Discriminate Analysis of Roasted Coffee Varieties for Trigonelline, Nicotinic Acid, and Caffeine Content

S. Casal,[†] M. B. P. P. Oliveira,*[,†] M. R. Alves,[‡] and M. A. Ferreira[†]

CEQUP/Serviço de Bromatologia, Faculdade de Farmácia, Universidade do Porto. Rua Aníbal Cunha 164, 4050-047 Porto, Portugal, and ESTG, Instituto Politécnico de Viana do Castelo, Av. do Atlântico, Apartado 574, 4900 Viana do Castelo, Portugal

Arabica and *robusta* roasted coffees from several geographical origins, in a total of 29 samples, were characterized for their contents in caffeine, trigonelline, and nicotinic acid by a recently developed HPLC/diode-array detector method. All samples were subjected to the same roasting procedure in order to eliminate the variations due to this process. Characterization was achieved by applying multivariate and nonparametric analysis to the chromatographic results. The two coffee varieties were clearly separated by their trigonelline and caffeine contents. Nicotinic acid could not be used as a variety discriminate factor. There was no association with the geographical origin of the samples.

Keywords: *Roasted coffee; caffeine; trigonelline; nicotinic acid; multivariate and nonparametric analysis*

INTRODUCTION

More than 80 species of the genus *Coffea* L. (Rubiaceae) are known. The most important are *Coffea arabica* and *Coffea canephora*, which account, respectively, for about 75% and 24% of the world production. Coffee is an expensive raw material, especially *arabica* coffee, and over the years many fraudsters have been tempted to falsify the product declaration due to the increasing practice of selling coffees on the basis of their botanical and/or geographic origin (Prodolliet, 1996).

International coffee trade is conducted almost exclusively with green coffee. In this state, *arabica* and *robusta* coffees are easily distinguished by their appearance (e.g., size, shape and color). Once roasted and/or ground, the form in which the coffee is commercially available to the consumers, this visual criterion is eliminated. Efficient methods are thus required for authentication of roasted coffee beans and for detecting trading fraud.

Efforts have been made to characterize the two coffee species using chemical data. Nevertheless, the chemical composition depends not only on the species and variety in question, but also on the degree of roasting and, to a lesser extent, on other factors such as agricultural practices, degree of maturation, storage conditions, and geographical origin. These factors are not under the analyst's control which makes it extremely difficult to quote average values for any type of coffee (Macrae, 1989). The main difference between the *arabica* and *robusta* roasted coffee seems to be the composition of the unsaponifiable matter (Folstar, 1989), specifically the presence of 16-*O*-methylcafestol in *robusta* coffee (Speer et al., 1991). Chlorogenic acids can also be successfully applied to this discrimination and, either alone or in conjunction with caffeine, have been used to characterize commercial and noncommercial species of coffee (Clifford, 1989; Correia et al., 1995; Andrade et al., 1997; Bicchi et al., 1995). Caffeine is probably the single most analyzed chemical factor in coffee and has also been frequently used in the discrimination of green coffee varieties (Clifford, 1987; Macrae, 1989). Several approaches to coffee authenticity have also been reported based on spectroscopic techniques (Kemsley et al., 1995; Briandet et al., 1996; Downey et al., 1997), aroma discrimination (Aishima, 1991), solid-phase microextraction−gas chromatography (Bicchi et al., 1997), and volatile profiles (Martín et al., 1996). Several of these works associated chemical analysis with pattern recognition techniques.

The aim of this work represents a contribution to the discrimination of the most representative coffee varieties and eventually to access their geographical origins. The compounds analyzed simultaneously were trigonelline (*N*-methylnicotinic acid), nicotinic acid, and caffeine by a recently developed HPLC/diode-array detector method (Casal et al., 1998). It was our objective to find out whether nicotinic acid could be used as a discriminate factor and to confirm the ability of caffeine and trigonelline contents to discriminate *arabica* and *robusta* roasted coffees.

MATERIALS AND METHODS

Coffee Samples. Roasted beans samples from both *Coffea canephora* var. *robusta* and *Coffea arabica* were studied. *Coffea robusta* samples ($n = 20$) were from several geographical origins: Ivory Coast (IC), India (IN), Honduras (H), Vietnam (VN), Angola (A), Uganda (UG), and Cameroon (CM). *Coffea arabica* samples ($n = 9$) were from Brazil (Br), Mexico (MX), Colombia (CO), Guatemala (GO), and Costa Rica (CR).

* To whom correspondence should be addressed (e-mail bromato@ff.up.pt; telephone 351-22-2078902; fax 351-22-2003977).

[†] CEQUP/Serviço de Bromatologia.

[‡] ESTG, Instituto Politécnico de Viana do Castelo.

10.1021/jf990702b CCC: $19.00
Published on Web 07/08/2000

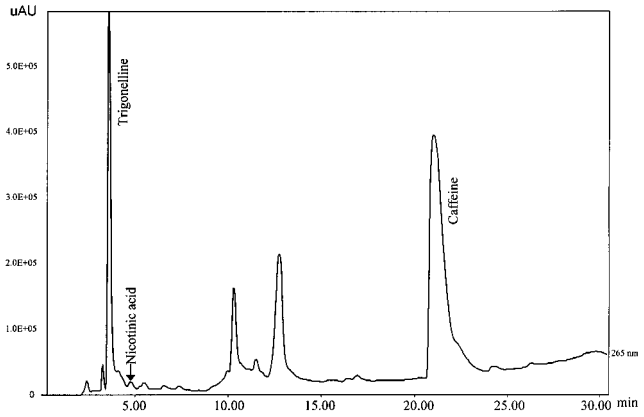

Figure 1. Typical HPLC profile of a roasted *arabica* coffee from Brazil.

A local importer and roaster of coffee supplied all coffee samples and was able to confirm their botanical and geographical origin. A standard method was used in the roasting procedure in order to eliminate the variations due to this process. Samples were hermetically sealed and stored at −20 °C until used for chemical analysis.

Sample moisture was determined in order to compare the results in a dry weigh basis.

Sample Preparation for Chemical Analysis. For extraction of the compounds in study, 2-g portions of each powdered coffee sample (25-mesh) were extracted with a total of 100 mL of boiling water (Casal et al., 1998). The solution was filtered into appropriate vials for use in the autosampler.

HPLC Analysis. HPLC analysis was achieved as described by Casal et al. (1998) with an analytical HPLC unit (Jasco) consisting of two PU-980 pumps, a MD-910 diode-array detector, and an AS-950 autosampler. A reversed-phase Spherisorb ODS2 (5 μm particle size, 25.0 × 0.46 cm) column was used. The solvent system used was a gradient of phosphate buffer pH 4.0 (0.1M) and methanol performed at a constant flow rate of 1.0 mL min^{-1} at room temperature. Detection was accomplished with a diode-array detector, and chromatograms were recorded at 265 nm.

The compounds were identified by their retention times, chromatographic comparisons with authentic standards, and their UV spectra. Quantification was based on the external standard method. Under the assay conditions described, a linear relationship between the concentration and the UV absorvance was obtained at 268 nm for trigonelline, 264 nm for nicotinic acid, and at 276 nm for caffeine.

To study the recovery of the procedure, known quantities of the three standards were added to one sample of roasted coffee, and the percentage recovery was calculated after triplicate analyses. These recovery values were 101 ± 1% for trigonelline, 98 ± 1% for nicotinic acid, and 99 ± 1% for caffeine.

Statistical Analysis. Principal Component Analysis was carried out after standardization of data to mean zero and unit variance and following standard procedures (Mardia et al., 1979).

Cluster analyses were performed after standardization of variables to unit variance and mean zero and using Euclidean distances as the distance measure and following the unweighted pair group average method as the linking method (Mardia et al., 1979) as implemented in the Statistica for Windows Statistical package.

Spearman's Rho was calculated by substitution of the actual values observed by the respective ranks, followed by computation of

$$\rho = \text{SUM}\,[R(x_i) - (n+1)/2][R(y_i) - (n+1)/2]/n(n^2-1)/12$$

where $R(x_i)$ and $R(y_i)$ represent the ranks for each pair of observations and n equals the total number of observations (or pairs).

Kendall's Tau was calculated by checking the number of concordant (Nc) and discordant (Nd) pairs of observations out of the total possible pairs of observations and computing the quantity

$$T = (\text{Nc} - \text{Nd})/n(n-1)/2$$

where n refers to the total number of possible pair of observations.

Both measures of correlations (ρ and T) were then compared with appropriate tables to check for significance (Conover, 1980).

RESULTS AND DISCUSSION

A typical chromatogram obtained with a roasted sample from Brazil (*Coffea arabica*) is represented in Figure 1.

Summary results are presented in Figure 2, in the form of conventional box and whiskers plots, displaying,

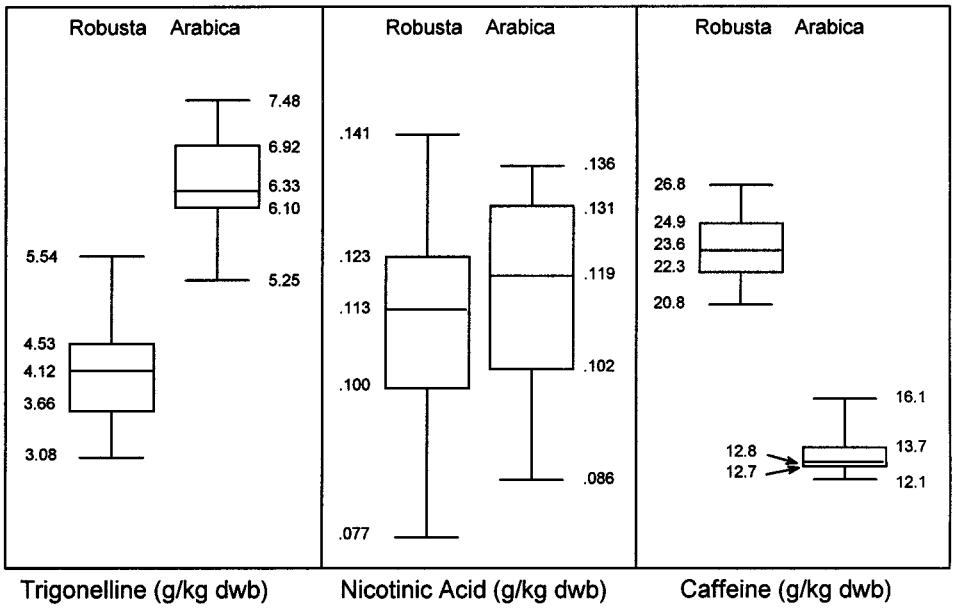

Figure 2. Univariate nonparametric analysis of trigonelline, nicotinic acid, and caffeine in roasted coffee samples.

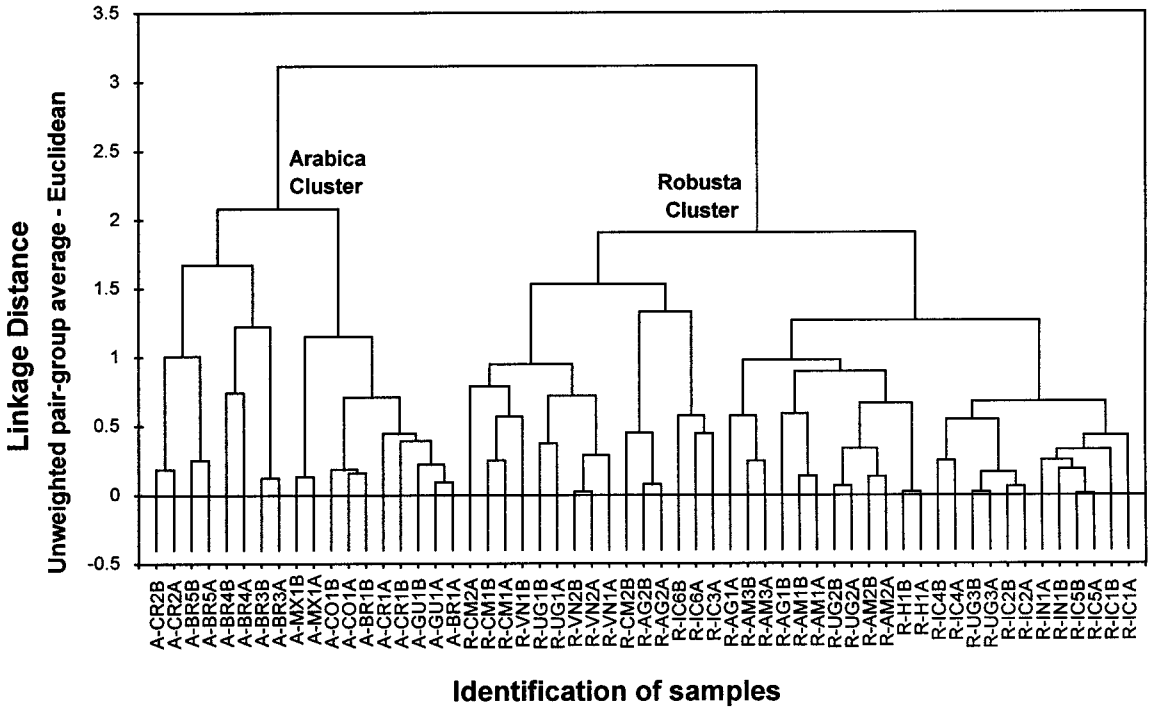

Figure 3. Cluster analysis of roasted coffee samples.

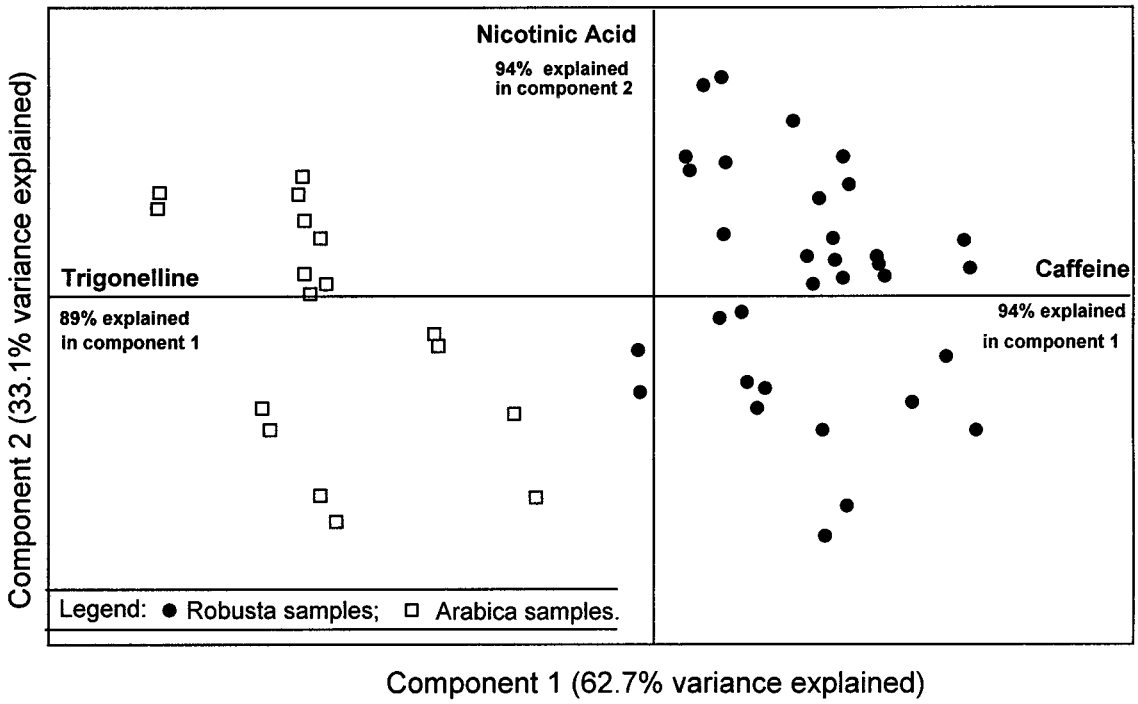

Figure 4. Principal component analysis of coffee samples on the three chemical parameters.

for each variable and for each coffee variety, the minimum values, first quartile, median, third quartile, and maximum values observed. All values are expressed in g·kg^{-1} on a dry weight basis (dwb).

It can be seen that there is a clear distinction between *arabica* and *robusta* samples in what concerns the caffeine levels. The levels found are in accordance with the literature (Clifford, 1987; Macrae, 1989). The variation between *arabica* and *robusta* coffees is mainly due to genetic differences found in the varieties of coffee beans, as it is stable on roasting (Macrae, 1989).

In the same figure we can observe that the differences in trigonelline between varieties are also apparent, but the nicotinic acid levels vary widely and assume similar levels in both varieties. Trigonelline and nicotinic acid

Table 1. Overall and Pooled within Groups Simple Linear Correlations of Trigonelline, Nicotinic Acid, and Caffeine in the Roasted Coffee Samples[a]

	overall correlations			pooled within groups correlations		
	trig	nic	caf	trig	nic	caf
trig	1.00			1.00		
nic	0.01	1.00		−0.11	1.00	
caf	−0.85[b]	−0.21	1.00	−0.29	−0.42[b]	1.00

[a] trig = trigonelline; nic = nicotinic acid; caf = caffeine. [b] Significant correlations at $p = 0.05$

Table 2. Spearman Rho and Kendall Tau Measure Correlations: (a) Overall Correlations, (b) Correlations within *arabica* Variety, and (c) Correlations within *robusta* Variety

		Spearman Rho			Kendall Tau		
		trigonelline	nicotinic acid	caffeine	trigonelline	nicotinic acid	caffeine
all samples	trigonelline	–			–		
	nicotinic acid	−0.03625	–		−0.018303	–	
	caffeine	−0.740277[a]	−2.51137[a]	–	−0.516291[a]	−0.207642[a]	–
arabica	trigonelline	–			–		
	nicotinic acid	0.108359	–		0.071895	–	
	caffeine	0.001032	−0.686275[a]	–	0.045752	−0.542484[a]	–
robusta	trigonelline	–			–		
	nicotinic acid	−0.276739	–		−0.158372	–	
	caffeine	−0.303441	−0.391752[a]	–	−0.192982	−0.232098[a]	–

[a] Significant correlations at $p = 0.05$

levels are known to be related with the intensity of roasting being the former progressively degraded and partially converted into nicotinic acid (Macrae, 1989). As our objective was to find correlation with the botanical and eventually geographical origin we have eliminated this variation factor by subjecting all samples to the same exactly roasting procedure (14 min from 160 to 220 °C).

A cluster analysis was carried out, with the corresponding dendogram show in Figure 3. Two distinct clusters representing each coffee variety are obvious, but other clusters, formed at lower levels, although apparent, could not be attributed to the geographical origin or to any other specific reason.

As cluster analysis shows the existence of clusters without explaining the reasons assisting their formation, a principal component analysis was done, as shown in Figure 4. It becomes clear that the first principal component, which explains two-thirds of the total information, represents an opposition, or strong negative correlation, between the levels of trigonelline and caffeine, the former being lower and the latter higher in *robusta* than in *arabica* samples. The second principal component, which explains roughly one-third of the total information, represents differences in composition of nicotinic acid, which, being represented in the second component, is not correlated with the first two parameters under study. It is important to note that, as already observed in Figure 2, there is a great variability within each variety cluster in what concerns the nicotinic acid levels.

The strong negative correlation between caffeine and trigonelline and the absence of correlation between these two variables and nicotinic acid, as displayed in Figure 4, must be seen with caution since all data is analyzed simultaneously, without taking notice of the existence of clusters. Consequently, a correlation analysis was carried out, with all coffee samples together and also analyzing the pooled within groups correlation, taking *arabica* and *robusta* samples as two distinct groups. The results are summarized in Table 1.

When coffee varieties are analyzed separately, the correlation between trigonelline and caffeine is not significant, and a significant negative correlation between nicotinic acid and caffeine is observed.

Using more powerful, nonparametric statistics, which overcome the high variations observed within each coffee group, such as the Spearman's Rho and Kendall's Tau, results are obtained as represented in Table 2 were found. Considering overall correlation, significant negative correlation between trigonelline and caffeine are observed, but also a negative correlation between caffeine and nicotinic acid becomes apparent even when all samples are considered together. Analyzing coffee varieties separately, only negative correlation between nicotinic acid and caffeine are significant, in accordance with the linear correlation results.

Therefore it seems that the negative correlation between trigonelline and caffeine is only expressing general differences between the levels of these substances in *arabica* and *robusta* coffees, i.e., is caused by the fact that caffeine is higher in *robusta* samples while trigonelline is lower. However, when the pooled within groups correlation are analyzed, one can see that there is no correlation between these two components. An important feature of these results is the negative correlation between nicotinic acid and caffeine, which can be observed within each coffee variety and also when all data is taken together.

CONCLUSIONS

From the results obtained we can concluded that trigonelline and caffeine can be used in the discrimination of pure roasted coffees. The results show that nicotinic acid cannot be used in the discrimination of these two coffee varieties contradicting what was initially thought (Macrae, 1989). Neither trigonelline nor caffeine can be used for identification of the geographical origin of the roasted coffee. As the nature of the green beans' processing employed may affect the final concentration in the roasted product, further studies are recommended to confirm this possibility.

The conjunction of these chemical parameters may be helpful in the identification of coffee varieties, and the methodology proposed is appropriate for routine analysis in the coffee industry.

ACKNOWLEDGMENT

S. Casal is indebted to Subprograma Ciência e Tecnologia do 2° Quadro Comunitário de Apoio (BD/9580/96).

LITERATURE CITED

Aishima, T. Aroma Discrimination by Pattern Recognition Analysis of Responses from Semiconductor Gas Sensor Array. *J. Agric. Food Chem*. **1991**, *39*, 752−756.

Andrade, P. B.; Leitão, R.; Seabra, R. M.; Oliveira, M. B.; Ferreira, M. A. 3,4-Dimethoxycinnamic acid levels as a tool for differentiation of *Coffea canephora* var. *robusta* and *Coffea arabica*. *Food Chem*. **1997**, *61* (4), 511−514.

Bicchi, C. P.; Panero, O. M.; Pellegrino, G. M.; Vanni, A. C. Characterization of Roasted Coffee and Coffee Beverages by Solid-Phase Microextraction − Gas Chromatography and

Principal Component Analysis. *J. Agric. Food Chem*. **1997**, *45* (12), 4680−4686.

Bicchi, C. P.; Binello, A.; Pellegrino, G. M.; Vanni, A. C. Characterization of Green and Roasted Coffees through the Chlorogenic Acid Fraction by HPLC−UV and Principal Component Analysis. *J. Agric. Food Chem*. **1995**, *43*, 1549−1555.

Briandet, R.; Kemsley, E. K.; Wilson, R. H. Discrimination of *Arabica* and *Robusta* in Instant Coffee by Fourier Transform Infrared Spectroscopy and Chemometrics. *J. Agric. Food Chem*. **1996**, *44*, 170−174.

Casal, S.; Oliveira, M. B.; Ferreira, M. A. Development of an HPLC/Diode-array detector method for simultaneous determination of trigonelline, nicotinic acid and caffeine in coffee. *J. Liq. Chrom., Relat. Technol*. **1998**, *21* (20), 3187−3195.

Clifford, M. N. Chemical and Physical aspects of green coffee and coffee products. In *Coffee: botany, biochemistry, and production of beans and beverage*; Clifford, M. N., Wilson, K. C., Eds.; Croom Helm: Beckenham, U.K., 1987; pp 305−374.

Clifford, M. N. Chlorogenic acids. In *Coffee*: *Chemistry*; Clarke, R. J., Macrae, R., Eds.; Elsevier Applied Science Publishers: London, 1989; Vol. 1, pp 153−202.

Conover, W. J. *Practical Nonparametric Statistics*; John Wiley & Sons: New York, 1980.

Correia, A. M. N. G.; Leitão, M. C. A.; Clifford, M. N. Caffeoyltyrosine and Angola II as characteristic markers for Angolan *robusta* coffees. *Food Chem*. **1995**, *53*, 309−313.

Downey, G.; Briandet, R.; Wilson, R. H.; Kemsley, E. K. Near- and Mid-Infrared Spectroscopies in Food Authentication: Coffee Varietal Identification. *J. Agric. Food Chem*. **1997**, *45*, 4357−4361.

Folstar, P. Lipids. In *Coffee*: *Chemistry*; Clarke, R. J., Macrae, R., Eds.; Elsevier Applied Science Publishers: London, 1989; Vol. 1, pp 203−222.

Kemsley, E. K.; Ruault, S.; Wilson, R. H. Discrimination between *Coffea arabica* and *Coffea canephora* variant *robusta* beans using infrared spectroscopy. *Food Chem*. **1995**, *54*, 321−326.

Macrae, R. Nitrogenous components. In *Coffee*: *Chemistry*; Clarke, R. J., Macrae, R., Eds.; Elsevier Applied Science Publishers: London, 1989; Vol. 1, pp 115−137.

Mardia, Kent and Bibby. *Multivariate Analyses*; Academic Press: London, 1979.

Martín, M. J.; Pablos, F.; González, A. G. Application of pattern recognition to the discrimination of roasted coffees. *Anal. Chim. Acta* **1996**, *320*, 191−197.

Prodolliet, J. Authenticity of coffee. In *Food Authentication*; Ashurst, P. R., Dennis, M. J., Eds.; Blackie Academic & Professional: London, 1996; pp 304−338.

Speer, K.; Tewis, R.; Montag, A. 16-O-methylcafestol. A quality indicator for coffee. In *14th Colloque Scientifique International sur le Café-San Francisco*; ASIC: Paris, 1991; pp 237−244.

Received for review June 28, 1999. Revised manuscript received May 11, 2000. Accepted May 13, 2000.

JF990702B