Detection of Olive Oil Adulteration Using Principal Component Analysis Applied on Total and Regio FA Content

V.G. Dourtoglou[a,*], Th. Dourtoglou[b], A. Antonopoulos[a], E. Stefanou[a], S. Lalas[a], and C. Poulos[c]

[a]Department of Food Technology and Nutrition, Technological Educational Institute (T.E.I.) of Athens, 12210, Athens, Greece, [b]Vioryl S.A., Kato Kifisia, Athens, Greece, and [c]Department of Chemistry, University of Patra, Patra, Greece

ABSTRACT: Principal component analysis (PCA) has been used to establish a new method for the detection of olive oil adulteration. The data set, composed of values obtained from the determination of the mole percentage of total FA and their regiospecific distribution in positions 1 and 3 in TG of oils (pure or mixtures) by GC analysis, was subjected to PCA. 3-D scatter plots showed clearly that it is possible to distinguish the pure oils from the mixtures. Moreover, it is possible to discriminate the different types of seed oil used for the adulteration.

Paper no. J10131 in *JAOCS 80,* 203–208 (March 2003).

KEY WORDS: Adulteration, olive oil, principal component analysis, seed oils.

Food fraud usually involves misleading the purchaser as to the true nature, composition, or quality of the goods changing hands. The offense can take the form of adulteration, which generally involves the dilution of a commodity with less expensive materials. Adulteration of olive oil with various seed oils is a common problem affecting the quality and commercial value of the product. This practice is a commercial exploitation of olive oil and causes major loss in economical value.

European Mediterranean countries producing olive oil have recently adopted a common legislation in order to protect olive oil growers and consumers from food fraud. Olive oil is the most expensive edible oil, and owing to the recent application of legal regulation, it is the most important one as far as adulteration is concerned (1). Olive oil is distinguished from other vegetable oils because it may be consumed without having been refined. When refining takes place, the product is considered less desirable since its attractive organoleptic properties are diminished during the process. Olive oil is subjected to two types of adulteration. The first is the blending of virgin olive oils with olive oils of lower grade (for example, refined olive oil or olive pomace oil). The second is the mixing of olive oil with other liquid vegetable oils.

Numerous researchers have proposed various methods to determine adulteration resulting from the mixing of the olive oil with other vegetable oils. The purity criteria for olive oils are based mainly on the FFA content, the PV, the specific extinction, the FA composition (especially in the 2-position of TG), and the sterol composition. A method for the detection of adulterated and misbranded olive oil products was proposed by Firestone *et al.* (2), based mainly on the analysis of different olive oil constituents, such as sterols and triterpenes and saturated FA in the 2-position of TG, and on other analyses such as acidity, color, and specific extinction. Other methods have also been proposed. Kapoulas and Andricopoulos (3) introduced a method for detecting adulteration of olive oil with very low levels of linoleic acid-rich oils, applying reversed-phase HPLC. Goodacre *et al*. (4) proposed a rapid assessment for olive oil adulteration. In this method, Curie-point pyrolysis mass spectra were obtained from a variety of extra-virgin olive oils prepared from various cultivars using several mechanical treatments. They proved that the major source of variation was the difference between the cultivars instead of whether the oils had been adulterated.

In our work, we propose a new method based on principal component analysis (PCA). The mole percentage of each FA plus the percentage of the same acid in the *sn*-1,3 positions of the TG are introduced as elements in the matrix. With this new method it is possible to distinguish mixtures of olive oil with seed oils even at a level of 5%. It is also possible to predict, by the applied model, the seed origin of the oil used for the adulteration. Stepwise discrimination analysis previously has been used to delineate the origin of wines (5) and to characterize of roasted coffee and coffee beverages (6). The method proposed in this work also may be applied to mixtures of other vegetable oils, after changes in the data introduced to the model.

MATERIALS AND METHODS

Vegetable oils (corn, soybean, sunflower, cottonseed, almond, and olive oils) were purchased from the local market.

1,3-Specific lipase from *Mucor miehei* (Lipozyme IM 20, Batch: LM7 0753) was obtained from Novo Nordisk A/S (Bagsvaerd, Denmark).

TLC was conducted on 0.25 mm Silica gel 60 plates F254 (5 × 20 cm) purchased from Merck Ltd. (Darmstadt, Germany). TLC plates were developed with hexane/ether/acetic acid (90:10:1) (all purchased from Sigma, St. Louis, MO), dried, and visualized by exposure to I_2 vapor.

Methyl ester formation. FAME were prepared by reaction of the oil (200 μL) with a small amount of sodium methoxide

*To whom correspondence should be addressed at Department of Food Technology and Nutrition, Technological Educational Institute (T.E.I.) of Athens, Ag. Spyridonos str., Egaleo, 12210, Athens, Greece.
E-mail: DOURT@ATH.FORTHNET.GR

TABLE 1
Temperature Program for GC Determination of FAME and FA Butyl Esters

Temp. (°C)	Rate (°C/min)	Hold (min)	Total (min)
150	0.0	0.00	0.00
212	2.0	1.00	32.00
240	10.0	5.00	39.80

in methanol in room temperature. The reaction lasted about 5 min with gentle agitation of the mixture periodically. After the reaction was completed, 10 mL of diethyl ether was added and the organic layer was washed three times with water, dried over Na_2SO_4, and reduced to 1 mL by heating in a water bath (40°C). The solution obtained was then subjected to GC analysis.

Determination of sn*-1,3 distribution of FA in TAG by butyl ester formation using a regiospecific lipase.* A mixture containing 200 µL of sample (oil) and 600 µL of *n*-butyl alcohol was placed in a test tube. Lipozyme IM (250 mg) was added, and the whole mixture was agitated. The oil was allowed to react with the above alcohol for 10 min under gentle agitation using a Vortex Genie 2 (Scientific Industries, Inc., Bohemia, NY). [The reaction was followed by TLC until the oil spot with R_f 0.5 was minimized. Three other spots with R_f values of 0.8, 0.2, and 0.1 (FA butyl ester, mono and diglycerides, respectively) were present on the plate at the end of the reaction.] Then 20 mL of pentane was added, followed by 10 mL of a saturated solution of NaCl. After gentle agitation and separation of the two phases, the water phase was removed. The organic phase was extracted twice with a saturated solution of NaCl and water and dried over Na_2SO_4. Most of the solvent was evaporated on a water bath at 40°C. During washings with water the two spots with R_f of 0.2 and 0.1 practically disappeared. Finally, the products were subjected to GC analysis.

Instrumentation. Analysis of methyl and butyl esters was carried out with a Hewlett-Packard HP 6890 gas chromatograph, equipped with a J&W (J&W Scientific, Köln, Germany) capillary DB-23 column (0.322 mm i.d., 30 m, 0.25 µm). Nitrogen flow rate was set at 1.72 mL/min. The column temperature program is detailed in Table 1. The injector and detector temperatures were maintained at 220 and 280°C, respectively.

Statistical analysis. The data set, which was composed of values obtained from GC analysis, was used for PCA and stepwise discriminant analysis (DA). The statistical program SPSS v10.0.7 for Windows (SPSS Inc., Chicago, IL) was used to calculate and plot the data from the PCA and DA.

Additionally, a classification matrix was calculated to evaluate the predictive accuracy of the discrimination model. The results were cross-validated, meaning that each case was classified by the functions derived from all cases other than that, because the original results may provide overly optimistic estimates and cross-validation attempts to remedy this problem (7–9). The classification results were compared with the ones that could be classified correctly by chance, taking into account their group sizes.

RESULTS AND DISCUSSION

As explained previously, some criteria for olive oil purity have been established. FFA content, PV, specific extinction, and regiospecific FA composition are the minimum required for the determination of the adulteration or misbranding of olive oil. Clearly, much analytical work is needed to achieve unambiguous results. In this work, a single and easily applied technique like GC, coupled with a statistical tool (DA), has been introduced for the determination of any adulteration of olive oil. The proposed method can be divided in three distinguishable parts. The first and second parts are the determination of the total FA as FAME and as *sn*-1,3 fatty acid butyl esters (FABE), respectively. The third part consists of a data matrix constructed using FAME and FABE mole percentage data measured on a total of 20 oils (10 pure oils and 10 mixed oils). The pure oils used were corn, soybean, sunflower, cottonseed, and olive oil, and the mixed oils were made with olive and soybean oil (95:5, 50:50, vol/vol), olive and cottonseed oil (95:5, 50:50, 25:75, vol/vol), olive and corn oil (95:5, 50:50, 25:75, vol/vol), and olive and sunflower oil (95:5, 50:50, 25:75, vol/vol) (see Tables 2 and 3). The method for

TABLE 2
Comparative FA Compositions (mol%) and Positional Distribution of FA in TAG (pure oils, vol/vol)

Mol%	$C_{16:0}$	$C_{18:0}$	$C_{16:1}$	$C_{18:3}$	$C_{18:2}$	$C_{18:1}$	Total
Olive oil (A)							
FA in 1,2,3	12.6	2.9	0.9	0.7	8.4	74.6	100.0
FA in 1,3	18.2	3.4	1.0	0.5	7.4	69.3	99.8
Olive oil (B)							
FA in 1,2,3	12.1	2.8	0.8	0.8	7.1	76.4	100.0
FA in 1,3	17.8	3.9	0.9	0.5	5.8	71.0	99.8
Corn oil (A)							
FA in 1,2,3	11.6	2.0	0.1	0.8	59.6	26.0	100.0
FA in 1,3	18.3	2.8	0.1	0.9	49.6	28.1	99.8
Corn oil (B)							
FA in 1,2,3	12.1	2.0	0.1	0.9	56.6	28.3	100.0
FA in 1,3	19.0	2.8	0.1	0.8	48.2	28.8	99.8
Soybean oil (A)							
FA in 1,2,3	11.9	4.3	0.1	6.9	54.4	22.5	100.0
FA in 1,3	17.7	6.8	0.1	7.0	46.1	21.2	98.8
Soybean oil (B)							
FA in 1,2,3	12.7	4.4	0.1	6.8	53.0	23.0	100.0
FA in 1,3	16.4	6.6	0.1	7.1	47.7	22.1	100.0
Sunflower oil (A)							
FA in 1,2,3	7.1	4.4	0.1	0.1	65.7	22.7	100.0
FA in 1,3	10.9	6.0	0.1	0.1	60.9	21.8	99.8
Sunflower oil (B)							
FA in 1,2,3	7.4	4.0	0.1	0.1	55.8	32.7	100.0
FA in 1,3	10.0	5.1	0.1	0.1	55.8	28.4	99.6
Cottonseed oil (A)							
FA in 1,2,3	25.3	2.4	0.6	0.2	54.2	17.3	100.0
FA in 1,3	38.8	3.6	0.7	0.1	41.6	15.0	99.8
Cottonseed oil (B)							
FA in 1,2,3	23.4	2.3	0.6	0.2	56.1	17.4	100.0
FA in 1,3	36.2	3.0	0.7	0.1	45.0	15.0	100.0

TABLE 3
Comparative FA Compositions (mol%) and Positional Distribution of FA in TAG (mixed oils, vol/vol)

Mol%	$C_{16:0}$	$C_{18:0}$	$C_{16:1}$	$C_{18:3}$	$C_{18:2}$	$C_{18:1}$	Total
Olive and soybean oil (95:5)							
FA in 1,2,3	12.7	2.4	0.8	1.0	11.2	71.9	100.0
FA in 1,3	17.0	3.5	1.0	0.5	11	67.0	100.0
Olive and cottonseed oil (95:5)							
FA in 1,2,3	13.5	2.4	0.9	0.7	11.5	71.2	100.0
FA in 1,3	20.1	3.5	1.0	0.4	9.5	65.3	99.8
Olive and cottonseed oil (50:50)							
FA in 1,2,3	19.0	2.4	0.8	0.4	31.7	45.7	100.0
FA in 1,3	29.0	3.5	0.8	0.3	25.5	40.7	99.8
Olive and cottonseed oil (25:75)							
FA in 1,2,3	22.3	3.5	0.7	0.3	43.7	29.4	100.0
FA in 1,3	31.9	4.7	0.7	0.1	30.4	32.1	99.9
Olive and corn oil A (95:5)							
FA in 1,2,3	12.8	2.5	0.9	0.7	11.6	71.7	100.0
FA in 1,3	18.5	3.6	0.9	0.4	10.4	65.9	99.8
Olive and corn oil B (95:5)							
FA in 1,2,3	13.9	2.7	0.9	0.7	12.9	68.9	100.0
FA in 1,3	18.7	4.0	0.9	0.5	9.6	66.5	100.2
Olive and corn oil (50:50)							
FA in 1,2,3	12.4	2.6	0.5	0.8	33.9	49.8	100.0
FA in 1,3	19.1	3.1	0.6	0.6	27.5	48.9	99.8
Olive and sunflower oil (95:5)							
FA in 1,2,3	12.7	2.6	0.9	0.7	12.6	70.6	100.0
FA in 1,3	18.9	3.5	1.0	0.5	9.6	66.4	99.8
Olive and sunflower oil (50:50)							
FA in 1,2,3	10.3	3.2	0.5	0.4	37.3	48.3	100.0
FA in 1,3	15.3	4.6	0.6	0.3	32.2	46.8	99.8
Olive and sunflower oil (25:75)							
FA in 1,2,3	9.0	4.6	0.3	0.3	50.1	35.8	100.0
FA in 1,3	12.1	6.1	0.4	0.2	45.0	36	100.0

FA analysis of these oils was presented by Dourtoglou *et al.* (10). This method is based on methyl ester formation for the determination of the total mole percentage of the FA in oil and on butyl ester formation mediated by 1,3-specific lipase for the regiodistribution of the same FA (10). For the regiodistribution of FA in pure or mixed oils, any other method could be used.

PCA has been applied to differentiate oils of different mixtures from pure oils. By choosing eigenvalues greater than one (>1), the dimensionality was reduced from 12 variables (methyl and butyl esters) (see Table 4) to four principal components (PC) with eigenvalues of 6.249, 2.863, 1.541, and 1.091. In choosing the first three PC, 88.782% of the total variability was explained (PC1 = 52.079%, PC2 = 23.857%, PC3 = 12.849%).

Figures 1 and 2 and Table 4 illustrate the strong correlation that exists between methyl and butyl esters for the same FA. Moreover, the same figures reveal the relative extent to which each original variable contributes to the variance contained in each PC (factor loadings). Methyl and butyl esters of $C_{18:1}$ and $C_{16:1}$ are positively correlated with the PC1, whereas the esters of $C_{18:2}$ are strongly negatively correlated to it. Methyl and butyl esters of $C_{18:3}$ are negatively correlated to PC2.

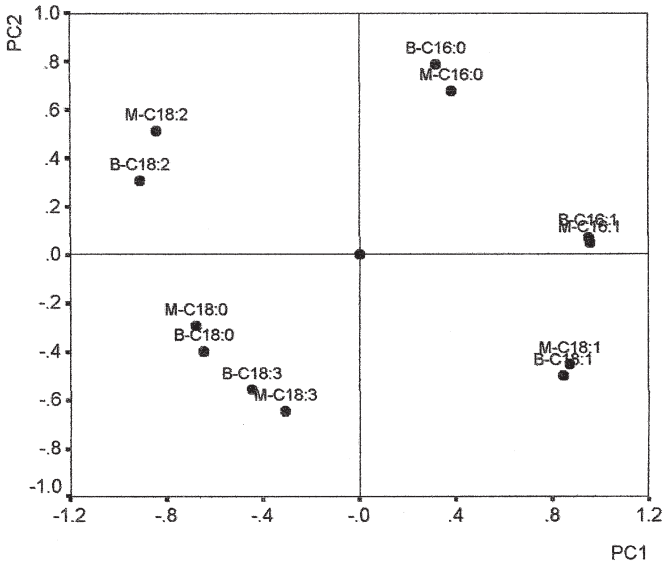

FIG. 1. Factor loadings for principal components (PC) PC1 and PC2. The prefix M or B indicates methyl or butyl esters, respectively.

The scores for the first three PC are plotted as a scatter diagram in Figure 3. In this graph one sees that PC1 is mainly responsible for the discrimination of all mixtures from pure oils. More specifically, the greater the quantity of a nonolive oil contained in a mixed sample, the greater the distance of this sample from pure olive oils. At the same time, the dispersion of samples along PC2 and PC3 for a specific mixture category is caused by the different types of oils mixed with the pure olive oil.

In applying a stepwise DA (stepwise method: Mahalanobis distance) on the oil sample groups using all 12 independent variables, five discriminant functions (DF) were deduced. The first DF accounts for 77.1% of the total variability, whereas DF2 and DF3 account for 21.9 and 0.8%, respectively (total variance explained, 99.8%). The variables that entered the discriminant model are shown in Table 5. All of them were significant at the 0.05 level. The scores for the first three discriminant functions are plotted as a scatter diagram in Figure 4. Figure 5 is an enlargement of a section of Figure 4.

According to the classification results, 100.0% of all original grouped cases were classified correctly before and after cross-validation. These percentages were far greater than those

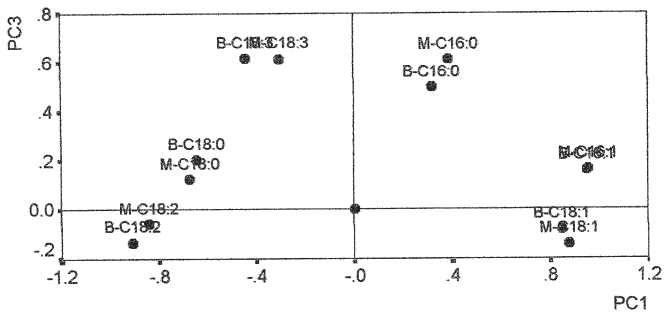

FIG. 2. Factor loadings for PC1 and PC3. For abbreviations see Figure

TABLE 4
Correlations of Transformed Variables from Principal Component Analysis

	$C_{16:0}$ methyl	$C_{16:0}$ butyl	$C_{18:0}$ methyl	$C_{18:0}$ butyl	$C_{16:1}$ methyl	$C_{16:1}$ butyl	$C_{18:3}$ methyl	$C_{18:3}$ butyl	$C_{18:2}$ methyl	$C_{18:2}$ butyl	$C_{18:1}$ methyl	$C_{18:1}$ butyl
M-C16:0[a]	1.000											
B-C16:0	0.948	1.000										
M-C18:0	–0.330	–0.356	1.000									
B-C18:0	–0.354	–0.380	0.933	1.000								
M-C16:1	0.515	0.424	–0.506	–0.479	1.000							
B-C16:1	0.528	0.434	–0.500	–0.493	0.991	1.000						
M-C18:3	–0.199	–0.319	0.279	0.397	–0.286	–0.296	1.000					
B-C18:3	–0.192	–0.284	0.346	0.453	–0.416	–0.434	0.953	1.000				
M-C18:2	–0.020	0.101	0.355	0.264	–0.814	–0.795	–0.078	0.085	1.000			
B-C18:2	–0.227	–0.134	0.413	0.340	–0.900	–0.886	0.052	0.199	0.938	1.000		
M-C18:1	–0.058	–0.139	–0.499	–0.423	0.778	0.759	–0.058	–0.216	–0.955	–0.914	1.000	
B-C18:1	–0.059	–0.159	–0.467	–0.408	0.746	0.727	0.027	–0.129	–0.954	–0.893	0.984	1.000
Dimension	1	2	3	4	5	6	7	8	9	10	11	12
Eigenvalue	6.249	2.863	1.541	1.091	0.082	0.060	0.040	0.030	0.024	0.009	0.006	0.004

[a]The prefix M or B indicates methyl or butyl esters, respectively.

that could be classified correctly by chance, and, according to that, the discriminant model is acceptable. Table 6 shows the predicted membership and prior probabilities for all groups.

As shown in Figure 4, olive oils are grouped together and differ from mixtures with other oils or other pure oils in an indistinguishable way. This spatial distribution reflects the classification results mentioned above. In the case of mixtures, the stereochemical distribution depends on the nature of the oils. A PCA has been applied using the same matrix of variables and individuals as the previous PCA. The only difference is that the individuals have been grouped according to the type of oil that was used in the mixture, whereas in the previous PCA the individuals were grouped according to the percentage of nonolive oil.

Thus, it is expected that all mixtures of oils should appear among the pure oils that were used to create the mixture. Fig-

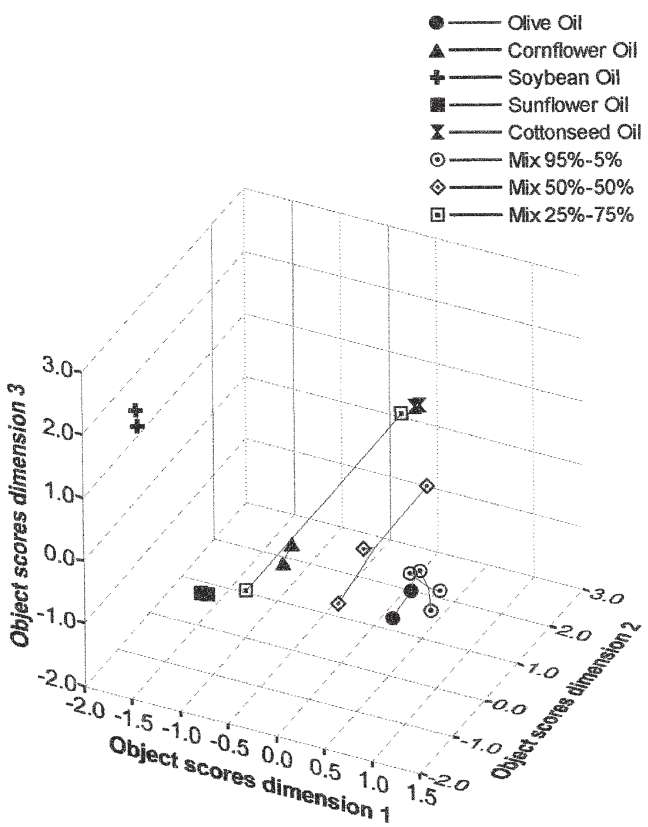

FIG. 3. Scatter plot of all oil samples for principal component analysis on PC1, PC2, and PC3.

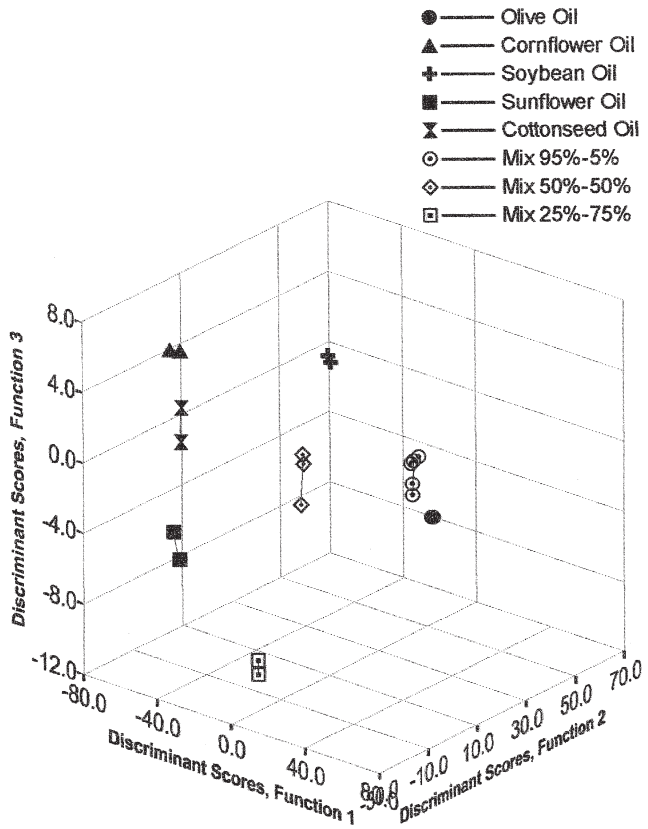

FIG. 4. Scatter plot of all oil samples from discriminant analysis on discriminant functions (DF) DF1, DF2, and DF3.

TABLE 5
Variables Entered/Removed[a–d] in the Stepwise Discriminant Analysis

			Min. D squared					
					Exact *F*			
Step[e]	Entered	Removed	Statistic	Between groups	Statistic	DF1	DF2	Sig.
1	B-C18:1		1.076	3 and 4	1.076	1	12	0.320
2	M-C18:2		5.376	2 and 3	2.464	2	11	0.131
3	M-C18:0		35.027	1 and 6	13.900	3	10	0.0006
4	M-C18:1		38.394	1 and 6	10.284	4	9	0.0020
5		B-C18:1	38.391	1 and 6	15.235	3	10	0.0004
6	M-C16:1		50.911	1 and 6	13.637	4	9	0.0007
7	B-C18:2		71.881	4 and 5	9.584	5	8	0.0031
8	M-C18:3		75.336	1 and 6	10.463	6	7	0.0033
9		M-C18:2	71.698	1 and 6	13.657	5	8	0.0009

[a]Maximum number of steps is 24.
[b]Maximum significance of *F* to enter is 0.05.
[c]Minimum significance of *F* to remove is 0.10.
[d]*F* level, tolerance.
[e]At each step, the variable that maximizes the Mahalanobis distance between the two closest groups is entered.

ure 6 presents this situation, where all mixtures start close to the olive oil position and extend to the nonolive position according to the type of nonolive oil used.

With this proposed method it is possible to distinguish pure oils from mixtures. It also should be possible to predict the type and the percentage of an oil used to adulterate pure olive oil. Additionally, many samples can be analyzed in a short time. Laboratories dedicated to the detection of adulteration can establish a data library and introduce new cases to it as they appear. Then this library can instantly give information on possible adulteration by simple incorporation of the GC results of the suspected oil into the data matrix.

REFERENCES

1. Codex Alimentarius Commission, *Report of the Fourteenth Session of the Codex Committee on Fats and Oils,* London, 27 September–1 October 1993, Alinorm 95/17, Rome, FAO/WHO.
2. Firestone, O., J.L Summers, R.J. Reina, and W.S. Adams, Detection of Adulterated and Misbranded Olive Oil Products, *J. Am. Oil Chem. Soc. 62*:1558–1562 (1985).

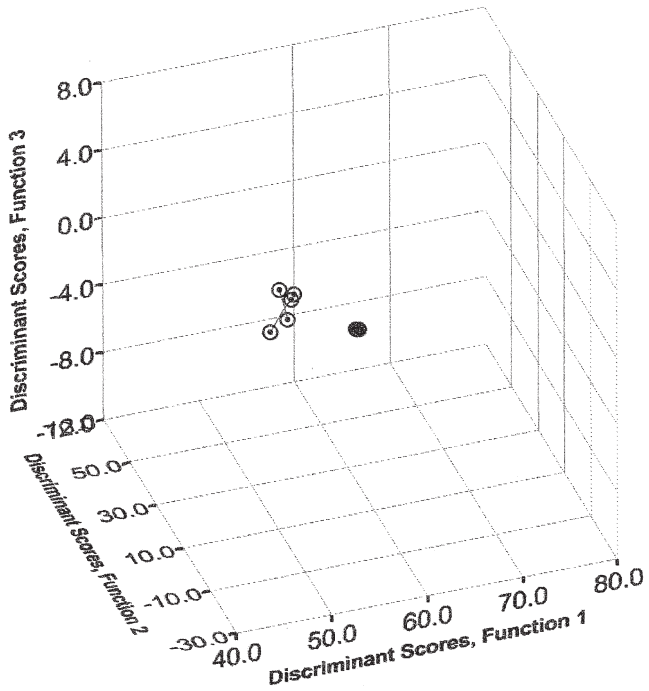

FIG. 5. Expansion of Figure 4 on groups 1 (pure olive oil, ●) and 6 (95% olive oil/5% nonolive oil, ⊙).

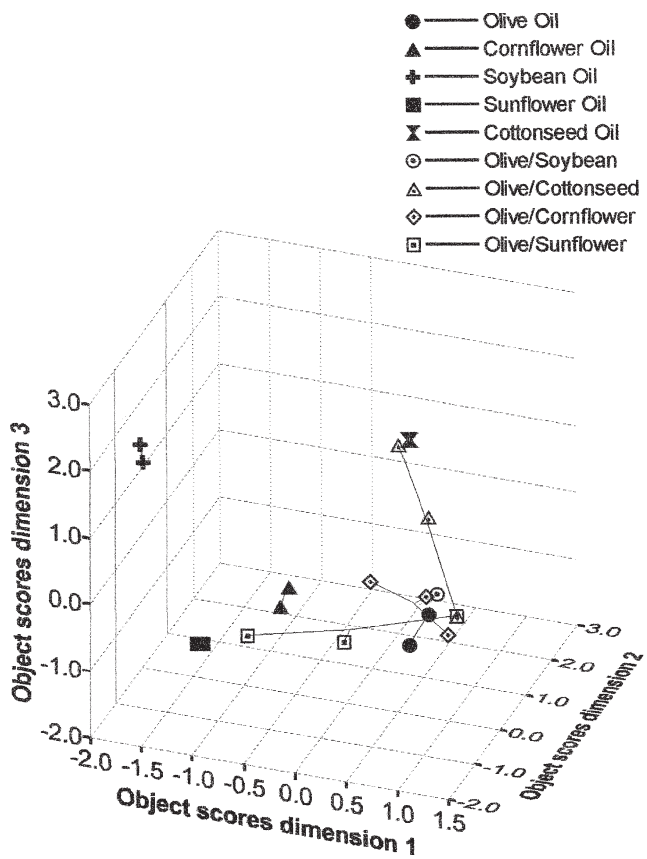

FIG. 6. Scatter plot of all oil mixtures grouped according to the type of oil used in the mixture.

TABLE 6
Classification Results[a,b]

		Predicted group membership							
		1	2	3	4	5	6	8	9
Original	Correct	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	Misclassified	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cross validated	Correct	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	Misclassified	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
		Prior probabilities							
		1	2	3	4	5	6	8	9
		10.0%	10.0%	10.0%	10.0%	10.0%	25.0%	15.0%	10.0%

[a]100.0% of original grouped cases correctly classified.
[b]100.0% of cross-validated grouped cases correctly classified.

3. Kapoulas, V.M., and N.K. Andricopoulos, Detection of Lipid Adulteration with Linoleic Acid-Rich Oils Reversed-Phase High-Performance Liquid Chromatography, *J. Chromatogr. 24*: 311–320 (1986).
4. Goodacre, R., D.B Kell, and G. Bianchi, Rapid Assessment of the Adulteration of Virgin Olive Oils by Other Seed Oils Using Pyrolysis Mass Spectrometry and Artificial Neural Networks, *J. Sci. Food Agric. 63*:297–307 (1993).
5. Huerta-Diaz Regano, M.D., M.R. Salinas Fernandez, and T. Masoud, Study of Major Volatiles in Wines and Discriminant Analysis Applied to Classification According to Region, *Boll. Chim. Farm. 136*:674–679 (1997).
6. Bicchi, C.P., O.M. Panero, G.M. Pellegrino, and A.C. Vanni, Characterization of Roasted Coffee and Coffee Beverages by Solid Phase Microextraction–Gas Chromatography and Principal Component Analysis, *J. Agric. Food Chem. 45*:4680–4686 (1997).
7. SPSS Inc., *Syntax Reference Guide for SPSS Base SPSS Regression Models SPPS Advanced Models,* ver. 10.0, Chicago, 1999.
8. Hair, J.F., Jr., R.E. Anderson, R.L. Tatham, and W.C. Black, Factor Analysis, in *Multivariate Data Analysis with Readings*, 4th edn., Prentice Hall, Upper Saddle River, NJ, 1995, pp. 365–418.
9. Hair, J.F., Jr., R.E. Anderson, R.L. Tatham, and W.C. Black, Multiple Discriminant Analysis, in *Ibid*, pp. 178–237.
10. Dourtoglou, T., E. Stefanou, S. Lalas, V. Dourtoglou, and C. Poulos, Quick Regiospecific Analysis of Fatty Acids in Triacylglycerols with GC Using Specific Lipase in Butanol, *Analyst 126*:1032–1036 (2001).

[Received October 23, 2001; accepted November 5, 2002]