Inductively Coupled Plasma Application for the Classification of 19 Almond Cultivars Using Inorganic Element Composition

Soledad Prats-Moya, Nuria Grané-Teruel, Vicente Berenguer-Navarro, and M. Luisa Martín-Carratalá*

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

Ten inorganic elements (Ca, Cu, Fe, S, K, P, Na, Mn, Mg, and Zn) have been determined in the kernel of 19 almond cultivars grown in the same field and year, but from different origins, by means of inductively coupled plasma (ICP) after a microwave digestion of the samples. Statistical methods have been applied to a batch of 171 samples to find relations among the cultivars. The *F* ratios obtained by applying ANOVA analysis are all significant, indicating that a classification of the cultivars using these variables is reasonable to pursue. The cluster analysis and the principal component analysis can distinguish the American cultivars Non Pareil, Wawona, and Texas from the remainder of almonds, most of them from the Mediterranean area, while the Titan cultivar, also an American one, shows a great similarity to Atocha, a Spanish almond. Further, linear discriminant analysis is used to identify the inorganic elements that contribute to a higher degree to find differences between the hypothesized groups of almonds obtained from a principal component analysis and cluster analysis.

Keywords: *Inductively coupled plasma; microwave digestion; multivariate analysis; almonds*

INTRODUCTION

Inductively coupled plasma (ICP) is a powerful technique for multielemental analysis. Multivariate analysis can take advantage of ICP's plentiful data for characterization of natural goods. The coupling of both techniques has been successfully applied to foods.

Almond cultivation occupies an important place in Spanish agriculture because mild weather conditions are favorable for its growth. Nevertheless, in recent years, the Spanish production of almonds has decreased considerably, while in the United States production has increased quickly due to the large amount of money invested for improving agricultural techniques and selecting new almond cultivars. This situation has led to a competition between American almonds and some Spanish ones, which are more expensive because of their appreciated sensory properties.

Quite a few studies have been made with the aim of finding composition parameters that could establish similarities between almonds with a common origin or similar quality by using multivariate statistical methods. Previous research has reported characterization of a set of different almond cultivars from diverse origins on the basis of their free amino acid (Prats-Moya and Berenguer-Navarro, 1994) and their complete fatty acid profiles (García-López et al., 1996), which allowed the assignment of cultivars to classes based on their similarity.

Macroelements and oligoelements in almonds have been determined using atomic spectroscopy after an acid digestion (Gómez-Martínez, 1976; Saura-Calixto et al., 1988), the results being that K, P, Mg, S, and Ca are present in high amounts in the kernel of the almonds, while other elements such as Na, Fe, Cu, and Mn are found only in minor or trace amounts. These studies conclude that the inorganic fraction varies not only between cultivars but also between the nuts of the same cultivar grown in different locations. The variations in the inorganic element composition are as important as the variations in protein and lipid content (Graselly and Grossa-Raynaud, 1980) among different cultivars. Gómez-Martínez (1976) studied the inorganic composition of almond leaves during the ripening process for the same almond cultivar in different locations. It appeared that N and P composition depends on the ripening stage but not on the location, while Ca, K, S, and Mg also depend on the location; so, this author has claimed that the type of soil and weather (i.e. location) exert a major influence on the inorganic composition than the cultivar.

In this work we have undertaken the study of the inorganic almond fraction in the kernel of 19 different almonds to evaluate more exactly the influence of the cultivar on the mineral composition. All of the almonds studied in this work were harvested in the same field and year and, therefore, were produced under the same conditions of climate, fertilization, and soil to reduce additional sources of variance.

EXPERIMENTAL PROCEDURES

Instrumentation. Digestions were carried out in a Milestone microwave oven (Model MLS-1200, MEGA) equipped with six 50 mL Milestone pressure decomposition Teflon vessels, a microwave digestion rotor (Model HPR 100016), and a remote control panel for the decomposition of the samples.

Analysis of all samples was performed using ICP atomic emission spectrometry (ICP-AES) (Perkin-Elmer, Optima 3000; Norwalk, CT), equipped with an autosampler and assisted by an IBM computer for data acquisition and readout system. The instrumental operating parameters are shown in Table 1.

Samples. The set of samples studied consists of 19 almond cultivars: 7 Spanish [Malagueña (MA), Peraleja (PE), Atocha (AT), Del Cid (DC), Desmayo-Largueta (DL), Ramillete (RA), and 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 Tunisian [Achaak (AC)], 1 from a Caucasian region [Primorskyi (PR)], 1 French [Ferragnes (FE)], and a hybrid, Clon Cebas (CE), obtained at Centro de Edafología y Biología

* Author to whom correspondence should be addressed [fax (+34) 65903464; e-mail maria.prats@ua.es].

Table 1. ICP-AES Instrumental Operating Condition

inductively coupled plasma	Perkin-Elmer, Optima 3000
operating power, W	1100
radiofrequency, MHz	40
coolant stream flow, L/min	15
torch type	quartz torch with 2.0 mm i.d. injector tube
nebulizer type	crossflow
nebulizer flow, mL/min	1.2
auxiliary flow, mL/min	1.0
flow rate, mL/min	1.0

Table 2. Milestone Microwave Oven Operating Program

program, stage	power, W	time, min
1	250	1
2	0	1
3	250	5
4	400	5
5	600	5
6	cooling	5

Aplicada del Segura, Murcia, Spain (CEBAS). The trees corresponding to the above-mentioned Spanish cultivars (generally, very diversified) were selected and accepted as the most authentic representatives of these cultivars some years ago by a technical committee (Felipe et al., 1984). Such selected cultivars have been replicated in other experimental research fields in Spain. All of the samples for the present study are a mixture of almonds taken from different trees grown under the same conditions in CEBAS experimental fields. All of the almonds belong to the 1993 harvest and since then have been carefully stored in a refrigerator, after grinding, drying, and defatting.

Sample Preparation. A general procedure was followed to digest the samples using the microwave oven. Half gram portions of ground, dried (1 h at 60 °C), and defatted almonds were weighed into 100 mL Teflon vessels, and thereafter 4 mL of concentrated nitric acid and 1 mL of hydrogen peroxide were added. The heating program (Table 2) was chosen to obtain clear solutions, to prevent blow-out and overheating of the vessels, and to reduce the digestion time to a minimum. The digested samples were allowed to cool in the vessels, and the resultant solutions were transferred into a calibrated flask (100 mL) and diluted to volume with ultrapure water. Samples were filtered before being analyzed by ICP-AES.

Statistical Data Processing. All of the statistical analyses were performed using SPSS release 6.0 Professional Statistics.

First, one-way ANOVA, according to the Tukey-B procedure (SPSS manual), was applied to explore the variables (inorganic elements) that could contribute most to find differences among the almond cultivars. The statistical significance ($p = 0.05$) was estimated by comparing *F*-observed with *F*-cited for ($g - 1$) and ($N - g$) degrees of freedom (where g is the number of groups and N the number of samples).

Afterward, principal component analysis (PCA) and cluster analysis were shown useful for finding associations among the almond cultivars studied taking into account the 10 elements determined. PCA was applied to data previously autoscaled, using the eigenvalues >1, i.e. the Kaiser criteria and the Scree test (Cattell, 1966). Cluster analysis was carried out by using the average linkage method for agglomeration and the square of Euclidean distance as criteria of proximity (Afifi and Clark, 1990).

Finally, a stepwise linear discriminant analysis (LDA) was used employing the Wilks' Lambda values (Tabachnick and Fidell, 1992) for the selection of features (elements) that are necessary and sufficient to obtain the hypothesized associations. The algorithms used have been applied to the individualized data obtained for each sample, but for the sake of clarity only the average values for each cultivar have been shown in the figures.

RESULTS AND DISCUSSION

Calibration curves are drawn with a blank and five mixed standard solutions and have very good correlation coefficients in all cases. The inorganic composition and the standard deviations (SD) for all of the samples are given in Table 3. The data are the mean of nine values, obtained from three different samples analyzed in triplicate.

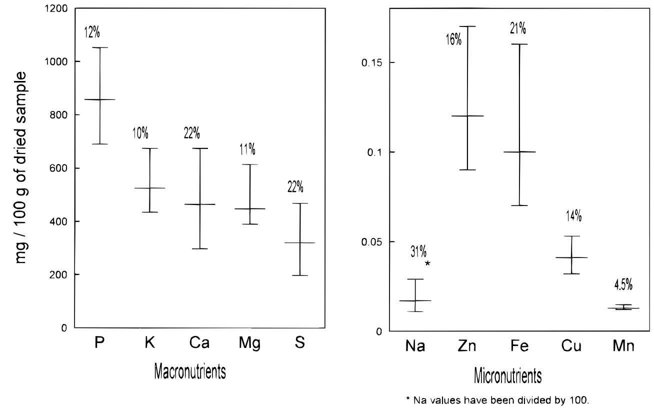

Figure 1. Intervals of inorganic content (mean values for each element). The crossed line shows the average of the mean values. Percentages on top of intervals are the corresponding relative standard deviations (% RSD).

P is the most abundant element in the almonds studied (Table 3), followed by K, Ca, Mg, and S, all of them macronutrients, while Na, Zn, Fe, Cu, and Mn are found in minor amounts. Figure 1 shows the intervals of mean values and their average, along with an indication of the relative standard deviation (% RSD) for each element. Na presents the higher variability among the cultivars (31% RSD), followed by S, Ca, and Fe (with 22%, 22% and 21% RSD, respectively), whereas Mn shows the lowest % RSD (4.5%).

Similar, but more precise, information is provided by one-way ANOVA using the Tukey test, which allows all pairwise comparisons of means. The *F* ratio values for each variable (Table 4) are all >1 at a significance level of 0.05, indicating that all of the inorganic elements show differences among the almond cultivars. It is likely that elements with the highest *F* ratio values, such as Ca (macroelement), Na, and Fe (oligoelements), could be used to mark differences among cultivars, when all variables are compared in pairs, independently of the location, since all of the samples used in this study were cultivated in the same field. The *F* ratio values for K, P, and Mg in the cultivars studied are very similar, maybe suggesting some common processes or species in which they take part. This suggestion is supported by correlation analysis on these three elements, whereby positive correlations between P and Mg of 0.70 and between P and K of 0.49 are found.

Table 4 also shows the percentage of significant differences when the cultivars are compared in pairs on the basis of their mean composition. This percentage can be given as an index of differentiating power for each element.

The meaning of the ANOVA analysis becomes apparent by grouping the cultivars with statistically comparable mean values for each element. Figure 2 displays in bars the associations that are likely to arise when Ca or Cu is considered. Cultivars belonging to the same bar significantly differ from cultivars of other bars if they are not overlapped, not even partially. So, Ca mean values are different enough to bring about 14 classes with only one or two cultivars each one. It can be concluded that Ca is a very suitable parameter for differentiating cultivars. On the contrary, Cu only

Table 3. Inorganic Content of Different Almond Cultivars[a]

sample		Ca	Cu (10^{-4})	Fe (10^{-3})	K	Mg	Mn (10^{-4})	Na	P	S	Zn (10^{-3})
Achaak (AC)	mean	549	502	100.4	435	393	125.6	1.130	709	281	138.1
	SD	3	1	0.4	3	11	0.4	4×10^{-3}	4	7	0.4
Atocha (AT)	mean	471	387	90.4	492	453	129	1.92	804	361	90
	SD	4	6	1.0	2	2	2	0.04	5	5	1
Clon Cebas (CE)	mean	439	0.373	99.5	438	517	124	2.11	841	322	113
	SD	9	3	0.8	5	5	1	0.02	9	6	4
Chellaston (CH)	mean	590	500	12.51	472	457	125.0	1.251	818	318	165
	SD	7	2	0.6	6	9	0.6	6×10^{-3}	13	4	5
Cristomorto (CR)	mean	557	376	100.3	519	439	125.3	1.504	772	197	132
	SD	10	2	0.4	8	12	0.5	6×10^{-3}	15	5	6
Del Cid (DC)	mean	407	383	76.5	674	445	127.49	2.45	978	269	114.7
	SD	2	1	0.4	5	2	0.07	0.06	6	5	0.1
D-Largueta (DL)	mean	36	526	111	541	466	136	1.17	991	445	136
	SD	6	5	1	8	4	3	0.01	25	7	3
Ferragnes (FE)	mean	523	368	74.0	517	438	122.8	1.25	880	373	122.8
	SD	5	2	0.4	9	7	0.7	0.02	22	3	0.7
Genco (GE)	mean	394	369	110.6	503	392	122.9	1.721	705	219	131
	SD	7	2	0.3	9	4	0.3	4×10^{-3}	11	5	6
Malagueña (MA)	mean	461	386	87	527	390	129	1.16	724	352	128.8
	SD	4	1	3	8	5	1	9.94	10	2	0.4
Marcona (MR)	mean	419	385	113.4	575	451	128	1.263	965	341	154
	SD	8	9	0.4	11	9	3	10^{-3}	19	3	4
Non Pareil (NP)	mean	503	390	104.0	551	614	130.1	2.861	1052	274	130.1
	SD	8	1	0.2	6	5	0.2	4×10^{-3}	17	7	0.2
Peraleja (PE)	mean	340	322	88.0	466	433	123	2.58	836	305	111.6
	SD	8	8	0.5	11	23	1	0.03	24	4	0.7
Primorskyi (PR)	mean	298	373	73	556	433	122	1.38	848	250	97
	SD	4		2	4	5	3	0.03	8	5	3
Ramillete (RA)	mean	357	497	99.4	543	453	124.3	1.13	977	468	124.3
	SD	4	3	0.6	6	8	0.7	0.02	9	6	0.7
Texas (TE)	mean	598	383	157	565	394	128	1.70	690	319	115
	SD	5	7	3	6	4	3	0.03	9	9	2
Titan (TI)	mean	497	384	89.6	546	448	128	20.48	897	252	102.4
	SD	2	1	0.1	2	3	2	3×10^{-3}	10	5	0.1
Tuono (TU)	mean	382	379	130	515	448	126	2.02	906	400	117
	SD	6	9	3	5	2	3	0.05	8	5	3
Wawona (WA)	mean	674	435	92.5	536	454	147	1.82	889	334	91
	SD	5	10	0.2	7	4	1	0.03	11	6	1

[a] All values are given in mg/100 g of dried sample for $n = 9$.

Table 4. *F* Ratio Values and Percent of Sample Significant Differences When the Cultivars Are Compared in Pairs for All of the Inorganic Elements Considered

element	*F* ratio	% sig diff[a]
Ca	2722	97.1
S	1515	95.3
Na	4024	94.2
Fe	2119	93.0
P	530	87.7
Zn	414	87.1
K	558	87.1
Cu	1071	73.1
Mg	495	73.1
Mn	101	54.5

[a] Percent significant differences ($p = 0.05$) = (N/M) × 100, where N is the number of significant differences found when samples were compared in pairs and M is the maximum number of significant differences that can be found.

separates completely four classes (Desmayo-Largueta, Wawona, Peraleja, and the group formed by Ramillete, Achaak, and Chellaston) from the rest of cultivars.

Since it was known that all of the variables can have a significant power, principal component analysis (PCA) was applied to the previously autoscaled composition values. The eigenvalues obtained from the correlation matrix are 2.52, 2.16, 1.52, 1.22, 0.95, 0.81, 0.33, 0.26, 0.16, and 0.07. Choosing only eigenvalues >1 led to the reduction of 10 variables to 4 principal components (PC), accounting for 74.4% of the total variability. The percentages of variance for the four principal components are 25.2% for the first one, 21.7% for the second one, 15.2% for the third one, and 12.3% for the last one.

The correlation between each principal component and each original variable is given in Table 5. Nearly all of the elements analyzed have an important correlation to principal component one (PC1) except S and Mn, which have coefficients <0.30. With the second principal component (PC2), P, S, and Cu are strongly correlated, as are Mn and Mg, although to a lesser extent. Ca and, somewhat less so, Mn are correlated with the third principal component (PC3). The fourth principal component (PC4) is positively correlated with Zn, Mg, and Fe and negatively with Mn.

The communality value of a variable is the percentage of variance of the variable that is explained by the retained principal components. These values can be used to remove those variables that do not contribute to the retained principal components. In this study all of them are >0.10, as shown in Table 5, and therefore, all of the elements analyzed have been considered in the next statistical approaches applied.

The scores for the first two principal components and for the first and third principal components are plotted as scatter diagrams in Figure 3. These graphics show the peculiarities of the single American cultivars Texas, Wawona, and Non Pareil, as well as the association between Achaak and Chellaston cultivars in both cases. From the remainder, the Spanish cultivars Desmayo-Largueta, Marcona, and Ramillete present some consistency as a separated group only when PC2 versus

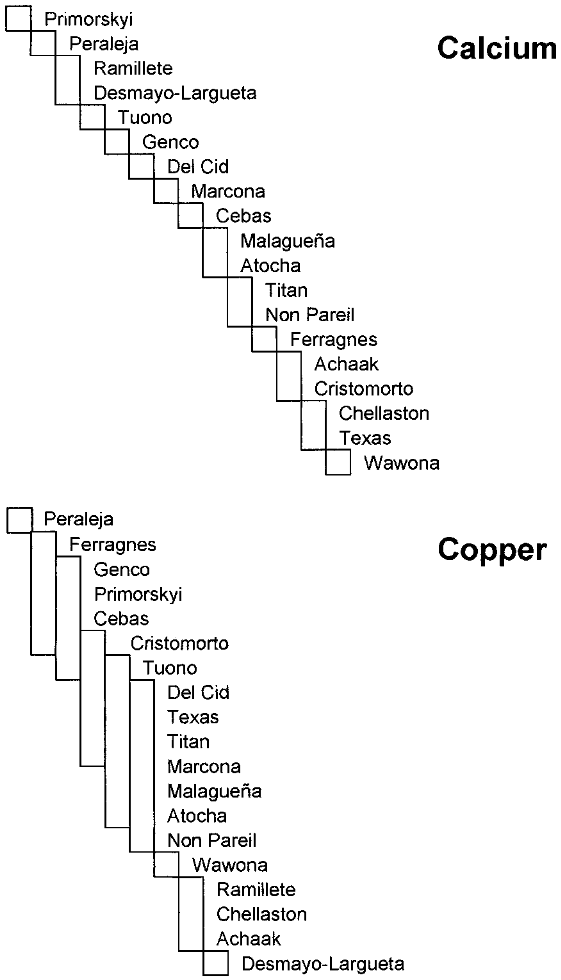

Figure 2. Representation of the results of one-way ANOVA for calcium and copper. Cultivars belonging to the same bar have mean values not significantly different from each other for the specified element.

Table 5. Principal Component Matrix and Communality Values for All of the Elements Studied

element	PC1	PC2	PC3	PC4	communality
Ca	−0.32	0.06	0.87	0.00	0.84
Fe	−0.46	0.19	0.29	0.44	0.52
K	0.46	0.25	−0.06	−0.27	0.35
Mg	0.60	0.47	0.15	0.51	0.86
Mn	0.10	0.54	0.63	−0.43	0.88
Na	0.80	−0.17	0.30	0.31	0.85
P	0.62	0.71	0.21	0.10	0.95
S	−0.17	0.68	0.32	−0.24	0.66
Zn	−0.53	0.28	0.18	0.59	0.75

PC1 is considered, but not when PC3 is taken into account.

Cluster analysis is a way to establish associations on the basis of similarity criteria between objects. As a result of the application of this algorithm, the dendrogram obtained is shown in Figure 4. The distances shown in the graph are rescaled to have all of them within the range from 1 to 25. At a distance value of 15, six associations are formed. Evidently, the American cultivars Non Pareil, Wawona, and Texas appear as singular ones and form three individual groups totally separated from the rest of the cultivars; a fourth group is constituted by Chellaston and Achaak; the fifth group consists of three Spanish cultivars Marcona, Ramillete, and Desmayo-Largueta, which display clear differences in relation to the rest of the Spanish almonds. The last group is made up of the rest of

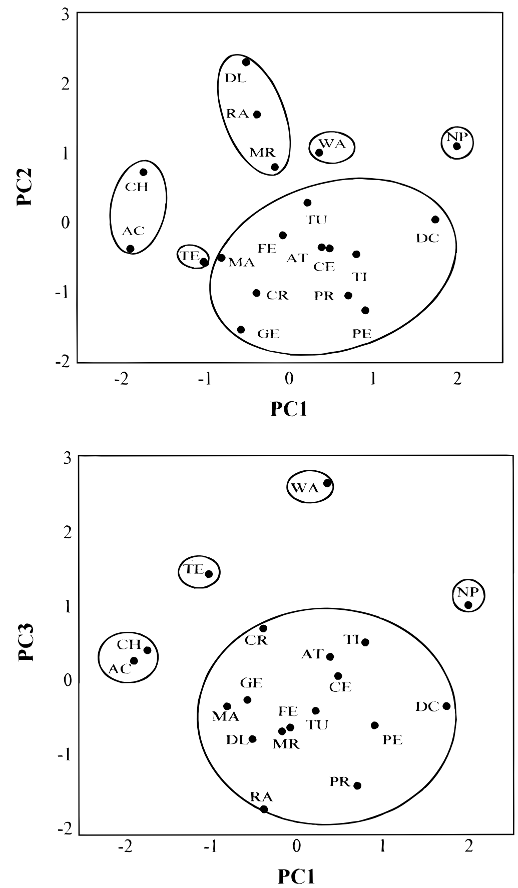

Figure 3. Projection of scores for cultivars in the reduced space determined by two principal components with hypothesized groups circled: (a) PC1 vs PC2; (b) PC1 vs PC3. See Table 3 for abbreviations of the cultivars used.

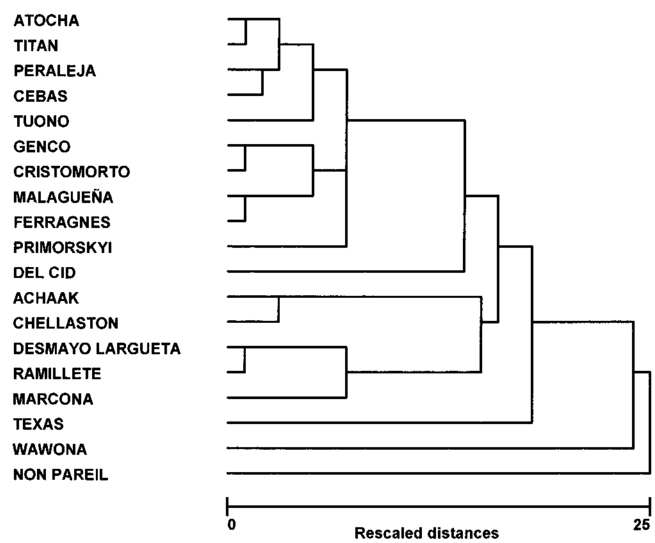

Figure 4. Dendrogram for hierarchical cluster analysis of the data set.

almonds cultivated in the Mediterranean area, except Titan (American), which shows a great similarity with the Spanish cultivar Atocha. This classification agrees with the one previously visualized by applying the PCA.

The classification that was found by applying PCA and cluster analysis to the inorganic element composition of the almonds separates the three American cultivars Non Pareil, Wawona, and Texas from the rest of almonds, most of them cultivated in the Mediterranean area. According to our previous research, neither Non Pareil and Texas, using the fatty acid composition (García-López et al., 1996), nor Non Pareil, using the amino acid composition (Prats-Moya and Berenguer-Navarro, 1994), can be separated from the Spanish cultivars.

Linear discriminant analysis (LDA) is an algorithm in which mathematical functions (linear combinations of the original variables) are constructed in such a way that the differences among the established groups are maximized.

First, this algorithm was applied for the consideration of only four associations, wherein Non Pareil, Wawona, and Texas are assumed to be independent classes to test their peculiarities and the fourth group is made up of the remaining cultivars. Three discriminant functions are obtained, explaining 44.74%, 32.22%, and 23.04% of the total variance. The first discriminant function is highly correlated with Ca and P and the second one with Mn and Zn, whereas the third function is correlated with the rest of the elements considered. By using the classification functions calculated, no cultivar was misclassified.

If the discriminant analysis algorithm is applied to the six associations deduced in the cluster analysis, instead of four as before, five discriminant functions are obtained, which explain 48.6%, 18.72%, 14.05%, 10.77%, and 7.95% of the total variance. Cu and Zn are strongly correlated with the first discriminant function, while the second and the third ones are correlated with P and Mn, respectively; Mg, Ca, Na, and S are correlated with the fourth discriminant function, and K and Fe with the fifth function.

An ANOVA analysis was performed using the scores obtained for the five discriminant functions, and the first discriminant function was confirmed to be able to establish significant differences among all of the groups formed, except between the groups Achaak and Chellaston and that of Desmayo-Largueta, Ramillete, and Marcona.

Classification power of different inorganic elements can be estimated by calculation of classification functions and from the variation in percentage of correct classification as long as variables are successively removed in the analysis. Thus, the first classification remains 100% correct after Cu, Na, S, and P are discarded, begins to fail significantly to 98.25% when Ca is not included in the computation, remains at the same level after Zn is also removed, and drops sharply to 82.46% when Fe is eliminated, leaving only Mn and Mg.

The case of the Cu variable is remarkable. While the F ratio value for Cu was one of the highest when six groups were submitted to the LDA, it becomes one of the lowest when only four groups are considered (Table 6).

From the comparison of the LDA analysis it follows that different variables are significant according to the cultivars we would like to distinguish. Ca and Fe can be considered important parameters to characterize the three American cultivars Non Pareil, Wawona, and Texas, but Cu and Zn can be paramount to get a complete classification and, in particular, to define the Mediterranean almonds as a group. Altogether, the mineral fraction of almonds can be considered a characteristic feature of them, which cannot be explained adequately by soil and climate factors alone.

Table 6. Comparison of *F* Ratios ($p = 0.05$) When Four and Six Groups Are Considered in the Discriminant Analysis

element	*F* ratio	
	four groups	six groups
Ca	36.4	59.3
Cu	1.6	98.6
Fe	42.8	43.6
K	3.1	12.8
Mg	113.0	75.9
Mn	105.7	82.9
Na	21.7	41.6
P	26.5	58.2
S	1.4	21.5
Zn	10.3	48.9

LITERATURE CITED

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

Cattell, R. B. Scree test for the number of factors. *Multivariate Behav. Res.* **1966**, *1*, 245−276.

Felipe, A.; Vargas, F. J.; García, J. E. *Variedades tipificadas de almendra en España*; Ed. 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.

Gómez-Martínez, B. *Bioquímica de la almendra*; Instituto de Estudios Alicantinos, Diputación Provincial de Alicante: Alicante, Spain, 1976.

Grasselly, C.; Crossa-Raynaud, P. *L'Amandier*; Maisonneuve et Larose: Paris, France, 1980.

Prats Moya, M. S.; Berenguer-Navarro, V. Characterization of some varieties of almonds using its composition in free amino acids. *Rev. Esp. Cienc. Tecnol. Aliment.* **1994**, *34*, 218−227.

Saura-Calixto, F.; Cañellas-Mut, J.; Soler, L. *La Almendra, composición, variedades, desarrollo y maduración*; Monografias I.N.I.A: Madrid, 1988.

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.

Received for review September 30, 1996. Revised manuscript received February 28, 1997. Accepted March 3, 1997.[⊗]

JF960740K

[⊗] Abstract published in *Advance ACS Abstracts*, April 15, 1997.