New Contribution to the Chemometric Characterization of Almond Cultivars on the Basis of Their Fatty Acid Profiles

M. L. Martín Carratalá,* C. García-López, V. Berenguer-Navarro, and N. Grané-Teruel

Department of Analytical Chemistry, University of Alicante, P.O. Box 99, 03080 Alicante, Spain

Eight minor fatty acids (C10:0, C12:0, C14:0, C15:0, C17:0, C17:1, C18:3, and C20:0) have been determined in the kernel of 19 almond cultivars grown in the same field and year, using gas chromatography. Submission of minor fatty acids along with the major ones (C16:0, C16:1, C18:0, C18:1, and C18:2) to multivariate techniques (principal component analysis, cluster analysis, discriminant analysis) contributes to characterizing this set. In comparison to previous studies, the minor fatty acids enable further subdivision within the group of the Mediterranean almonds. The American cultivar Titan, though classified near the cultivars Wavona, Genco and Tuono, presents the fatty acid profile most similar to the Spanish cultivars. The Spanish cultivar Marcona, which is classified with the Spanish Ramillete, Cebas, Desmayo Largueta, and Del Cid and the American Non Pareil, shows some particular features. Discriminant analysis was applied to the four groups previously established by cluster analysis.

Keywords: *Chemometrics; almond; fatty acids; characterization; gas chromatography; multivariate techniques*

INTRODUCTION

Component profiles of foods and beverages have often provided useful information for their characterization, and, in this way, they help to prevent adulterations and fraudulent practices. Effective statistical methods are able to reveal features that allow the establishment of associations among samples of a common origin or of a similar quality. In particular, multivariate analysis has been applied to differentiate Chinese tea (Lin et al., 1987), wild oysters (Favreto et al., 1991), spirits (Tetsuo, 1991), apricot purées (Parolari et al., 1992), Italian mandarin (Campisi et al., 1995), mushrooms (White et al., 1993), and wines (Ortiz et al., 1996).

Fat is the main fraction in almond (50−65%, w/w, referred to the dried weight of the seed). A few fatty acids (palmitic, palmitoleic, stearic, oleic, and linoleic) account for ∼99.5% of fat content in almonds. In a previous paper (Garcia et al., 1996), we reported the major fatty acid profiles of a set of 19 almond cultivars, and it was proved that this fraction can be used to associate them in groups, pointing out relationships among them. This conclusion could be enriched by the study of the rest of the fatty acids (minor fatty acids), which include the acids C10:0, C12:0, C14:0, C15:0, C17:0, C17:1, C18:3, and C20:0. Undoubtedly, these acids are not important from the point of view of composition, but their significance for discriminating purposes might well be considered.

Information on minor fatty acid content in European almonds is very scarce and often useless for our purpose. Some of these data refer to some very local or undefined cultivars, and others are imprecise, due to the limitation of the techniques employed to measure them. No reference is found on statistical multivariate evaluation of data of minor fatty acids.

In this paper we determine the minor fatty acid composition of the same set of almonds previously studied (García et al., 1996) and submit to multivariate techniques the whole information available on the fatty acid fraction of the set.

EXPERIMENTAL PROCEDURES

Samples. The set studied consists of 19 almond cultivars: 7 Spanish [Malagueña (MA), Peraleja (PE), Atocha (AT), Del Cid (DC), Desmayo-Largueta (DL), Ramillete (RA), Marcona (MR)], 3 Italian [Genco (GE), Tuono (TU), Cristomorto (CR)], 1 Australian [Chellaston (CH)], 4 American [Texas (TE), Non Pareil (NP), Titan (TI), Wawona (WA)], 1 Tunician [Achaak (AC)], 1 from a Caucasian region [Primorskyi (PR)], 1 French [Ferragnes (FE)], and a hybrid, CE, obtained at CEBAS Center (Centro de Edafología y Biología Aplicada del Segura, Murcia, Spain). For all of the cultivars, three independent previously blended samples were taken from different trees. All of the almond trees are cultivated in the same conditions in the CEBAS experimental station, and in this way additional sources of variance are avoided. The Spanish cultivars come from a vegetal material accepted as the most authentic representative of these cultivars some years ago by a Spanish technical committee (Felipe et al., 1984).

Extraction and Gas Chromatography. Fat extraction and transesterification to methyl esters of the almond kernels were carried out according to the method described in our previous paper (AOCS, 1969). Chromatographic data were obtained with a Carlo Erba series 8000 gas chromatograph equipped with a split/splitless injector and a flame ionization detector, interfaced to a computer provided with the analytical Chrom-Card program for data acquisition and processing. The chromatographic column was a capillary DB23 (J&W Scientific), 30 m × 0.32 mm i.d.

Experimental conditions for determining minor fatty acid methyl esters were conveniently adapted to their low level content, and the thermal program was selected so that the retention times of methyl esters were ∼10 times the dead time. The time selected for valve closure (45 s) ensured a correct area measurement of peaks, a reduced purge time, and a high sampling rate.

---

* Author to whom correspondence should be addressed.

964 *J. Agric. Food Chem.,* Vol. 46, No. 3, 1998

Martín Carratá et al.

Table 1. Calibration Data for the Standard Minor Fatty Acid Methyl Ester Mixture

fatty acid	sensitivity (mV/ppm)	detection limit (g/100 g of oil)
C10:0	0.958	0.002
C12:0	1.002	0.003
C14:0	1.059	0.007
C15:0	1.090	0.002
C17:0	1.107	0.007
C17:1	1.119	0.005
C18:3	1.148	0.005
C20:0	1.177	0.015

The optimal temperatures found for injector and detector were 275 and 300 °C, respectively. Oven temperature was kept at 50 °C for 2 min, increased from 50 to 180 °C at 10 °C min^{-1}, held at 180 °C for 5 min, increased from 180 to 240 °C at 5 °C min^{-1}, and finally maintained at 240 °C for 2 min. Helium at 2 mL min^{-1} flow rate was used as carrier gas. Samples of 1 µL were injected.

Standard Materials. Standard fatty acid methyl esters were acquired from Sigma and chromatographically run as the samples. Fatty acid methyl ester C9:0, which is absent in almonds, was used as the internal standard.

Statistical Analysis and Data Processing. Several statistical methods in the SPSS statistical package (SPSS, 1994) were used for the analysis of data. First, one-way ANOVA (using the Tukey-B procedure) was applied to test the variables (fatty acid methyl esters) that could contribute most to differentiating almond cultivars. The significance level ($p = 0.05$) was estimated by comparing F_{obsd} with F_{cited} for ($g - 1$) and ($N - g$) degrees of freedom (where $g = 19$ is the number of groups and $N = 57$ is the total number of samples). Starting from this information, we sequentially applied principal component, cluster, and linear discriminant analyses.

The principal component analysis (PCA) was applied to autoscaled correlations, using different criteria: scree test (Cattel, 1966), mean eigenvalues (Cela, 1994), and indicator function (Malinowski, 1990). The cluster analysis was carried out by applying the average linkage method for agglomeration and the square of Euclidean distance as the criterion of proximity (Afifi and Clark, 1990). The linear discriminant analysis (LDA) was conducted stepwise by employing the Wilks' lambda statistic (Tabachnick and Fidell, 1992) for variable selection.

In all cases, the algorithms used were applied to the individualized data obtained for each sample, but the representative points of cultivars shown in the figures of this paper are average values for the sake of the clarity.

RESULTS AND DISCUSSION

Chromatographic peaks are well resolved, and calibration curves of peak areas relative to internal standard against the methyl ester content are linear ($r > 0.996$) for each fatty acid in the range of work. Table 1 shows some important calibration parameters.

Reproducibilities of analytical results within a day and between days for every minor fatty acid methyl ester are similar for different cultivars. Representative values obtained from six analyses of a sample of Marcona cultivar within a day or once a week appear in Table 2. Relative standard deviations (RSD) for minor fatty acid methyl esters are logically higher than those for major ones, because of the lower contents of the former, but, in general, they can be considered moderate.

Mean values from three totally independent determinations of minor fatty acid methyl esters for every cultivar studied are given in Table 3. The data on major fatty acid methyl esters used in the statistical processing are taken from our previous paper (Garcia et al., 1996).

Table 2. Reproducibility Data for Minor Fatty Acid Methyl Esters of the Marcona Almond Cultivar

fatty acid	reproducibility in the same day (g/100 g of oil) mean ($n = 6$)	RSD (%)	reproducibility week to week (g/100 g of oil) mean ($n = 6$)	RSD (%)
C10:0	0.067	0.6	0.070	4.7
C12:0	0.041	2.7	0.041	4.8
C14:0	0.052	5.8	0.052	8.4
C15:0	0.011	7.7	0.011	10.8
C17:0	0.042	8.7	0.041	12.0
C17:1	0.060	8.7	0.059	11.7
C18:3	0.009	10.7	0.009	13.5
C20:0	0.040	9.4	0.040	12.9
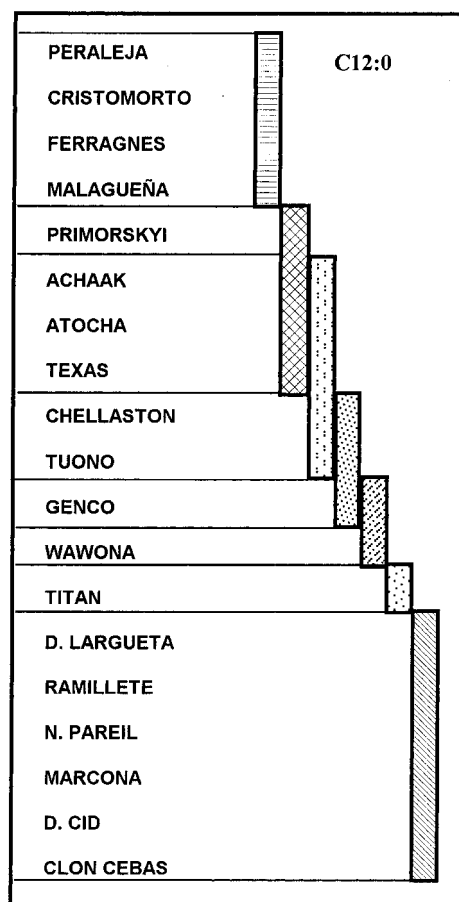
Figure 1. Graphical representation of cultivars associated on the basis of a Tukey multiple comparison of mean values for fatty acid C12:0. Cultivars connected by a common bar are not significantly different at $p = 0.05$.

It is apparent that five acids (palmitic, palmitoleic, stearic, oleic, and linoleic) account for >99% of total esterified fat; their contribution considerably decreases in the following order: C18:1 > C18:2 > C16:0 > C18:0 > C16:1. The relative importance of the rest of the fatty acids is C17:1 (0.06%) > C10:0 > C14:0 ≫ C20:0 ≫ C17:0 > C12:0 > C18:3 ≫ C15:0 (<0.01%).

By means of ANOVA analysis, F ratio values for every minor fatty acid methyl ester (or simply for every fatty acid, as stated in the rest of the paper) were obtained (Table 4). Since all values, including those for minor acids, are >1 at a significance level of 0.05, it follows that all of the fatty acids can contribute to a greater or lesser extent to the discrimination of the almond cultivars considered.

Figure 1 shows the associations arising when the cultivars are compared on the basis of their mean

Chemometric Characterization of Almond Cultivars

J. Agric. Food Chem., Vol. 46, No. 3, 1998 **965**

Table 3. Mean Concentration of Minor Fatty Acid Methyl Esters

sample	fatty acid methyl ester (g/100 g of oil)							
	C10:0	C12:0	C14:0	C15:0	C17:0	C17:1	C18:3	C20:0
Genco	0.058	0.24	0.027	0.007	0.023	0.041	0.005	0.016
Marcona	0.068	0.043	0.58	0.012	0.45	0.069	0.010	0.039
Del Cid	0.071	0.043	0.050	0.010	0.044	0.064	0.006	0.044
D. Largueta	0.074	0.040	0.048	0.010	0.035	0.067	0.011	0.028
Texas	0.040	0.019	0.043	0.011	0.050	0.071	0.018	0.042
C. CEBAS	0.069	0.044	0.043	0.009	0.034	0.062	0.011	0.037
Peraleja	0.016	0.007	0.039	0.010	0.056	0.085	0.016	0.077
Atocha	0.048	0.017	0.024	0.006	0.024	0.041	0.005	0.025
Achaak	0.038	0.017	0.034	0.009	0.034	0.055	0.008	0.024
Wawona	0.058	0.026	0.040	0.010	0.034	0.053	0.009	0.028
Malagueña	0.017	0.007	0.036	0.012	0.052	0.067	0.017	0.052
Non Pareil	0.072	0.043	0.055	0.011	0.048	0.060	0.017	0.065
Feragnes	0.019	0.007	0.028	0.008	0.043	0.080	0.017	0.045
Cristomorto	0.016	0.007	0.034	0.010	0.042	0.098	0.024	0.063
Tuono	0.050	0.020	0.029	0.008	0.031	0.042	0.007	0.034
Chellaston	0.051	0.020	0.037	0.008	0.039	0.050	0.006	0.026
Primorskyi	0.038	0.016	0.030	0.009	0.032	0.058	0.009	0.024
Ramillete	0.072	0.042	0.047	0.009	0.039	0.077	0.010	0.039
Titan	0.065	0.032	0.057	0.012	0.046	0.084	0.014	0.043

Table 4. F_{ratio} Values from ANOVA Analysis for the Minor Fatty Acid Methyl Esters

fatty acid	F_{ratio}	fatty acid	F_{ratio}
C10:0	345.9	C17:0	7.3
C12:0	373.8	C17:1	4.8
C14:0	36.9	C18:3	22.6
C15:0	8.0	C20:0	9.4

content in fatty acid C12:0. It is apparent that this fatty acid, despite of its low content in almond, suggests seven associations of cultivars. The almonds within groups that partially overlap, such as those of Primorskyi and Chellaston, do not differ from the point of view of C12:0. Similar graphics can be obtained for other fatty acids, suggesting a basis for corresponding associations.

From the co-variance matrix of the means of all the variables, the PCA demostrates that the first four principal components retained explain 86.6% of the total variance. A varimax rotation reveals that the first two principal components are strongly associated with the major fatty acids, whereas the other two are associated with the minor ones.

The analysis of variance on the scores for a cultivar in the space determined by the four rotated principal components shows that the principal component associated with the minor fatty acids C15:0, C17:0, C17:1, C18:3, and C20:0 has a small *F*-ratio value compared it with the other three principal components. For this reason, this group of fatty acids has been omitted with a view to obtaining a more simplified visualization of the mean values in the reduced space of the first two principal components.

By reapplying the PCA to this new set formed by the major fatty acids along with the minor C10:0, C12:0, and C14:0 and retaining the first three rotated principal components, 84.7% of the total variance is accounted for. Table 5 shows the communality values and the partial variance explained by the principal components retained. Figure 2 shows the projections of different almond cultivars on the reduced space determined by the first two rotated principal components, where the four gross associations suggested by the cluster analysis are outlined.

The cluster analysis is an effective means of establishing associations on the basis of nearness criteria between objects. As a result of the application of this algorithm, the dendrogram, from the mean values, of

Table 5. Communality Values for Each of the Fatty Acids Studied and the Individual Variance Accounted for by the Principal Components Retained

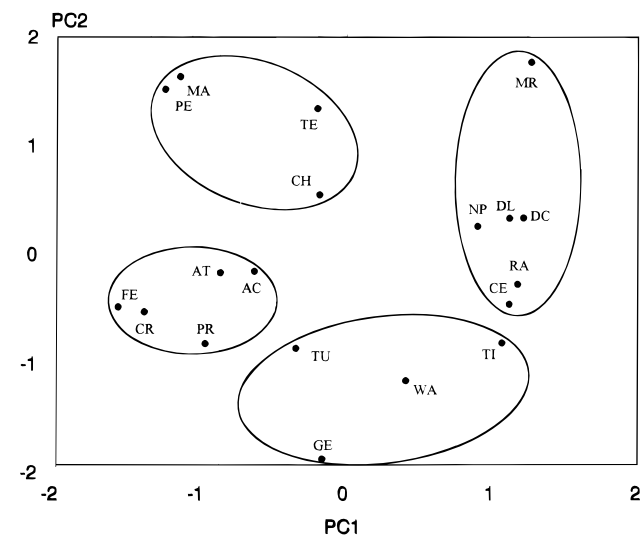| fatty acid	communality	PC	eigenvalue	% var	cum var
C10:0	0.91974	1	3.25660	40.7	40.7
C12:0	0.96353	2	2.12176	26.5	67.2
C14:0	0.80447	3	1.69858	17.5	84.7
C16:0	0.90217				
C16:1	0.69642				
C18:0	0.83385				
C18:1	0.87499				
C18:2	0.77899				

Figure 2. Projections of the average scores of almond cultivars on the reduced space of the two rotated principal components. Groupings are derived from a separate cluster analysis.

Figure 3 arises. The first association consists of Genco, Tuono, Wawona, and Titan cultivars, and the second of the Spanish cultivars Peraleja and Malagueña, together with the American Texas and Chellaston. The other two groups arise from the breaking down of the large association, which included the majority of Mediterranean almonds in a previous paper, in which only the major fatty acids were processed (Garcia et al., 1996). The inclusion of data on minor fatty acids provides a further basis to state similarities and dissimilarities among strongly related cultivars. The third group
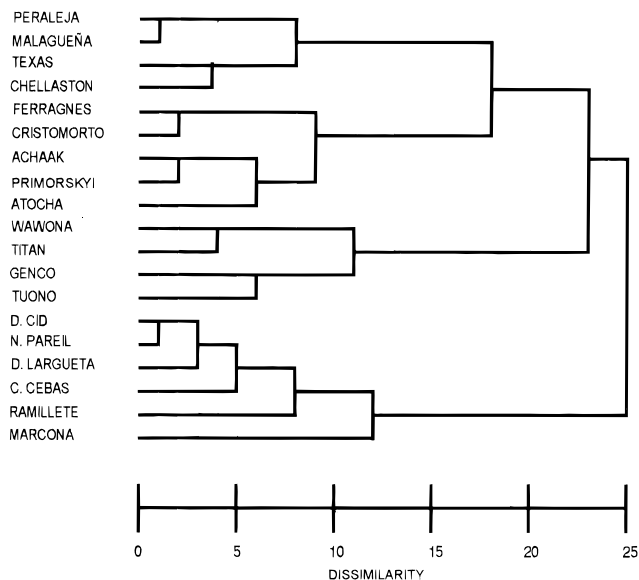
**Figure 3.** Dendrogram from an average linkage cluster analysis.

Table 6. Rotated Correlation between Discriminating Variables and Canonical Discriminant Functions

variable	DF1	DF2	DF3
C18:1	0.73138*[a]	−0.11123	−0.15315
C16:1	0.46695*	0.24092	0.16287
C18:0	−0.11164	0.69128*	−0.00468
C18:2	0.10548	0.65952*	0.01688
C16:0	0.32903	0.46075*	0.06009
C12:0	0.09889	−0.00374	0.71175*
C10:0	−0.01065	−0.11450	0.68781*
C14:0	−0.08507	0.24405	0.53968*

[a] Asterisks denote largest absolute correlation between each variable and discriminant functions.

consists of the Ferragnes, Cristomorto, Achaak, and Primorskyi cultivars, to which the Spanish Atocha is associated. In the fourth group, we include very typical Spanish cultivars, cultivated in rather large areas of Spain. It is not surprising that the American cultivar Non Pareil also belongs to this last group, associated with Spanish cultivars, as already observed in an early study on amino acids (Prats and Berenguer, 1994). The dendrogram shows that some cultivars, such as Marcona and Titan, continue showing singular features within the groups in which they are included.

The descriptive discriminant analysis is an algorithm in which mathematical functions (linear combinations of original variables) are calculated, so that these functions maximize the differences between the established groups. In this way and through a convenient varimax rotation, three discriminant functions that explain 34.65, 33.77, and 31.58% of the total variance, respectively, are deduced. These functions assign every cultivar to a single group at a 98.25% level of success. Rotated correlations among discriminating variables and discriminant functions are shown in Table 6. There it can be seen that the first function is strongly correlated with the fatty acids C16:1 and C18:1; the second one, with the fatty acids C16:0, C18:0, and C18:2; and finally, the third one with C10:0, C12:0, and C14:0. The classification power of the different fatty acids can be estimated by the calculation of classification functions and from the variation of the percentage of correct classification on the basis of the inclusion or exclusion of the fatty acid in the computation. Considering only
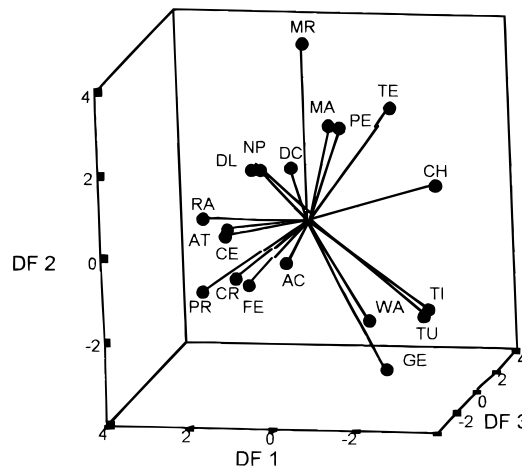
the fatty acids C:10 and C:12, a correct classification of >84% is obtained. By including, however, the fatty acids C10:0, C12:0, and C18:0, correct classification increases to 95%, and when the rest of the fatty acids are added, the classification can be said to be 100% correct.

Figure 4 shows the projections of all the samples in the reduced space determined by the first three rotated discriminant functions. The cultivars are associated in a way similar to that we obtained above but from different algorithms, affording a confirmation of their validity. The distance between Marcona and the rest of the cultivars in the assigned groups is again remarkable.

Analysis of variance on the score values for all the samples in relation to the three discriminant functions shows that the second rotated discriminant function is the only one that establishes significant differences among all groups of almond cultivars, when these are compared two by two. The first rotated discriminant function, however, yields the least number of differences between the groups at the level $p = 0.05$.

The consistency of results from the different approaches used is worth noting. As far as the minor fatty acids are concerned, the undoubtful significance of including at least some of them (C10:0 and C12:0) in the data processing for classification purposes must be noted. Turning back to Figure 1, we can observe that according to ANOVA and on the basis of the minor acid C12:0, the Spanish cultivars appeared associated with the American Non Pareil and clearly separated from the others studied here.

The significance of analytical results so far discussed is, of course, open to discussion and demands further research. Nevertheless, on the basis of the consistent associations observed in this study (e.g. among some extended cultivars in Spain, on one hand, or among Titan, Wawona, Genco, and Tuono, on the other hand) the existence of convenient patterns for classification seems more than a good heuristic hypothesis.

In fact, we have found a proper discriminant function for the group of all the Spanish cultivars studied (Atocha, Peraleja, Malagueña, Marcona, Desmayo Largueta, Del Cid, Ramillete, Cebas) using only data from fatty acids C16:0, C16:1, C18:1, and C18:3. This function permits a classification with 100% certainty for the mentioned cultivars, distinguishing them from the others studied. The structure matrix and the standard-

Figure 4. Mean scores of almond cultivars projected on the reduced space of the three rotated discriminant functions.

Chemometric Characterization of Almond Cultivars

J. Agric. Food Chem., Vol. 46, No. 3, 1998 **967**

Table 7. Correlation between Discriminating Variables and Discriminant Function and Standardized Discriminant Function Coefficients for Each Fatty Acid

fatty acid	correlations function 1	standardized coeff DF
C16:0	0.7544	0.8730
C16:1	0.5819	0.6457
C18:1	0.1042	−0.6375
C18:3	−0.0855	−0.3762

ized coefficients for the discriminant function are shown in Table 7. If the fatty acid C18:3 is left out, the degree of certainty decreases slightly to 96.5%, whereby, however, two of the three samples of Cebas cultivar are misclassified.

Given the above-mentioned singularity of the Spanish cultivar Marcona, a discriminant analysis for differentiating this from the rest of the studied cultivars was tested. In fact, a discriminant function can be calculated, successfully differentiating this cultivar at 100%. This function is strongly associated with the fatty acid C16:1 and, to a lesser extent, with the fatty acids C15:0, C17:1, and C18:1, which could probably be considered characteristic chemical parameters for this cultivar.

A similar test has been done for the American cultivar Titan, which, after PCA, appears somehow far from Genco, Tuono, and Wawona, and after cluster analysis, it is associated with the group of Spanish cultivars or with the American ones, depending on the fatty acids taken into account. A discriminant function can be found to differentiate this cultivar from the rest, with 100% success, this function is strongly associated with the fatty acids C14:0, C15:0, C16:1, and C18:0 and, to a lesser extent, with the fatty acid C12:0. If the fatty acid C15:0 is dropped from the analysis, the discrimination is only 98.25% successful.

CONCLUSION

The value of statistical multivariate techniques to establish relationships among almond cultivars on the basis of their fatty acid profiles has been confirmed. Inclusion of some minor fatty acids in this approach provides a further subdivision of groups earlier claimed on the basis of major fatty acid composition. Such analysis reveals a rather strong association among Tuono, Genco, Titan, and Wawona on one hand, and among typical Spanish cultivars such as Del Cid, Desmayo Largueta, Ramillete, Cebas, and Marcona, on the other. Convenient discriminant functions can be deduced, by means of which a successful differentiation at 100% of all the Spanish cultivars studied is achieved. For cultivars Marcona and Titan, discriminant functions can also be obtained, suitable to differentiating them from the rest of the cultivars.

LITERATURE CITED

Afifi, A. A.; Clark, V. *Computer-Aided Multivariate Analysis*, 2nd ed.; New York, 1990; pp 429−462.

AOCS. In *Official and Tentative Methods of the American Oil Chemists' Society*; American Oil Chemists' Society: Chicago, IL, 1969; Method Ce-2-66.

Campisi, B.; Dugo, G.; Crotoneo, A.; Favreto, L. Chemometric analysis and extraction processes of mandarin essential oils. *Anal. Chim. Acta* **1995**, *312*, 199−205.

Cattel, R. B. The scree test for the number of factors. *Mult. Behav. Res.* **1966**, *1*, 245−276.

Cela, R. *Avances en Quimiometría Práctica*; Servicio de Publicaciones e Intercambio Científico: Santiago de Compostela, Spain, 1994; pp 241−353.

Favretto, L.; Gabrielli, L.; Ceccon, L.; Vojnovic, D. Differentiation of wild oysters from three sites by chemometric procedures. *Anal. Chim. Acta* **1991**, *248*, 51−58.

Felipe, A.; Vargas, F. J.; García, J. E. *Variedades Tipificadas de Almendra en España*; Fundación Caja de Pensiones: Barcelona, 1984.

García-López, C.; Grané-Teruel, N.; Berenguer-Navarro, V.; García-García, E.; Martín-Carratalá, M. L. Major fatty acid composition of 19 almond cultivars of different origins. A chemometric approach. *J. Agric. Food Chem.* **1996**, *44* (7), 1751−1755.

Lin, X.; Van Espend, P.; Adams, F.; Yan, S.; Vanbelle, M. Classification of Chinese tea samples according to origin and quality by principal component techniques. *Anal. Chim. Acta* **1987**, *200*, 421−430.

Malinowski, E. R. *Factor Analysis in Chemistry*, 2nd ed.; Wiley: New York, 1990.

Ortiz, M. Cruz; Sarabia, Luis A.; Symington, Charles; Santamaría, Fernando; Iñíguez, Monserrat. Analysis of aging and typification of vintage Ports by partial least squares (PLS) and soft independent modelling class analogy (SIMCA). *Analyst* **1996**, *121* (8), 1009−1013.

Parolari, G.; Virgili, R.; Bolzoni, L. Analysis of sensory and instrumental data on apricot purées with pattern recognition techniques. *Anal. Chim. Acta* **1992**, *259*, 257−265.

Prats, M. S.; Berenguer, V. Caracterización de algunas variedades de almendra por su composición en aminoácidos libres. *Rev. Esp. Cienc. Tecnol. Aliment.* **1994**, *34*, 218−227.

SPSS. Release 6.0.1; SPSS Inc., Chicago, IL, 1994.

Tabachnick, B.; Fidell, L. *Using Multivariate Statistics*, 12th ed.; Harper Collins Publishers: New York, 1992; pp 505−596.

Tetsuo, A. Discrimination of liquor aromas by pattern recognition analysis of responses from a gas sensor array. *Anal. Chim. Acta* **1991**, *243*, 293−300.

White, R.; Wentzell, P.; Beasy, M. Taxonomy of Amanita mushrooms by pattern recognition of amino acid chromatographic data. *Anal. Chim. Acta* **1993**, *277*, 333−346.