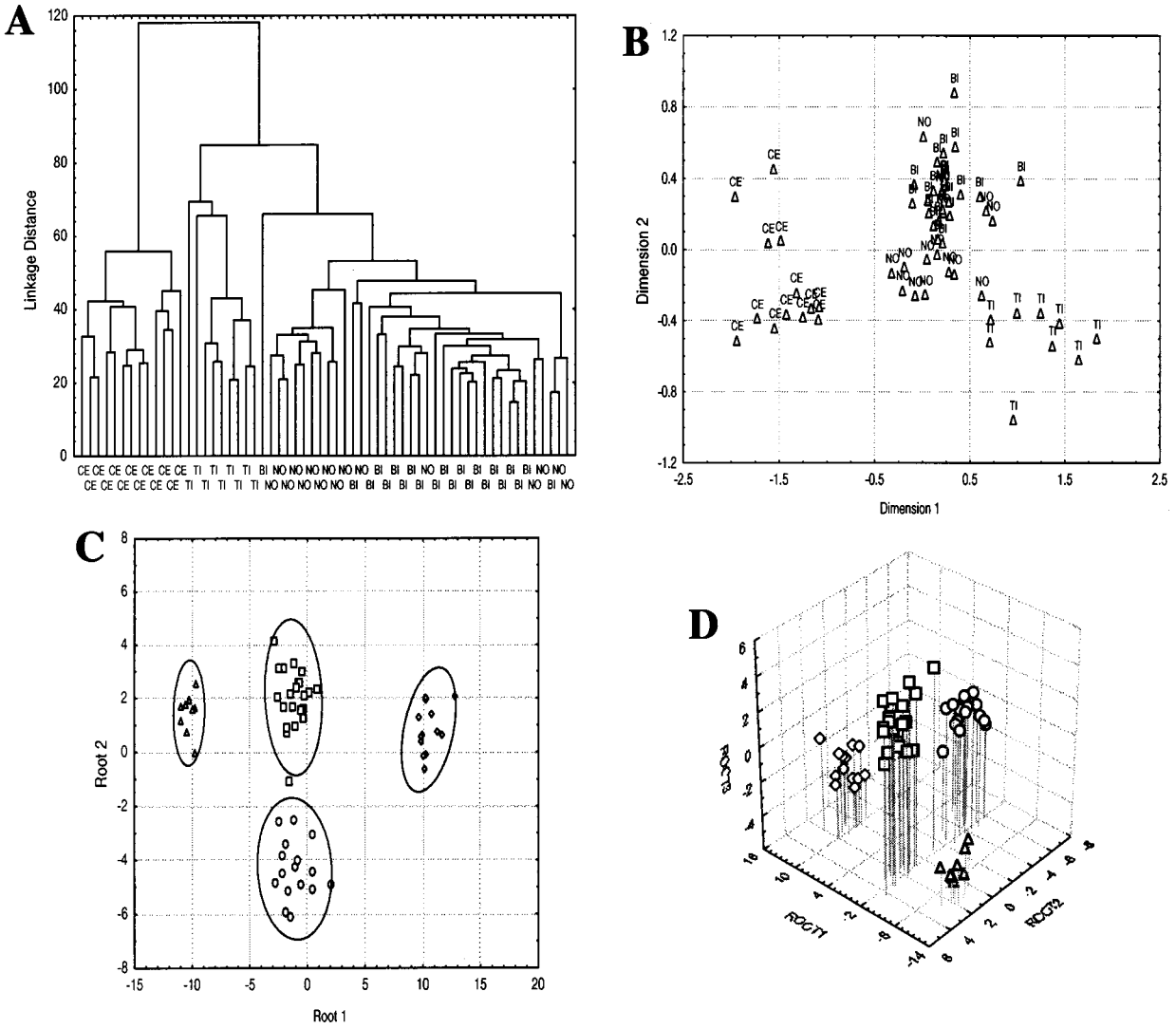

Figure 6. TCA (A), MDS (B), and LDA (C and D) of 60 Sicilian extra virgin olive oils based on 36 ^{13}C peaks. In panel C, the two-dimensional plot was obtained with the canonical scores for the two discriminant equations (roots 1 and 2). In panel D, the three-dimensional plot was obtained with the canonical scores for the three discriminant equations (roots 1, 2, and 3). Samples labeled with the same symbol are from the same cultivar: ○, NO, Nocellara; □, BI, Biancolilla; ◇, CE, Cerasuola; △, TI, Tonda Iblea.

indicates that such ordination is a good representation of the dissimilarities among the samples.

In **Figure 6C**,**D**, the results of LDA, applied on the 36 variables, are reported. In a three-dimensional plot with the three roots on the axes (see **Figure 6A**), a perfect separation among all the cultivars is shown. The observation of four groups fully separated is also clear in a root2/root1 bidimensional plot (see **Figure 6B**).

Conclusions. The results obtained in this study show that the gas chromatographic analysis gives a better separation between cultivars than NMR does.

On the other hand, in contrast to GC, NMR is able to distinguish between fatty acids in position *sn* 2 and in position *sn* 1,3. Further, it shows that the acyl composition of the *sn* 1,3 position has the major role in the cultivar selection: in fact, 36 out of 46 selected variables are due to resonances on the *sn* 1,3 position, and by using only these resonances, it is possible to obtain a good separation among cultivars, even if not so good, like the one obtained with all 46 variables.

In this study we have achieved our purpose of finding direct evidence of a relationship between the fatty acid composition and some specific cultivars grown within a well-limited geographical region; this evidence, to our knowledge, had not ever been found previously.

Further studies, based on the large chromatographic library available, are in progress.

LITERATURE CITED

(1) Alessandri, S.; Cimato, A.; Modi, G.; Mattei, A.; Crescenzi, A.; Caselli S.; Tracchi, S. Univariate models to classify Tuscan virgin olive oils by zone. *La Riv. Ital. Sostanze Grasse* **1997**, *74*, 155−163.

(2) Aparicio, R.; Albi, T.; Cert, A.; Lanzon, A. SEXIA expert system: canonical equations to characterize Spanish olive olis by varieties. *Grasas Aceites* **1988**, *39*, 219−228.

(3) Forina M.; Tiscornia, E. Pattern recognition methods in the prediction of Italian olive oil origin by their fatty acid content. *Ann. Chim.* **1982**, *72*, 143−155.

(4) Tsimidou, M.; Macrae, R.; Wilson, I. Authentication of virgin olive oils using principal components analysis of triglyceride and fatty acid profiles: part 1. Classification of Greek olive oils. *Food Chem.* **1987**, *25*, 227−239.

(5) Aparicio, R.; Calvente, J. J.; Morales, M. T. Sensory Authentication of European Extra-Virgin Olive Oil Varieties By Mathematical Procedures. *J. Sci. Food Agric.* **1996**, *72*, 435−439.

(6) Alessandri, S.; Cimato, A.; Modi, G.; Mattei, A.; Crescenzi, A.; Caselli S.; Tracchi, S. Univariate models to classify Tuscan virgin olive oils by zone. *La Riv. Ital. Sostanze Grasse* **1997**, *74*, 155−163.

(7) Sacchi, R.; Patumi, M.; Fontanazza, G.; Barone, P.; Fiordiponti, P.; Mannina, L.; Rossi, E.; Segre, A. L. A high-field ^{1}H Nuclear Magnetic Resonance study of the minor components in virgin olive oils. *J. Am. Oil Chem. Soc*. **1996**, *23*, 747−758.

(8) Sacchi R.; Mannina, L.; Fiordiponti, P.; Barone, P.; Paolillo, L.; Patumi, M.; Segre A. L. Geographical classification of Italian extra-virgin olive oils by high field ^{1}H-NMR spectroscopy. *J. Agric. Food Chem*. **1998**, *46*, 3947−3951.

(9) Segre, A. L.; Mannina, L. ^{1}H-NMR study of edible oils. *Recent Res. Dev. Oil Chem*. **1997**, *1*, 297−308.

(10) Mannina, L.; Patumi, M.; Proietti, N.; Segre, A. L. P.D.O. (Protected Designation of Origin): Geographical characterization of Tuscan extra virgin olive oils using high-field ^{1}H-NMR spectroscopy. *Ital. J. Food Sci.* **2001**, *13*, 53−63.

(11) Mannina, L.; Patumi, M.; Proietti, N.; Bassi, D.; Segre, A. L. Geographical characterization of Italian extra virgin olive oils using high field ^{1}H-NMR spectroscopy. *J. Agric. Food Chem*. **2001**, *49*, 2687−2696.

(12) Mannina, L.; Barone, P.; Patumi, M.; Fiordiponti, P.; Emanuele, M. C.; Segre A. L. Cultivar and pedoclimatic effect in the discrimination of olive oils: A high-field NMR study. *Recent Res. Dev. Oil Chem*. **1999**, *3*, 85−92.

(13) Mannina, L.; Luchinat, C.; Emanuele, M. C.; Segre, AL. Acyl positional distribution of glycerol tri-esters in vegetable oils: a ^{13}C NMR study. *Chem. Phys. Lipids* **1999**, *103*, 47−55.

(14) Mannina, L.; Luchinat, C.; Patumi, M.; Emanuele, M. C.; Rossi, E.; Segre, A. L. Concentration dependence of ^{13}C NMR spectra of triglycerides: implications for the NMR analysis of olive oils. *Magn. Reson. Chem.* **2000**, *38*, 886−890.

(15) Ng, S. Analysis of positional distribution of fatty acids in palm oil by ^{13}C NMR spectroscopy. *Lipids* **1985**, *20*, 778−782.

(16) Lanteri, S.; Armanino, C.; Perri, E.; Palopoli, A. Study of oils from calabrian olive cultivars by chemometric methods. *Food Chem.* **2002**, *76*, 501−507.

(17) Shaka, J.; Barker, S. B.; Freeman, R. J. *J. Magn. Reson*. **1985**, *64*, 574−614.

(18) Clarke and Warwick. *Change in marine communities: an approach to statistical analysis and interpretation*; Plymouth Marine Laboratory: 1994.

(19) Romesburg, H. C. *Cluster Analysis for Researchers*; Krieger Publishing: Malabar, FL, 1984.

(20) Morrison D. F. *Multivariate statistical methods*, 3rd ed.; McGraw-Hill: New York, 1990; p 269.

(21) Wollenberg, K. F. Quantitative high resolution ^{13}C nuclear magnetic resonance of the olefinic and carbonyl carbons of edible vegetable oils. *J. Am. Oil Chem. Soc*. **1990**, *67*, 487−494.

(22) Boschelle, O.; Giomo, A.; Lercker, G. Alcuni parametri di valutazione del periodo ottimale di raccolta delle olive. Esperienza triennale su alcune cultivars del Golfo di Trieste. "*Olive Oil Quality*" proceedings, 1992.

Received for review May 7, 2002. Revised manuscript received October 14, 2002. Accepted October 28, 2002. This work received partial financially support from EEC in the "Promotion of Oliviculture" program "Ricerca Sviluppo Tecnologico e Alta Formazione" and by P.O.M. B02 Misura 2 in the research project "Reduction of production costs, improving of the quality and protection of environment in the olive oil processing chain".

JF025656L