JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Detection and Identification of Extra Virgin Olive Oil Adulteration by GC-MS Combined with Chemometrics

Yang Yang,[†] Miguel Duarte Ferro,[§] Isabel Cavaco,[§] and Yizeng Liang*[,†]

[†]Research Center of Modernization of Traditional Chinese Medicine, College of Chemistry and Chemical Engineering, Central South University, Changsha 410083, China

[§]DQF-FCT, Universidade do Algarve, 8005-139 Faro, Portugal

ABSTRACT: In this study, an analytical method for the detection and identification of extra virgin olive oil adulteration with four types of oils (corn, peanut, rapeseed, and sunflower oils) was proposed. The variables under evaluation included 22 fatty acids and 6 other significant parameters (the ratio of linoleic/linolenic acid, oleic/linoleic acid, total saturated fatty acids (SFAs), polyunsaturated fatty acids (PUFAs), monounsaturated fatty acids (MUFAs), MUFAs/PUFAs). Univariate analyses followed by multivariate analyses were applied to the adulteration investigation. As a result, the univariate analyses demonstrated that higher contents of eicosanoic acid, docosanoic acid, tetracosanoic acid, and SFAs were the peculiarities of peanut adulteration and higher levels of linolenic acid, 11-eicosenoic acid, erucic acid, and nervonic acid the characteristics of rapeseed adulteration. Then, PLS-LDA made the detection of adulteration effective with a 1% detection limit and 90% prediction ability; a Monte Carlo tree identified the type of adulteration with 85% prediction ability.

KEYWORDS: *olive oil, adulteration, univariate analysis, multivariate analysis, PLS-LDA, Monte Carlo tree*

INTRODUCTION

The olive tree is one of the most important crops in Mediterranean countries. Olive oil has the properties of preventing cardiovascular disease, neurological disorders, and breast and colon cancers due to its well-balanced fatty acid composition and natural antioxidants.[1] As olive oil is usually much more expensive than other edible oils, its adulteration is generally motivated to maximize profits by replacing olive oil with cheaper vegetable oils, which may be a serious problem for regulatory agencies and may threaten the health of consumers.

Detection of olive oil adulteration is a difficult and challenging analytical problem because olive oil consists of complex mixtures of triacylglycerols (TAGs), partial glycerides, hydrocarbons, tocopherols, pigments, sterols, alcohols, triterpene acids, volatile compounds, phenolic compounds, and phospholipids.[2,3] Much work has been done to detect and quantify other vegetable oils in olive oil. Spectroscopic methods are commonly used due to their rapid and nondestructive advantages, such as total synchronous fluorescence (TSyF) spectra,[4] mid-infrared,[5] FT-IR,[6] NMR,[7] and Raman.[2]

However, spectroscopy can only give information related to the whole fingerprinting characteristics, whereas information regarding its components cannot be obtained. Analytical methods focused on the component difference between pure and adulterated olive oil are highly demanded. Those methods can be divided into two groups: one is based upon the analysis of a particular compound such as filbertone,[8] and the other is to analyze a specific fraction of the oil sample, such as sterols,[9] volatile compounds,[10] and fatty acids.[11−13] As TAGs, composed of three esterified fatty acids with an attached glycerol backbone, are the main component of vegetable oil (∼95%)[14] and because it is the well-balanced fatty acid composition that confers to the olive oil its high nutritional values, the determination of fatty acids with GC or GC-MS becomes favorable.

Most studies based on the fatty acid compositions were focused only on the detection of adulterations without considering the type of vegetable oil involved. Only Capote et al.[13] investigated different kinds of adulterants (sunflower, corn, peanut, coconut oils) in olive oil. However, despite its effective identification, only eight different samples of pure olive oil were considered, which could not eliminate the variability between them,[15−18] and also peak areas were used to do the quantitative analysis without the use of an internal standard. More importantly, it was not investigated which fatty acids the adulterated samples changed significantly.

In this study, more kinds of pure olive oils and different types of adulterants were considered to be able to construct a more effective model. As for the internal standard, heptadecanoic acid methyl ester is commonly used for this purpose. However, because olive oil contains heptadecanoic acid (≤0.3%),[19] in this study, tridecanoic acid methyl ester was chosen as the internal standard . Moreover, SIM mode was applied to get a lower detection limit and, thus, to improve the sensitivity of the detector.

The aims of this paper are (1) to detect whether an olive oil sample is adulterated or not, (2) to identify the type of adulterants, and (3) to explore the significant parameters that can distinguish pure olive oil from adulterated olive oils and classify different types of adulterants using the fatty acid profile by GC-MS coupled with chemometrics.

Received: August 2, 2012
Revised: March 23, 2013
Accepted: March 25, 2013
Published: March 25, 2013

dx.doi.org/10.1021/jf4000538 | *J. Agric. Food Chem.* 2013, 61, 3693−3702

MATERIALS AND METHODS

Samples and Chemicals. Fifty nonadulterated olive oil samples were included. Sixteen of these samples were bought in the Chinese market, which were imported from Italy (4), Greece (4), and Spain (8); 20 other samples were provided by the University of Algarve, Portugal, and 14 by the University of Cadiz, Spain. Both Portuguese and Spanish olive oils were collected from different regions.

Six corn oil samples, 4 sunflower oils, 11 rapeseed oils, and 3 peanut oils were purchased from local supermarkets. All of the samples were kept in dark glass bottles and stored at room temperature.

For the investigation of the olive oil adulteration, five different adulteration levels (1, 5, 10, 20, and 50% (w/w)), were prepared for each type of adulterant. Both pure olive oils and adulterants used for the mixing process were randomly selected. These mixtures were analyzed immediately after preparation.

A 37-component fatty acid methyl esters (FAMEs) mix and methyl tridecanoate were purchased from Sigma-Aldrich (Shanghai) Trading Co., Ltd. The solution of 0.4 M NaOH/CH_3OH was freshly prepared in the laboratory by dissolving a reagent grade NaOH in methanol.

Fatty Acid Methyl Esters Preparation. Fatty acid methyl esters were prepared by alkaline transmethylation.[20] Aliquots (50 μL) of sample were spiked with internal standard working solution (50 μL tridecanoic acid methyl ester, 100 g/L). One milliliter of 0.4 M NaOH–CH_3OH was added and reacted for about 5 min in an ultrasonic bath. After the solution turned clear and transparent, the methyl esters were extracted with 1 mL of hexane twice and diluted to a final volume of 2 mL.

GC-MS Instrument and Analytical Conditions. All GC-MS analyses were performed with Shimadzu GC2010A (Kyoto, Japan) gas chromatography instrument coupled with a GCMS-QP2010 quadrupole mass spectrometer (Shimadzu). In the gas chromatographic system, a DB-23 capillary column, 30 m length, 0.25 mm i.d., and 0.25 μm film, consisting of 50% cyanopropyl 50% dimethyl polysiloxane (Agilent) was used. Column temperature was programmed from 120 to 160 °C at the rate of 20 °C/min, from 160 to 190 °C at the rate of 6 °C/min, from 190 to 220 °C at the rate of 20 °C/min, and then held for 6 min at 220 °C. The injection temperature was kept at 250 °C, the carrier gas was helium, and the column flow was 1.0 mL/min. A sample of 1 μL was injected with a split ratio of 500:1.

Mass Spectroscopy Conditions. The ion source temperature was 200 °C, and the interface temperature was 250 °C. Ionization voltage was 70 eV; single-ion monitoring was performed at m/z 55, 67, 74, 79, and 87.

Multivariate Statistical Analysis. CARS-PLS-LDA and Monte Carlo tree (MCTree) were used to detect and identify olive oil adulteration.

CARS-PLS-LDA. Partial least-squares (PLS) is a supervised method, and one of the most commonly used multivariate analysis methods. It is primarily concerned with the transformation of a large set of related variables into a new, smaller set of orthogonal variables, which are called latent variables.

Linear discriminant analysis (LDA) is a statistical technique that can be used for the classification of individuals into mutually exclusive and exhaustive groups, based on a set of independent variables.[21] The latent variables obtained by PLS are usually used as the input data to LDA.

Including too many variables into a model may cause overfitting.[22] Therefore, variable selection is always performed before putting the data into the model. Competitive adaptive reweighted sampling (CARS) is a variable selection method;[23] it selects N subsets of variables by N sampling runs in an iterative manner and finally chooses the subsets with the lowest root-mean-square error of CV (RMSECV) value as the optimal subset.

MCTree. MCTree provides a feasible way to uncover the predictive structure of data by the establishment of cross-predictive models.[24] First, a number of classification trees with different structures are grown by means of the Monte Carlo strategy. As each tree is constructed, classification tree-based sample similarity can be computed to generate a sample proximity matrix. The sample proximity matrix obtained can give an intrinsic measure of similarities between samples. Meanwhile, variable importance ranking allows the discovery of informative variables. Nonlinear mapping algorithms can be employed to convert the high-dimensional space into a low-dimensional space. This approach shares many outstanding advantages; the one we are most attracted to is that it can effectively cope with multiclass data, which are not easily solved by other statistical approaches.

RESULT AND DISCUSSION

Identification and Quantification of FAMEs. The FAMEs were identified by retention time and mass spectra comparison with those of the corresponding standards. Finally, 22 fatty acid methyl esters were identified.

The ratios of the FAMEs' areas to that of the internal standard were calculated for the quantitative analysis. Besides the single component, total saturated fatty acids (SFAs), polyunsaturated fatty acids (PUFAs), monounsaturated fatty acids (MUFAs), and the ratios of oleic/linoleic acid, linoleic/linolenic acid, and MUFAs/PUFAs were also under evaluation.

To ensure the quality of the olive oils used to obtain the mixtures, PCA was first applied to olive oils, and the score plot is shown in Figure 1. Some samples were located far from the concentrated region, and so discarded for this purpose. Only samples located inside the ellipse were selected as candidates to make the mixtures. To further minimize the effect of different varieties of olive oil in the adulteration, 30 olive oils, which were obtained from different regions, were finally selected for the mixing process. Eleven of these samples were bought in the Chinese market, which were imported from Italy (4), Greece (4), and Spain (4); 10 other samples were provided by the University of Algarve, Portugal, and 9 by the University of Cadiz, Spain.

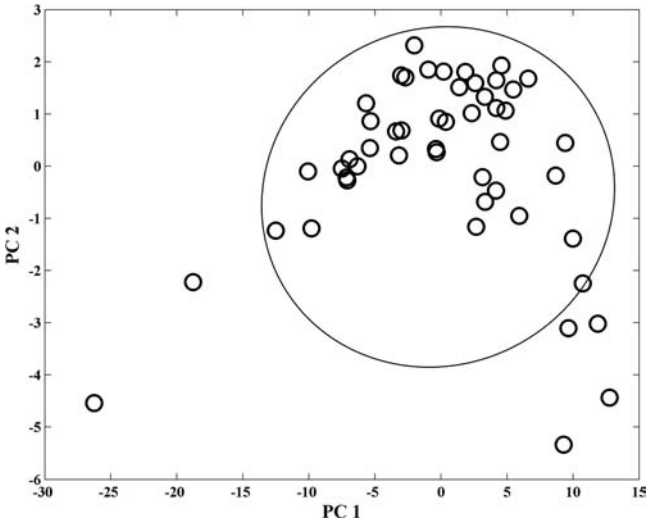

Figure 1. PCA scores plot of the olive oils.

The average total ion current (TIC) chromatograms for the 30 pure olive oils and the average TIC for each type of other analyzed oils are shown in Figure 2. The fatty acid composition of the five pure oils is shown in Table 1.

As shown in Figure 2 and Table 1, each of the four cheaper vegetable oils is different from the olive oil in the fatty acid profile, and they change the olive oil fatty acid profile in different ways.

Univariate Analysis. The change trends of each variable for different adulteration types and at different adulteration levels are shown in Figure 3, where each subfigure represents a variable and

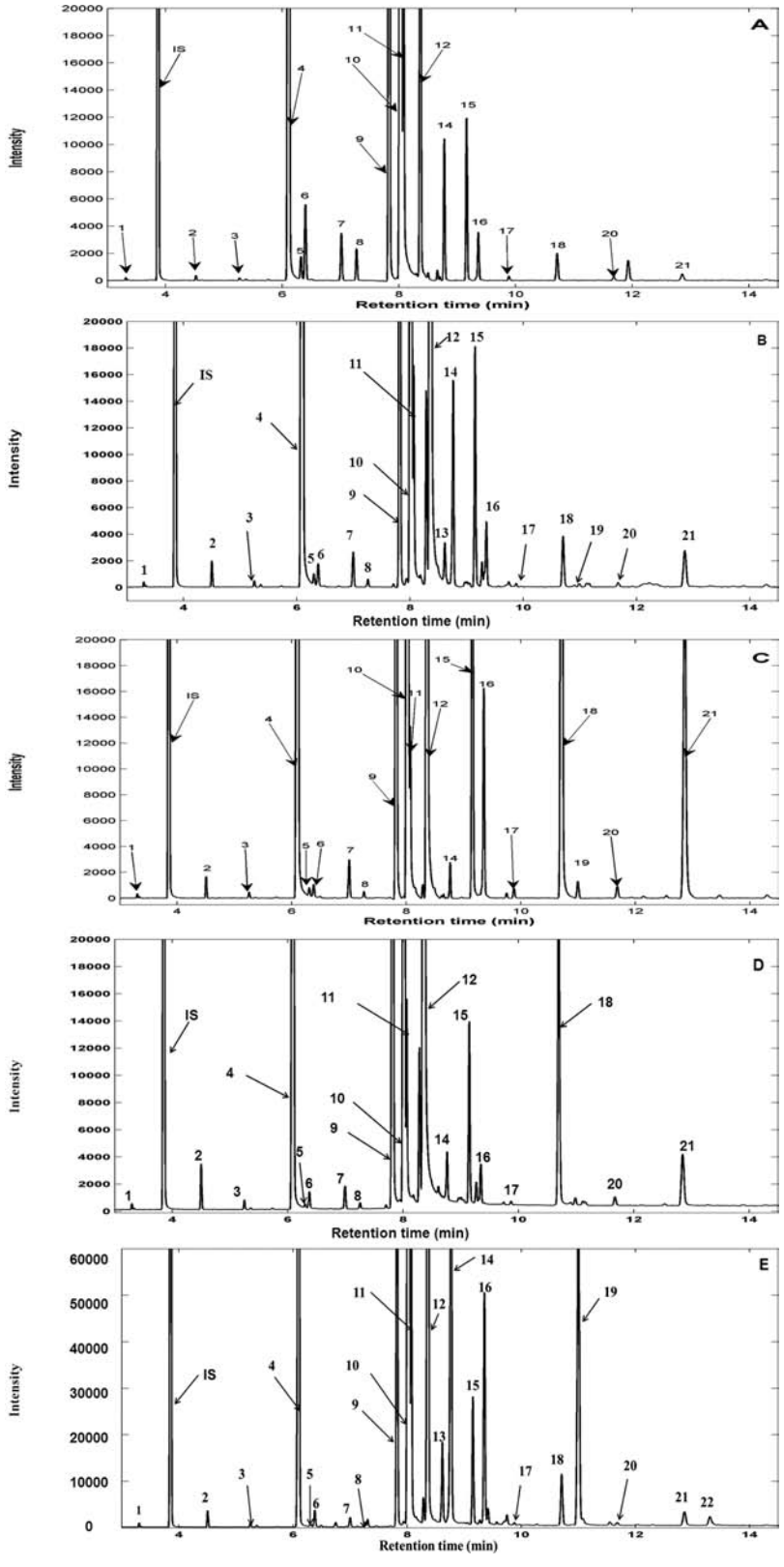

Figure 2. Average chromatograms (TICs) of each kind of pure oil (A, olive oil; B, corn oil; C, peanut oil; D, sunflower oil; E, rapeseed oil). Peaks: 1, 12:0; 2, 14:0; 3, 15:0; 4, 16:0; 5, 16:1n-9c; 6, 16:1n-7c; 7, 17:0; 8, 17:1n-7c; 9, 18:0; 10, 18:1n-9c; 11, 18:1n-7c; 12, 18:2n-6c; 13, 18:3n-6c; 14, 18:3n-3c; 15, 20:0; 16, 20:1n-9c; 17, 21:0; 18, 22:0; 19, 22:1n-9c; 20, 23:0; 21, 24:0; 22, 24:1n-9c; IS, internal standard.

each line represents a type of adulteration. Most of the variables have undergone a change to a different degree.

It was observed that all of the adulterated samples increased the contents of myristic acid, linoleic acid, and PUFAs and reduced the levels of 7-hexadecenoic acid, 9-hexadecenoic acid, 10-heptadecenoic acid, and oleic acid as well as the oleic/linoleic acid ratio, MUFAs, and MUFAs/PUFAs ratio. Besides, peanut–olive significantly raised the contents of eicosanoic acid,

dx.doi.org/10.1021/jf4000538 | J. Agric. Food Chem. 2013, 61, 3693−3702

Table 1. Summary of the Relative Levels of Different Fatty Acids in the Five Pure Vegetable Oils[a]

component	name	olive	corn	peanut	sunflower	rapeseed
12:0	dodecanoic acid	0.0005 ± 0.0001	0.0007 ± 0.0001	0.0007 ± 0.0001	0.0008 ± 0.0003	0.0015 ± 0.0003
14:0	myristic acid	0.0013 ± 0.0002	0.0043 ± 0.0001	0.0051 ± 0.0021	0.0072 ± 0.0009	0.0072 ± 0.0035
15:0	pentadecanoic acid	0.0006 ± 0.0002	0.0012 ± 0.0001	0.0013 ± 0.0003	0.0017 ± 0.0001	0.0023 ± 0.0003
16:0	palmitic acid	1.2046 ± 0.1374	1.5218 ± 0.1085	1.3437 ± 0.0275	0.6897 ± 0.0322	0.6278 ± 0.3395
16:1n-9c	7-hexadecenoic acid	0.0062 ± 0.0010	0.0031 ± 0.0011	0.0027 ± 0.0005	0.0010 ± 0.0004	0.0019 ± 0.0003
16:1n-7c	9-hexadecenoic acid	0.0363 ± 0.0118	0.0048 ± 0.0005	0.0038 ± 0.0007	0.0034 ± 0.0006	0.0091 ± 0.0014
17:0	heptadecanoic acid	0.0092 ± 0.0036	0.0075 ± 0.0005	0.0092 ± 0.0014	0.0047 ± 0.0005	0.0053 ± 0.0016
17:1n-7c	10-heptadecenoic acid	0.0070 ± 0.0028	0.0016 ± 0.0001	0.0019 ± 0.0005	0.0012 ± 0.0002	0.0025 ± 0.0005
18:0	stearic acid	0.3239 ± 0.0569	0.2051 ± 0.0075	0.4409 ± 0.0922	0.5253 ± 0.0931	0.2400 ± 0.0601
18:1n-9c	oleic acid	4.0339 ± 0.2082	1.6202 ± 0.1084	2.0994 ± 0.4530	1.2163 ± 0.0601	2.7017 ± 0.5855
18:1n-7c	11-octadecenoic acid	0.1395 ± 0.0214	0.0364 ± 0.0013	0.0461 ± 0.0221	0.0319 ± 0.0046	0.1489 ± 0.0316
18:2n-6c	linoleic acid	0.2909 ± 0.0570	2.7993 ± 0.1359	2.1244 ± 0.4255	3.3801 ± 0.3064	1.1978 ± 0.5347
18:3n-6c	γ-linolenic acid	0.0000 ± 0.0000	0.0077 ± 0.0048	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0421 ± 0.0245
18:3n-3c	linolenic acid	0.0364 ± 0.0034	0.0394 ± 0.0051	0.1000 ± 0.1536	0.0085 ± 0.0036	0.4055 ± 0.0845
20:0	eicosanoic acid	0.0410 ± 0.0048	0.0449 ± 0.0034	0.1357 ± 0.0498	0.0325 ± 0.0041	0.0621 ± 0.0077
20:1n-9c	11-eicosenoic acid	0.0122 ± 0.0015	0.0132 ± 0.0005	0.0356 ± 0.0120	0.0425 ± 0.0006	0.1168 ± 0.0600
21:0	heneicosanoic acid	0.0014 ± 0.0002	0.0008 ± 0.0002	0.0023 ± 0.0005	0.0008 ± 0.0001	0.0015 ± 0.0005
22:0	docosanoic acid	0.0107 ± 0.0022	0.0132 ± 0.0021	0.2143 ± 0.0966	0.0756 ± 0.0083	0.0347 ± 0.0054
22:1n-9c	erucic acid	0.0000 ± 0.0000	0.0009 ± 0.0008	0.0052 ± 0.0013	0.0000 ± 0.0000	0.2361 ± 0.2119
23:0	tricosanoic acid	0.0017 ± 0.0003	0.0013 ± 0.0002	0.0035 ± 0.0005	0.0025 ± 0.0002	0.0023 ± 0.0008
24:0	tetracosanoic acid	0.0040 ± 0.0008	0.0141 ± 0.0017	0.0879 ± 0.0403	0.0187 ± 0.0011	0.0143 ± 0.0010
24:1n-9c	nervonic acid	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0097 ± 0.0053
18:2n-6c/18:3n-3c		8.0104 ± 1.4671	72.2838 ± 12.2890	138.8884 ± 142.8961	473.6135 ± 262.1779	3.3338 ± 2.6452
18:1n-9c/18:2n-6c		14.4615 ± 3.2772	0.5790 ± 0.0320	1.0372 ± 0.3737	0.3628 ± 0.0446	2.5586 ± 0.8394
SFAs	saturated fatty acids	1.5989 ± 0.1299	1.8148 ± 0.1150	2.2445 ± 0.1718	1.3593 ± 0.0929	0.9989 ± 0.4003
PUFAs	polyunsaturated fatty acids	0.3273 ± 0.0586	2.8464 ± 0.1325	2.2245 ± 0.5791	3.3887 ± 0.3050	1.6454 ± 0.4981
MUFAs	monounsaturated fatty acids	4.2350 ± 0.2115	1.6801 ± 0.1095	2.1947 ± 0.4437	1.2614 ± 0.0657	3.2267 ± 0.5978
MUFAs/PUFAs		13.4006 ± 2.7742	0.5905 ± 0.0327	1.0563 ± 0.4080	0.3752 ± 0.0463	2.1210 ± 0.6000

[a]The results indicate the ratios of the FAMEs' areas to that of the internal standard, and the data are presented as the mean ± SD.

docosanoic acid, tetracosanoic acid, and SFAs, whereas rapeseed−olive elevated the contents of linolenic acid, γ-linolenic acid, 11-eicosenoic acid, erucic acid, and nervonic acid. It is worth noting that corn−olive, peanut−olive, and sunflower−olive had a higher ratio of linoleic/linolenic acid than pure olive oil, whereas rapeseed−olive slightly lowered this ratio.

The oleic/linoleic acid ratio is frequently used as a stability parameter, and the oils with higher ratio are also those with a higher oxidative stability.[1,25] As for the adulterated oils, a decrease in the oxidative stability was observed. Moreover, linoleic (18:2n-6c) and linolenic (18:3n-3c) acids are the starting points to manufacture other n-6 and n-3 fatty acids.[26,27] They are involved in similar metabolic systems for the synthesis of their active metabolites, in which linolenic acid (n-3) has a higher affinity for those enzyme systems. Any alteration in the concentration of n-3 fatty acids will have an effect on the metabolism of n-6 fatty acids. The daily recommendation intake for the n-6/n-3 ratio is between 4/1 and 10/1.[26] This ratio in pure olive oil was around 8, whereas the adulterated oils mostly increased this ratio significantly, except for the rapeseed−olive. As for PUFAs, their high consumption induces a decrease in HDL cholesterol, along with the known effects on the oxidizability of LDL,[26,28,29] so they should not exceed the recommended 7% of calories. Conversely, MUFAs may increase HDL cholesterol, and its high amounts confer to the olive oil a high nutritional value.[30] Therefore, the higher levels of PUFAs and the lower levels of MUFAs in the adulterated oils are not a good sign.

Furthermore, the higher content of erucic acid was characteristic of rapeseed−olive. However, erucic acid has been identified as an antinutritional compound and has been proscribed since previous studies found an association between erucic acid consumption and cardiac lipidosis in rats.[31−34]

Multivariate Analysis. The application of univariate analysis is useful for the study of olive oil adulteration; it can help us to clearly know the components' variations in adulterated olive oils, but it alone is not enough. It is difficult to define a quality index to detect an adulteration, because the frauds could find a way to make this signal index acceptable, whereas a lower overall nutritional value would still occur in the product. Therefore, the multivariate analysis is necessary.

Detection of the Adulteration. Autoscaling was used to preprocess the data before multivariate statistical analysis.

Adulterated samples were first grouped according to the type of vegetable oil used in the adulteration. PCA was performed separately for each adulterated group with nonadulterated olive oil samples and indicated that the adulterated samples were different from the pure olive oil (figure not shown). To more clearly visualize the discrimination of adulterated samples from pure olive oil samples, the CARS-PLS-LDA model was constructed.

For each model, one-fifth of the samples (six from pure olive oil and one from each adulteration level) were randomly selected as test set, and the remaining samples were combined to get the training set to construct the models.

Variable selection was first carried out in the training set. The variables selected for constructing models were palmitic acid, stearic acid, oleic acid, and the ratios of oleic/linoleic acid and MUFAs/PUFAs for corn−olive; stearic acid, oleic acid, 11-octadecenoic acid, eicosanoic acid, docosanoic acid, and

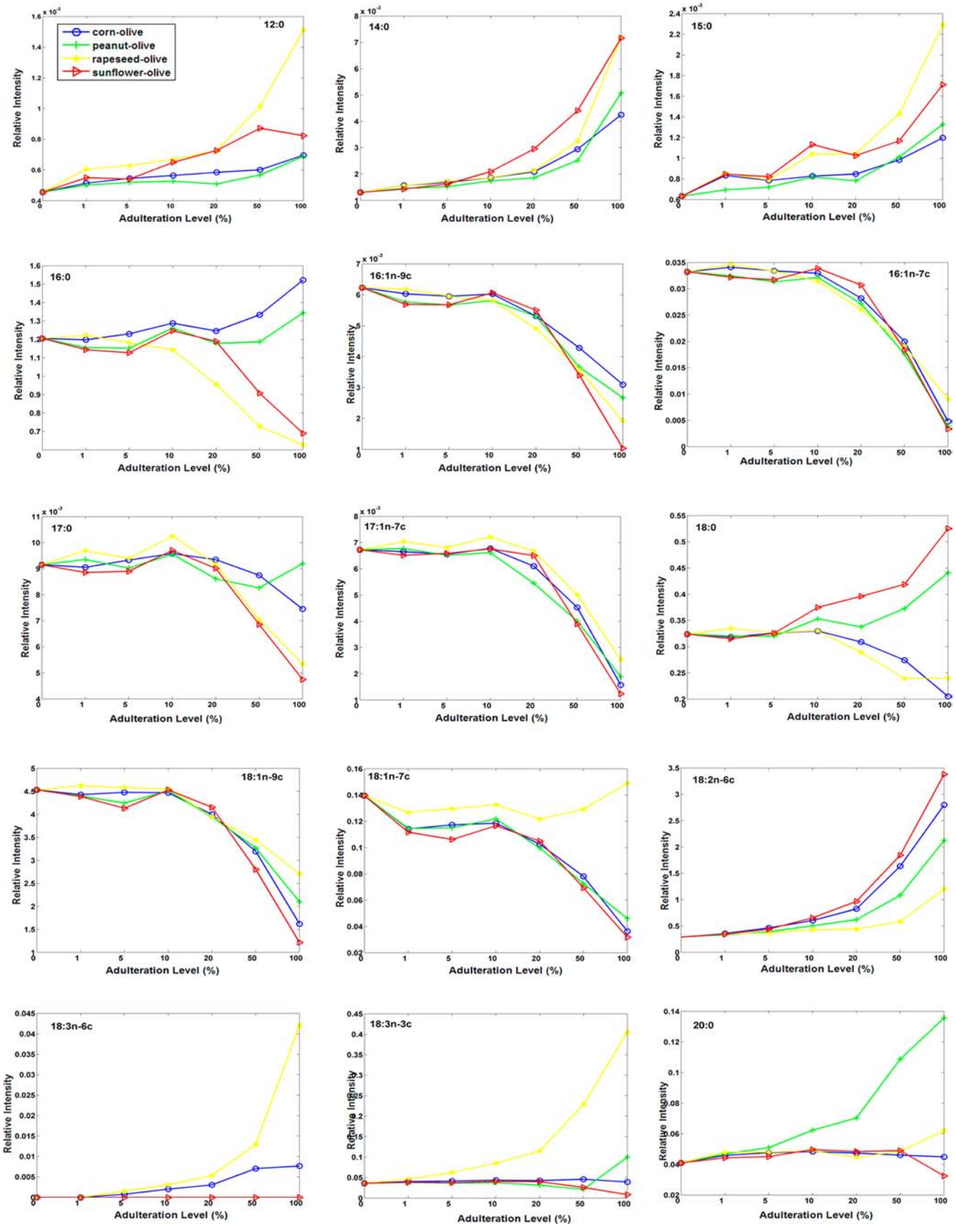

Figure 3. continued

dx.doi.org/10.1021/jf4000538 | *J. Agric. Food Chem.* 2013, 61, 3693−3702

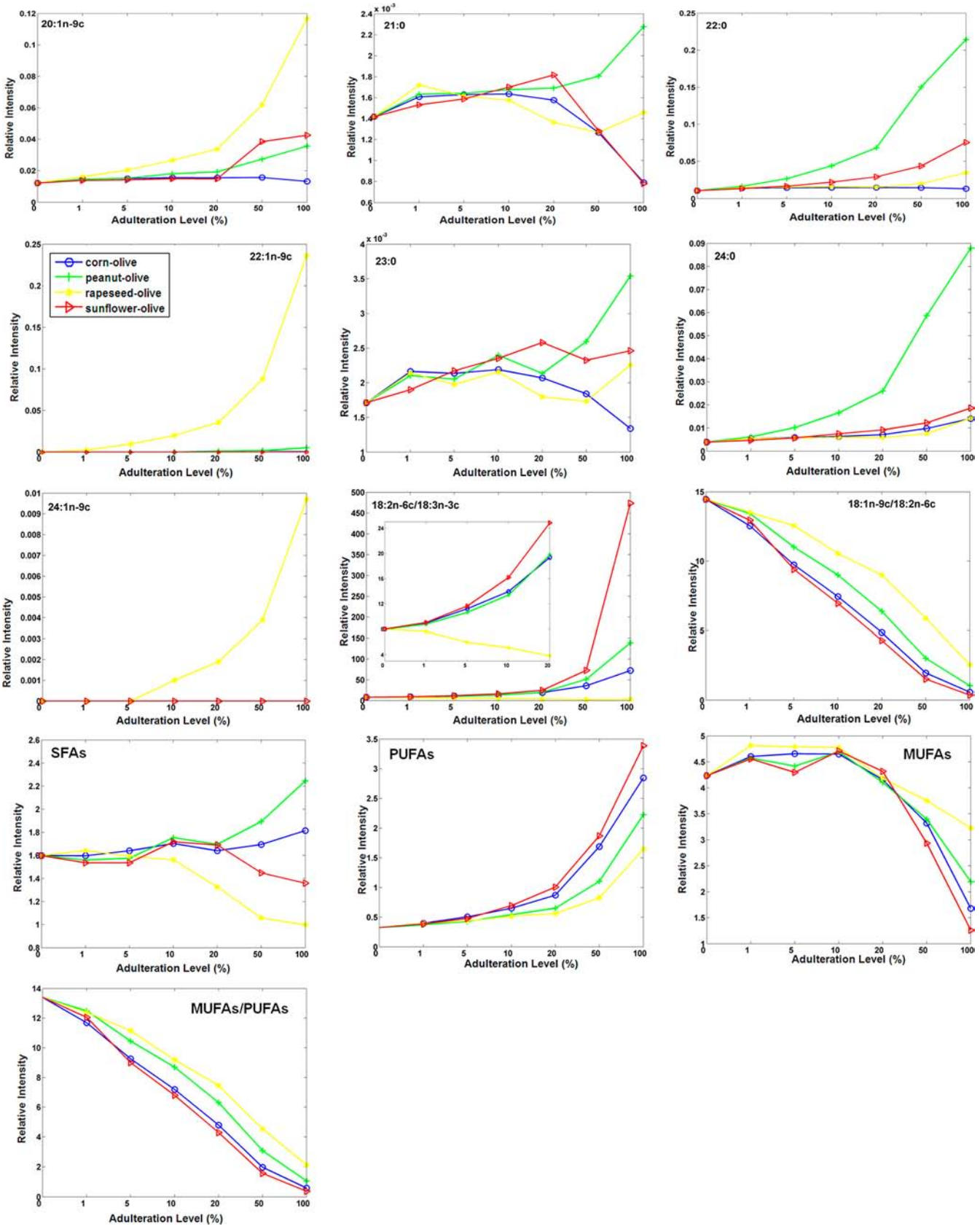

Figure 3. Relative intensity of different variables at different adulteration level: (blue) corn−olive; (green) peanut−olive; (yellow) rapeseed−olive; (red) sunflower−olive. 0% represents pure olive oil, and 100% represents pure adulterant.

tetracosanoic acid for peanut−olive; 7-hexadecenoic acid, 9-hexadecenoic acid, oleic acid, erucic acid, and MUFAs for rapeseed−olive; and 9-hexadecenoic acid, heptadecanoic acid, oleic acid, 11-octadecenoic acid, MUFAs, and the ratios of oleic/

dx.doi.org/10.1021/jf4000538 | J. Agric. Food Chem. 2013, 61, 3693−3702

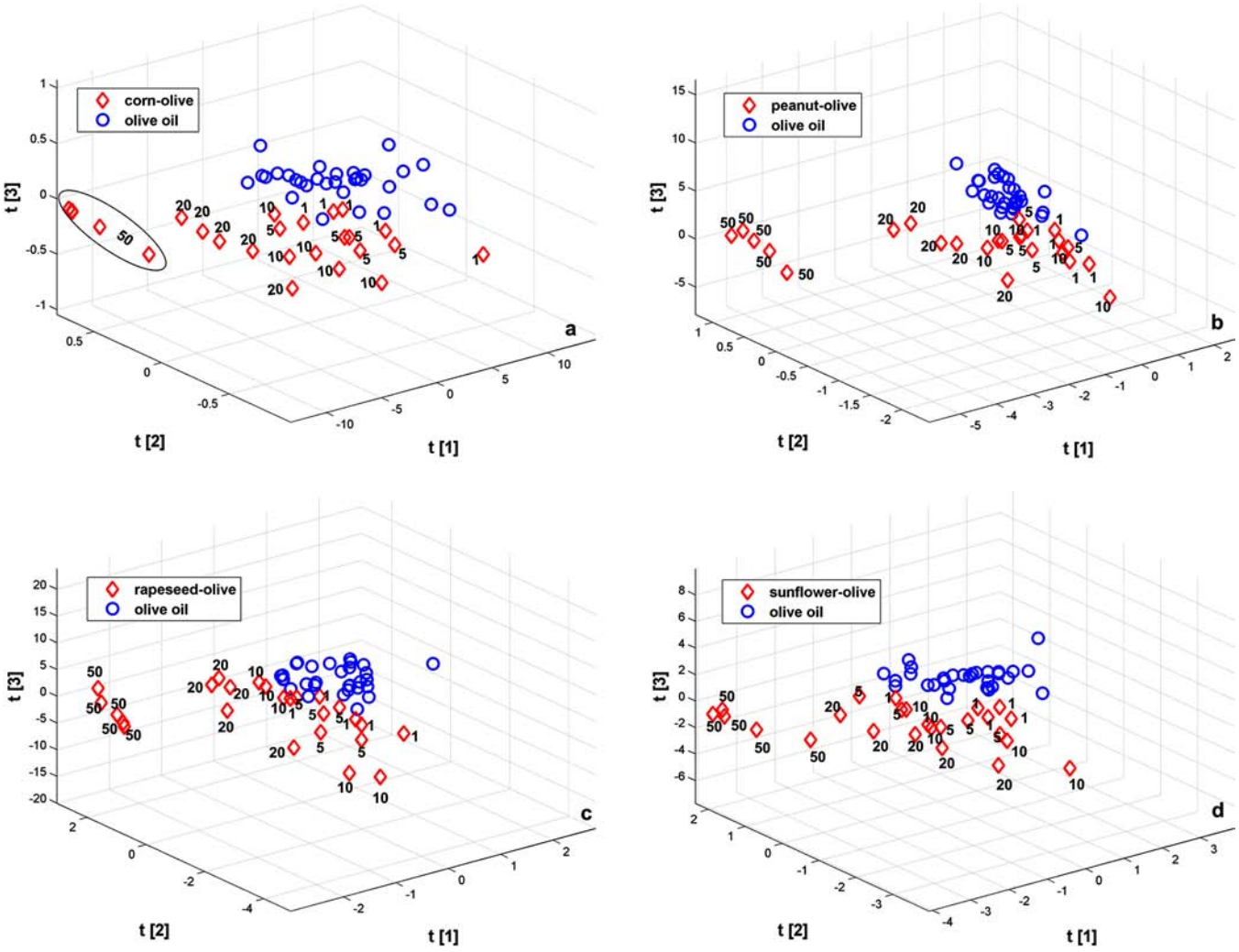

Figure 4. PLS-LDA scores plots of (a) corn−olive (red), (b) peanut−olive (red), (c) rapeseed−olive (red), and (d) sunflower−olive (red) versus nonadulterated olive oil samples (blue). 1, 5, 10, 20, and 50 represent 1, 5, 10, 20, and 50% adulteration levels, respectively.

Table 2. Classfication of Different Groups by PLS-LDA Method

group vs pure olive oil	recognition rate[a] (%)	prediction rate[b] (%)	sensitivity[c]	specificity[d]
corn−olive	97.73 (43/44)	100.00 (11/11)	1.000	0.9500
peanut−olive	97.73 (43/44)	90.91 (10/11)	0.9583	1.0000
rapeseed−olive	97.73 (43/44)	90.91 (10/11)	0.9583	1.000
sunflower−olive	100.00 (44/44)	100.00 (11/11)	1.0000	1.000
1% adulterated olive (sunflower−olive as the test set)	95.00 (38/40)	90.00 (9/10)	0.9200	1.0000
1% adulterated olive (rapeseed−olive as the test set)	90.00 (36/40)	80.00 (8/10)	0.8400	1.0000
1% adulterated olive (peanut−olive as the test set)	100.00 (40/40)	100.00 (10/10)	1.000	1.0000
1% adulterated olive (corn−olive as the test set)	100.00 (40/40)	90.00 (9/10)	1.000	1.0000

[a]Recognition rate is the correct classification of the training set. [b]Prediction rate is the rate of the correct classification of the predicting set. [c]Sensitivity is the number of true positives classified as positive. [d]Specificity is the number of true negative classified as negative.

linoleic acid and MUFAs/PUFAs for sunflower−olive with pure olive, respectively.

The selection results are in good accordance with the univariate analysis previously performed, which demonstrates that eicosanoic acid, docosanoic acid, and tetracosanoic acid are the particularities of peanut−olive and that erucic acid is a particularity of rapeseed−olive. As for other variables selected by CARS, their particularities are not notable, but it is the synergy effect among the variables that makes the adulterated oils separate from pure olive oil.

In these models constructed by the selected variables, a good separation can be seen between pure olive oil with each kind of adulteration. Moreover, each grouping of sample pointed in adulterated olive oil class model represents different adulteration percentages in ascending order from right to left. The PLS-LDA scores plots of these models are shown in Figure 4, and the classification results are summarized in Table 2. Then, the test sets were used to validate these models, and the prediction results are also shown in Table 2, which indicate quite a successful discrimination and high prediction ability of the models.

dx.doi.org/10.1021/jf4000538 | J. Agric. Food Chem. 2013, 61, 3693−3702

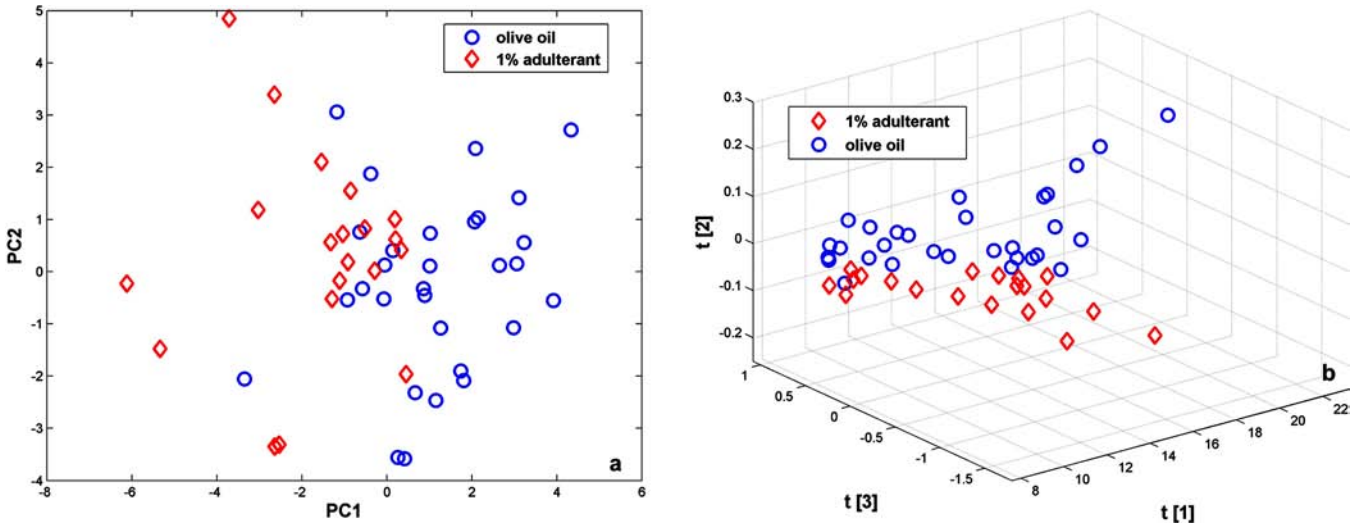

Figure 5. PCA (a) and PLS-LDA (b) scores plots of 1% adulterated olive oils versus nonadulterated olive oil samples (blue, nonadulterated; red, 1% adulterated olive oil samples).

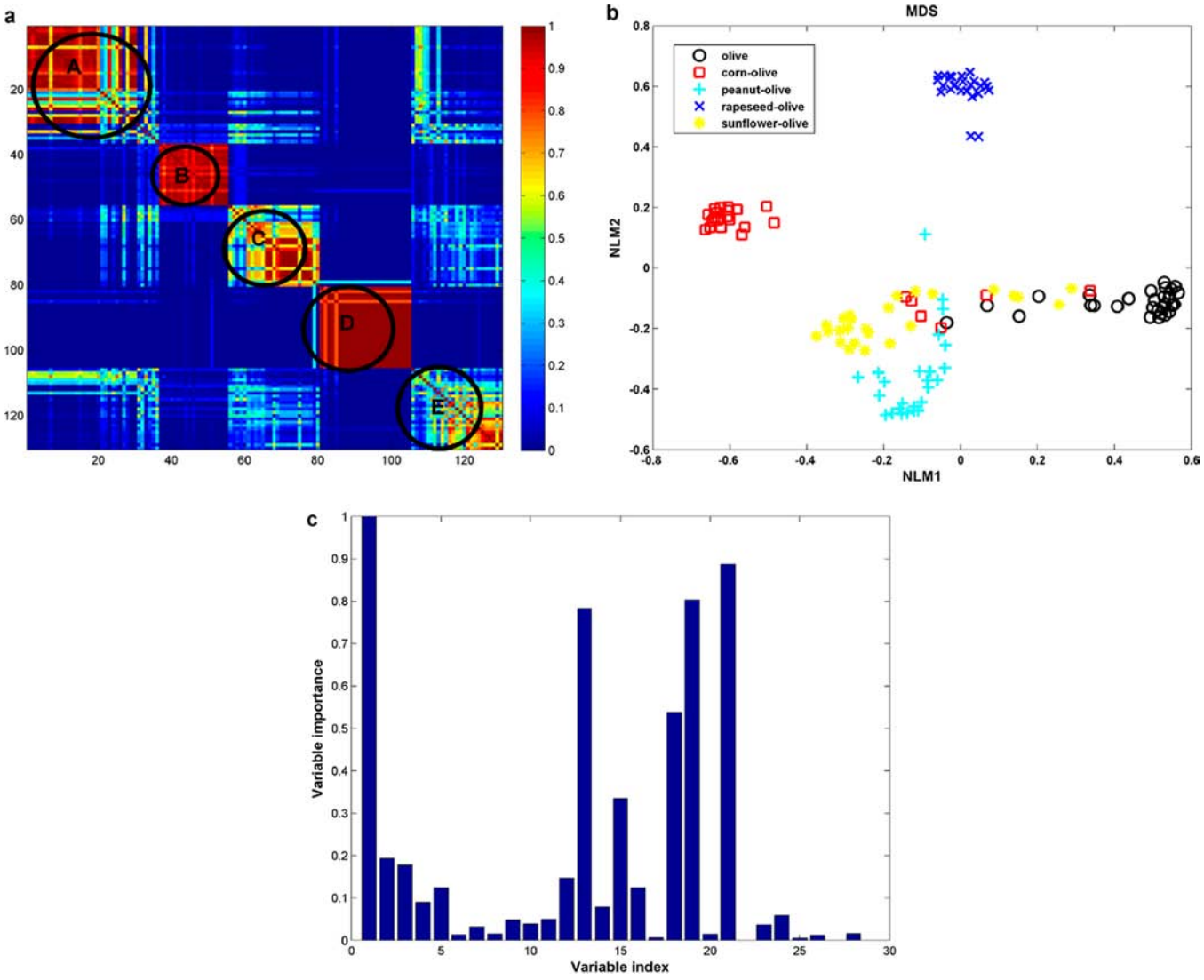

Figure 6. (a) Sample proximity plot of the oils' data: A, pure olive oil; B, corn−olive; C, peanut−olive; D, rapeseed−olive; E, sunflower−olive. (b) Nonlinear mapping plot of the oils data. (c) Variable importance measure obtained by the MCTree.

dx.doi.org/10.1021/jf4000538 | J. Agric. Food Chem. 2013, 61, 3693−3702

Next, to determine if this model is able to discriminate an olive oil adulterated with an unknown adulterant from pure olive oils, the 1% adulterated oils, regardless of the type of adulterant, were grouped as one adulterant class and subjected to statistical analysis. The union set of the selected variables for each kind of adulteration above (palmitic acid, 7-hexadecenoic acid, 9-hexadecenoic acid, heptadecanoic acid, stearic acid, oleic acid, 11-octadecenoic acid, eicosanoic acid, docosanoic acid, erucic acid, tetracosanoic acid, the ratio of oleic/linoleic acid, MUFAs, and the ratio of MUFAs/PUFAs) was used to construct the model.

Modeling adulterated oil samples together means more variable chemical information, making them more difficult to be placed in the same group. According to the PCA scores plots (Figure 5a), the pure olive oil tended to cluster to the right, whereas the adulterated samples tended to move to the left. However, some samples were plotted in the region where they cannot be classified as pure or adulterated, and so PLS-LDA was applied (Figure 5b). In this model, 3 types of adulteration samples and 25 pure olive oils were treated as the training set to construct the model; the remaining samples were the test set. In such a case, the adulteration type in the test set was independent from the three adulteration types in the training set. Four different test sets were selected, and the results are also shown in Table 2, which demonstrated a very good separation between the 1% adulterant oils and the pure olive oils.

To sum up, PLS-LDA was able to identify the adulterations in olive oil regardless of the type of adulterant, and even when this adulterant was unknown. The selected variables proved to be important for olive oil quality control.

Identification of Adulterants. The PLS-LDA model enabled us to detect whether a sample was adulterated or not, but it cannot provide reliable information about what type of vegetable oil has been used as adulterant. Thus, the next step was to identify the adulterants using the MCTree.[24]

The parameters of MCTree were manually set to ntree = 500, R = 0.5. Figure 6a shows the sample proximity plot of the oils' data. Clearly, the corn−olive denoted B and the rapeseed−olive denoted D were significantly different from the other three groups. Among the other three groups denoted A, C and E, there exists some overlapping to some extent. To directly and conveniently observe the classification of the data, the nonlinear mapping plot (Figure. 6b) was plotted to directly visualize the data in low-dimensional space. The results of Figure 6a coupled with Figure 6b sufficiently indicated that the type of adulterants was possible to identify using the fatty acid profile by MCTree.

To demonstrate the predictivity of this model, the training and validation sets (4/5 and 1/5) were therefore divided into five groups: one for pure olive oil and one for each type of adulteration (corn−olive, peanut−olive, rapeseed−olive, and sunflower−olive). Thus, the training samples were used to establish the MCTree model and the validation samples to evaluate the model. Pure olive oil, rapeseed−olive, and sunflower−olive oil were all accurately predicted. For corn−olive and peanut−olive groups, the numbers of misclassification samples were 2 and 1, respectively. Finally, the total prediction rate for the five patterns was 88.46% (23/26).

The variable importance calculated by the MCTree is shown in Figure 6c, from which one can see that there are six fatty acids (dodecanoic acid, γ-linolenic acid, eicosanoic acid, docosanoic acid, erucic acid, tetracosanoic acid) that can be very predictive for the oils' data. γ-Linolenic acid and erucic acid are the peculiarities of rapeseed−olive, and eicosanoic acid, docosanoic acid, and tetracosanoic acid are the ones of peanut−olive. The change tendency of dodecanoic acid was also different among different adulteration types.

The ratios of oleic/linoleic acid, linoleic/linolenic acid, and MUFAs/PUFAs and PUFAs were not important for this model, which seems to conflict with variable selection results using CARS. As all of the adulterated oils, regardless of adulteration type, significantly changed the four variables in the same direction, this may mean that the four variables can discriminate pure olive oil from adulterated oils but cannot distinguish the type of adulterant, making them trivial in this point.

All in all, an analytical method using the fatty acid profile to make the detection and identification of extra virgin olive oil adulteration was proposed in this paper. This enabled us to (1) detect whether an olive oil sample was adulterated or not, (2) identify the type of adulterant, and (3) select significant variables to distinguish pure olive oil from adulterated olive oils and to classify different types of adulterants. As a result, it was found that each kind of adulteration had its particularities and changed the content of different fatty acids, whereas they also had some common characteristics, such as that all of them changed significantly the nutritional ratio. The models constructed were able to overcome the variabilities among pure olive oils and permitted the detection of adulteration and the identification of the type of adulterant with prediction abilities above 90 and 85%, respectively, with a detection limit of 1%.

AUTHOR INFORMATION

Corresponding Author

*Phone: +86-731-88830824. Fax: +86-731-88830831. E-mail: yizeng_liang@263.net.

Funding

This work was financially supported by the National Nature Foundation Committee of the People's Republic of China (Grant 21075138) and the graduate degree thesis Innovation Foundation of Central South University (No. 2011ssxt081). We thank the European Commission through the Erasmus Mundus Master in Quality in Analytical Laboratories, FPA 2008-0095, for funding M.D.F.'s mobility, which made the olive oil samples available.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the University of Algarve, Portugal, and the University of Cadiz, Spain, for helping us to collect the olive oil samples.

ABBREVIATIONS USED

SFAs, saturated fatty acids; PUFAs, polyunsaturated fatty acids; MUFAs, monounsaturated fatty acids; FAMEs, fatty acid methyl esters; TIC, total ion current; 12:0, dodecanoic acid; 14:0, myristic acid; 15:0, pentadecanoic acid; 16:0, palmitic acid; 16:1n-9c, 7-hexadecenoic acid; 16:1n-7c, 9-hexadecenoic acid; 17:0, heptadecanoic acid; 17:1n-7c, 10-heptadecenoic acid; 18:0, stearic acid; 18:1n-9c, oleic acid; 18:1n-7c, 11-octadecenoic acid; 18:2n-6c, linoleic acid; 18:3n-6c, γ-linolenic acid; 18:3n-3c, linolenic acid; 20:0, eicosanoic acid; 20:1n-9c, 11-eicosenoic acid; 21:0, heneicosanoic acid; 22:0, docosanoic acid; 22:1n-9c, erucic acid; 23:0, tricosanoic acid; 24:0, tetracosanoic acid; 24:1n-9c, nervonic acid; IS, internal standard

REFERENCES

(1) Matos, L. C.; Cunha, S. C.; Amaral, J. S. Chemometric characterization of three varietal olive oil (cvs. Cobrancosa, Madural and Verdeal Transmontana) extracted from olives with different maturation indices. *Food Chem.* **2007**, *102*, 406−414.

(2) Frankel, E. N. Chemistry of extra virgin olive oil: adulteration, oxidative stability, and antioxidants. *J. Agric. Food Chem.* **2010**, *58*, 5991−6006.

(3) Angerosa, F.; Campestre, C.; Glasante, L. Chapter 7: analysis and authentication. In *Olive Oil, Chemistry and Technology*, 2nd ed.; Boskou, D., Ed.; AOCS Press: Champaign, IL, 2006; pp 113−172.

(4) Poulli, K. I.; Mousdis, G. A.; Georgiou, C. A. Rapid synchronous fluorescence method for virgin olive oil adulteration assessment. *Food Chem.* **2007**, *105*, 369−375.

(5) Gurdeniz, G.; Ozen, B. Detection of adulteration of extra-virgin olive oil by chemometric analysis of mid-infrared spectral data. *Food Chem.* **2009**, *116*, 519−525.

(6) García, M. J. L.; Ramos, G. R.; Martínez, J. M. H. Authentication of extra virgin olive oils by Fourier-transform infrared spectroscopy. *Food Chem.* **2010**, *118*, 78−83.

(7) Šmejkalová, D.; Piccolo, A. High-power gradient diffusion NMR spectroscopy for the rapid assessment of extra-virgin olive oil adulteration. *Food Chem.* **2010**, *118*, 153−158.

(8) Flores, G. F.; Castillo, M. L. R.; Herraiz, M. Study of the adulteration of olive oil with hazelnut oil by on-line coupled high performance liquid chromatographic and gas chromatographic analysis of filbertone. *Food Chem.* **2006**, *97*, 742−749.

(9) Ismail, K. M. A.; Alsaed, A. K.; Ahmad, R. Detection of olive oil adulteration with some plant oils by GLC analysis of sterols using polar column. *Food Chem.* **2010**, *121*, 1255−1259.

(10) Szkudlarz, S. M.; Jeleń, H. H. The potential of different techniques for volatile compounds analysis coupled with PCA for detection of the adulteration of olive oil with hazelnut oil. *Food Chem.* **2008**, *110*, 751−761.

(11) Vázquez, J. G.; Falcón, M. S. G.; Gándara, J. S. Control of contamination of olive oil by sunflower seed oil in bottling plants by GC-MS of fatty acid methyl esters. *Food Control* **2003**, *14*, 463−467.

(12) Christopoulou, E.; Lazaraki, M.; Komaitis, K. Effectiveness of determinations of fatty acids and triglycerides for the detection of adulteration of olive oils with vegetable oils. *Food Chem.* **2004**, *84*, 463−474.

(13) Capote, F. P.; Jiménez, J. R.; Castro, M. D. L. Sequential (step-by step) detection, identification and quantitation of extra virgin olive oil adulteration by chemometric treatment of chromatographic profiles. *Anal. Bioanal. Chem.* **2007**, *388*, 1859−1865.

(14) Chapagain, B. P.; Wiesman, Z. MALDI-TOF/MS fingerprinting of triacylglycerols (TAGs) in olive oil produced in the Israeli Negev Desert. *J. Agric. Food Chem.* **2009**, *57*, 1135−1142.

(15) Issaoui, M.; Flamini, G.; Brahmi, F. Effect of the growing area conditions on differentiation between Chemlali and Chétoui olive oils. *Food Chem.* **2010**, *119*, 220−225.

(16) Aguilera, M. P.; Beltrán, G.; Ortega, D. Characterisation of virgin olive oil of Italian olive cultivars: 'Frantoio' and 'Leccino', grown in Andalusia. *Food Chem.* **2005**, *89*, 387−391.

(17) Lanteri, S.; Armanino, C.; Perri, E. Study of oils from Calabrian olive cultivars by chemometrics methods. *Food Chem.* **2002**, *76*, 501−507.

(18) Stefanoudaki, E.; Williams, M.; Chartzoulakis, K. Effect of irrigation on quality attributes of olive oil. *J. Agric. Food Chem.* **2009**, *57*, 7048−7055.

(19) Mandia, S.; Bauer, B. *International Olive Oil Council*; http://www.internationaloliveoil.org/.

(20) Liang, N. N.; Zhang, L. X.; Wang, X. L.; Zhang, L. X. Identification of fatty acids in vegetable oils by mass spectrometry and equivalent chain length. *Chinese J. Anal. Chem.* **2011**, *39*, 1166−1170.

(21) Yi, L.; Yuan, D.; Che, Z. Plasma fatty acid metabolic profile coupled with uncorrelated linear discriminant analysis to diagnose and biomarker screening of type 2 diabetes and type 2 diabetic coronary heart diseases. *Metabolomics* **2008**, *4*, 30−38.

(22) Li, H.; Liang, Y.; Xu, Q. Model population analysis for variable selection. *J. Chemom.* **2010**, *24*, 418−423.

(23) Li, H.; Liang, Y.; Xu, Q. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal. Chim. Acta* **2009**, *648*, 77−84.

(24) Cao, D.; Wang, B.; Zeng, M. A new strategy of exploring metabolomics data using a Monte Carlo tree. *Analyst* **2011**, *136*, 947−954.

(25) Velasco, J.; Dobarganes, C. Oxidative stability of virgin olive oil. *Eur. J. Lipid. Technol.* **2002**, *104*, 661−676.

(26) Cruz, M. A. M.; Brocca, M. M. B.; Leyba, C. O. Olive oil in clinical nutrition. *Grasas Aceites* **2004**, *55*, 76−83.

(27) Frankel, E. N. Nutritional and biological properties of extra virgin olive oil. *J. Agric. Food Chem.* **2011**, *59*, 785−792.

(28) Swinburn, B. A.; Boyce, V. L.; Bergman, R. N. Deterioration in carbohydrate metabolism and lipoprotein changes induced by modern, high fat diet in Pima Indians and Caucasians. *J. Clin. Endocrinol. Metab.* **1991**, *73*, 156−165.

(29) Hayne, R. J.; Mulder, C. M.; Popp-Snijders, C. P. Linoleic-acid-enriched diet: long-term effects on serum lipoprotein and apolipoprotein concentrations and insulin sensitivity in non-insulin dependent diabetic patients. *Am. J. Clin. Nutr.* **1989**, *49*, 448−456.

(30) Montealegre, C.; Alegre, M. L. M.; Ruiz, C. G. Traceability markers to the botanical origin in olive oils. *J. Agric. Food Chem.* **2010**, *58*, 28−38.

(31) Gandhi, S. D.; Kishore, V. K.; Crane, J. M. Selection for low erucic acid and genetic mapping of loci affecting the accumulation of very long-chain fatty acids in meadowfoam seed storage lipids. *Genome* **2009**, *52*, 547−556.

(32) Borg, K. Physiopathological effects of rapeseed oil: a review. *Acta Med. Scand. Suppl.* **1975**, *585*, 5−13.

(33) Charlton, K. M.; Corner, A. H.; Davey, K. Cardiac lesions in rats fed rapeseed oils. *Can. J. Comp. Med.* **1975**, *39*, 261−269.

(34) Christophersen, B. O.; Norseth, J.; Thomassen, M. S. Metabolism and metabolic effects of $C_{22:1}$ fatty acids with special reference to cardiac lipidosis. In *Nutritional Evaluation of Long Chain Fatty Acids in Fish Oil*; Barlow, S. M., Stansby, M. E., Eds.; Academic Press: New York, 1982; pp 89−139.