Chlorogenic Acids as a Potential Criterion in Coffee Genotype Selections

G. Guerrero and M. Suárez*

Departamento de Química, Universidad Nacional de Colombia, A.A 14490 Bogotá, Colombia

G. Moreno

Disciplina de Mejoramiento Genético, Cenicafe, Chinchiná, Caldas, Colombia

A systematic study by HPLC was conducted to determine the content of chlorogenic acids in green coffee beans of the Variedad Colombia (*Coffea arabica* Caturra var. × Híbrido de Timor) and in other genotypes of interest of *C. arabica*, *Coffea canephora*, Híbrido de Timor, and the F1 offspring derived from the crossing Caturra × Híbrido de Timor. Comparisons were made of the total content of these acids, their quantitative differences, and the presence or absence of some of them, using parametric statistical techniques and multivariate analysis. Total content differences were found between *C. canephora* accessions, between Típica and Caturra varieties of *C. arabica*, and between Híbrido de Timor accessions. The chromatographic profile comparisons with principal component analysis separated in the first component *C. canephora* accessions from the rest of the genotypes, whereas the second component separated *C. canephora* accessions.

Keywords: *Coffea; C. arabica; C. canephora; Variedad Colombia; chlorogenic acids; multivariate analysis; HPLC; Rubiaceae*

INTRODUCTION

The chlorogenic acids (CGA) content of the two main species of coffee grown, *Coffea arabica* and *Coffea canephora*, varies between 7 and 10% on dry basis (dmb) (*1*). In the Híbrido de Timor, a coffee population probably of interspecific origin (*C. arabica* × *C. canephora)*, this content is intermediate (*2*). The chlorogenic acids content was used as a taxonomic criterion in *Coffea* and *Psilanthus* and in interspecific hybrids such as Arabusta (*C. arabica* × *C. canephora*), which showed an intermediate concentration relative to that of its parents (*3*). The content of the major chlorogenic acids subgroups [feruloylquinic (FQA), caffeoylquinic (CQA), and dicaffeoylquinic (diCQA)] varies with maturity (*4*, *5*), whereas the minor compounds, usually not identified, are associated with the origin of the bean; some of their variations may account for the kind of genotype (*6*). Furthermore, studies on the possibility of distinguishing green or roasted coffees of different origins by the CGA fraction have been carried out through high-performance liquid chromatography with ultraviolet detection (HPLC-UV) and principal component analysis (PCA) (*7*). Taking this information as a basis, the Variedad Colombia, a composite cultivar resistant to rust (*Hemileia vastatrix*), which was obtained at Cenicafé, Colombia (*8*), was studied as a model population together with other *C. arabica* and *C. Canephora* genotypes in order to determine similarities or differences that might be relevant as selection guidelines in genetic breeding programs.

* Author to whom correspondence should be addressed [telephone (571) 3165000, ext. 14440; fax (571) 3165220; e-mail msuarez@ciencias.ciencias.unal.edu.co].

MATERIALS AND METHODS

Plant Material. Collections were made of physiologically mature fruit (32 weeks after flowering) of Típica, Caturra, and Colombia (*C. arabica* var. Caturra × Híbrido de Timor) varieties of *C. arabica*, three Híbrido de Timor accessions, the first generation (F1) of Caturra × Híbrido de Timor crossing, and three accessions of *C. canephora.* These materials were planted in the Cenicafe germplasm collection in Chinchiná, Caldas, at an altitude of 1400 m above sea level (mean temperature of 21.7 °C, annual rainfall of 2258.6 mm, and relative humidity of 75%). The fruits were subjected to wet processing, and the seeds were dried in the sun. Parchment-free beans known as green coffee were used for the study. The beans were ground in the presence of liquid nitrogen and were stored at −20 °C in order to perform the analyses.

Chlorogenic Acids Extraction. Four successive extractions of 2 g of ground coffee were made with 25 mL of aqueous ethanol (80%); alcohol fractions were combined (100 mL) and concentrated under reduced pressure. The aqueous residue was extracted with petroleum ether four times to eliminate lipids and pigments. Ammonium sulfate was added to the aqueous extract to a final concentration of 20 g/L for protein precipitation (*9*). Most of the caffeine was extracted from the aqueous fraction with chloroform (4 × 20 mL), and another four extractions were made with 20 mL of ethyl acetate to extract the chlorogenic acids. The ethyl acetate phases were combined, and the residual water was eliminated using sodium sulfate. The extract was dried under a nitrogen stream, and the residue was dissolved in 10 mL of methanol for HPLC analysis. Extraction of total chlorogenic acids content was made using Lehman's technique (*10*).

Identification. Chlorogenic acids (CQA, FQA, and diCQA) showed the same UV spectra (220−340 nm) and were identified by comparison of retention times in relation to 5-caffeoylquinic acid (*11*, *12*). Separation was accomplished in a Varian 500 HPLC with a diode array detector (Hewlett-Packard 1040 A), a Lichrosorb RP-18 precolumn, and a Hibar Lichrospher RP-18 column (250 mm × 4.6 mm; 5 μm). These conditions were previously reported (*6*).

10.1021/jf001286u CCC: $20.00
Published on Web 04/20/2001

Table 1. Total Dry Base Content of Chlorogenic Acids[a]

genotype	chlorogenic acids (%)	group[b]
C. canephora BP46	13.5	A
C. canephora Centro1	13.36	A
C. canephora BP4	12.09	B
Var. Típica of *C. arabica*	11.13	C
Híbrido de Timor 832	11.21	C
F1 Caturra × Híbrido de Timor	10.51	D
Variedad Colombia	10.06	E
Híbrido de Timor 2252	9.29	F
Híbrido de Timor 1343	9.18	F
Var. Caturra of *C. arabica*	7.87	G

[a] Groupings were done using the Duncan test with a 99% probability; values correspond to the average of three determinations. [b] A−G, classes based on statistically significant differences.

The main isomers were identified by HPLC−mass spectrometry (HPLC-MS) by comparing them with reference spectra (*13*) and on the basis of analysis of spectral data. A system consisting of an HP particle beam 5988A and a Hibar Lichrospher RP-18 column (250 mm × 4.6 mm; 5 μm) was used. The elution solvents were (A) 2 mM H_3PO_4 and (B) MeOH, and the elution program at room temperature (25 °C) was as follows: 0−10 min, 60% B (isocratic); 10−25 min, 100% B (linear gradient); flow rate = 1.0 mL/min.

Quantification. The total chlorogenic acids quantification was achieved using ultraviolet spectrophotometry (*10*). The quantification of each isomer was performed using the external standard method, integrating the area of each of the chromatograph peaks, and relating them to the 5-caffeoylquinic acid pattern (1 mg/mL). For the extraction of total chlorogenic acids, the materials were extracted for 3 h. Three replications per genotype were made.

Statistical Analysis. Variance analysis was used for the total chlorogenic acids content of the genotypes, and mean separation techniques were applied (Duncan, P = 99%) to identify statistically different genotypes. Multivariate analysis tests were used for chromatographic profile comparisons, using chromatograph peaks as variables and evaluating them on the basis of their concentrations. The tests used were PCA, correspondence factorial analysis (CFA), hierarchical ascending classification (HAC), and discriminating factorial analysis (DFA). Analyses were performed using SAS and STAT-ITCF statistical software.

RESULTS AND DISCUSSION

Table 1 shows the average total chlorogenic acids content, together with the statistical classification of the genotypes on the basis of the percentage of these acids. The highest contents were found in the three *C. canephora* accessions and the lowest in the Caturra variety. These values are consistent with those reported by other authors to *C. arabica* and *C. canephora* coffees (*14*). Statistically significant differences were found among total chlorogenic acids content, which allowed the discrimination of the genotypes into seven clearly defined groups (A−G). Important differences were observed between the Caturra (7.87%) and Típica (11.36%) varieties of *C. arabica*. Variedad Colombia lies between these two (10.06%). Differences between Híbrido de Timor accessions confirm the diversity of this population. Híbrido de Timor accession 832 shows a relatively high content and was classified in the same group (C) as the Típica variety. This result could be explained as suggested in another study (*15*) that one of the Híbrido de Timor's parents could have been this variety. Although the Variedad Colombia and the F1 generation of Caturra × Híbrido de Timor belong to statistically different groups, their contents are very similar, suggesting that total quantities have remained without noticeable modifications throughout the selection process.

Table 2 shows the quantification and identification results for the CGA chlorogenic acids. Under the HPLC conditions used (*6*), 16 compounds were separated, of which the following acids were identified: 5-CQA, 5-FQA, 3,4-diCQA, 3,5-diCQA, 4,5-diCQA, 3,4-CFQA, and 4,5-CFQA. The first five are present in all of the genotypes, whereas 3,4-CFQA was not found in the Caturra and Colombia varieties, the Caturra × Híbrido de Timor F1 generation, or Híbrido de Timor 1343. 4,5-CFQA was not detected in Caturra or Típica. The remaining nine compounds were not identified. Among that group, Clifford has also reported isomers with relative retention times of 2.04, 2.13, and 2.21 as unidentified. The isomers corresponding to retention times of 0.72, 1.13, 1.26, 1.35, 1.44, and 2.57 showed ultraviolet spectra characteristic of chlorogenic acids but were not reported by this author. These latter compounds could be characteristics of some of the studied genotypes.

When the data were analyzed by regions, it was found that in the range corresponding to CFQA (t_R = 2.03−

Table 2. Identification and Quantification of Chlorogenic Acids in Each Coffee Genotype, Comparing Retention Times Relative to 5-Caffeoylquinic Acid Reported by Clifford (11) (Grams per Kilogram on Dry Basis) and Relative Retention Times (t_R), for Variety (Var.), Híbrido de Timor (HT), Canephora (CAN), and First Generation of Caturra × Híbrido de Timor (F1)

		coffee genotype									
compound	t_R	Var. Caturra	Var. Tipica	Var. Colombia	F1	HT 1343	HT 832	HT 2252	CAN BP4	CAN BP46	CAN CEN1
unknown	0.72	0.756	1.253	1.161	0.996	1.542	1.391	1.605	0.405	2.002	1.265
5-CQA	1.00	21.371	23.914	19.004	24.167	27.074	25.675	21.352	11.577	27.408	23.996
unknown	1.13	0.036	0.035	0.068	0.085	0.059	0.087	0.045	0.206	–[a]	–
4-FQA	1.26	0.206	0.090	0.107	0.130	0.118	0.057	0.078	–	–	–
unknown	1.30	0.193	0.176	0.262	0.268	0.225	0.261	0.260	0.170	0.422	0.341
5-FQA	1.35	2.079	1.634	1.988	2.695	2.218	2.483	1.620	2.436	4.763	4.401
unknown	1.44	–	0.439	0.529	0.106	0.244	0.453	0.679	–	0.345	–
3,4-diCQA	1.73	1.161	0.916	0.761	1.105	1.152	1.351	0.878	2.012	3.140	3.136
3,5-diCQA	1.82	3.642	3.696	2.222	3.090	2.591	3.085	2.039	2.646	4.085	4.191
4,5-diCQA	1.92	1.126	0.964	0.673	1.170	0.915	0.951	0.821	1.550	2.567	2.430
3,4-CFQA	2.03	–	0.137	–	–	–	0.216	0.081	0.667	0.590	0.697
unknown	2.04	0.1	–	0.037	–	0.032	0.323	0.100	–	–	–
unknown	2.13	–	–	–	–	0.189	–	–	0.852	0.658	1.032
unknown	2.21	–	–	–	–	0.049	0.161	0.062	0.641	0.599	0.833
4,5-CFQA	2.30	–	–	0.135	0.175	0.139	0.166	0.159	1.269	1.574	1.431
unknown	2.57	–	–	–	–		–	–	0.325	–	0.257

[a] Not detected.

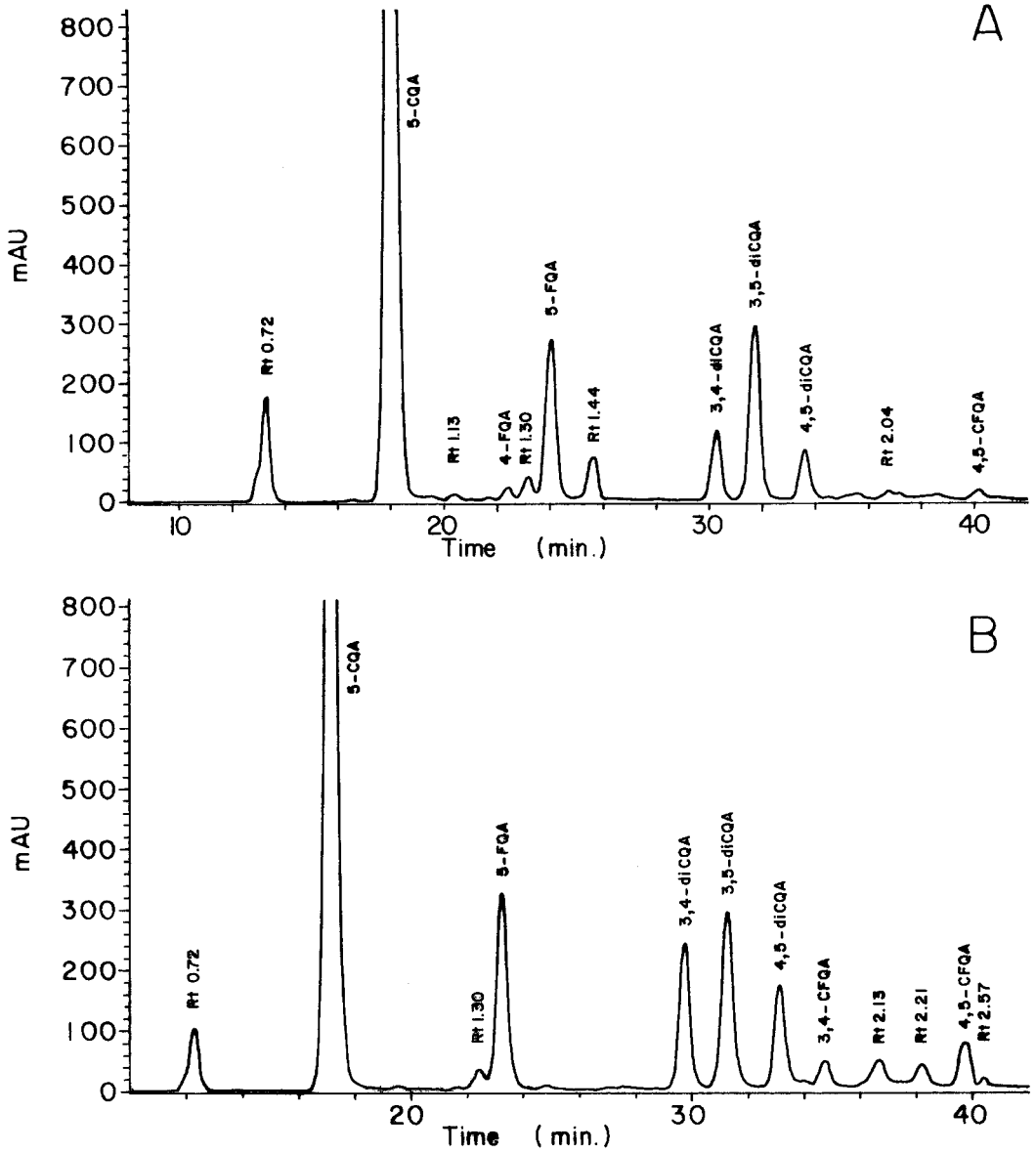

Figure 1. HPLC chromatograms of chlorogenic acids in green coffee beans: (A) Var. Caturra of *C. arabica*; (B) BP46 introduction of *C. canephora*. Conditions are described under Materials and Methods.

2.57, according to Clifford), there are six compounds among which 3,4-CFQA and 4,5-CFQA were identified. The highest contents of these two acids were found in the *C. canephora* accessions. In the other genotypes these compounds were found in very low concentrations or were not detected at all. The major compound in this group of isomers was 4,5-CFQA. In the region corresponding to diCQA (t_R = 1.73−1.92), 3,4-diCQA, 3,5-diCQA, and 4,5-diCQA were identified and were found present in all of the genotypes. 3,5-diCQA is a major component in this region, and it was found in high content in the BP46 and Centro 1 accessions of the *canephora* species. Four compounds were found in the FQA region (t_R = 1.26−1.44), among which 5-FQA and 4-FQA were identified. The major compound in this region is 5-FQA, which is present in all of the genotypes, with its highest concentrations found in introductions BP46 and Centro 1 of *C. canephora*. Finally, three compounds were found in the CQA region (t_R = 0.72−1.13). 5-CQA was identified and is the major compound in all of the coffees studied, with a content ranging between 45 and 75% of the total CGA measured by HPLC. The compound at t_R = 1.13 was not found in introductions BP46 and Centro 1 of *C. canephora*, whereas in introduction BP4 the content was the highest.

The presence or absence of some compounds offers the potential for using them as discriminating compounds in certain cases. In this regard, isomer 4,5-CFQA, present in all of the materials except for the Caturra and Típica varieties, could be used to separate these varieties from the rest of the genotypes. Its presence in Variedad Colombia and in F1 of Caturra × Híbrido de Timor suggests that the compound was derived from the *C. canephora* species through the Híbrido de Timor parent. Likewise, the Típica and Caturra varieties could be distinguished by means of 3,4-CFQA and from the compound with t_R = 1.44, present in the former and absent in the latter. In addition, 4-FQA could be used for differentiating *canephora* coffees from the rest of the genotypes, considering that it was not detected in any of the three accessions of that species.

Híbrido de Timor accessions 832 and 2252 exhibit the same behavior in relation to the presence or absence of

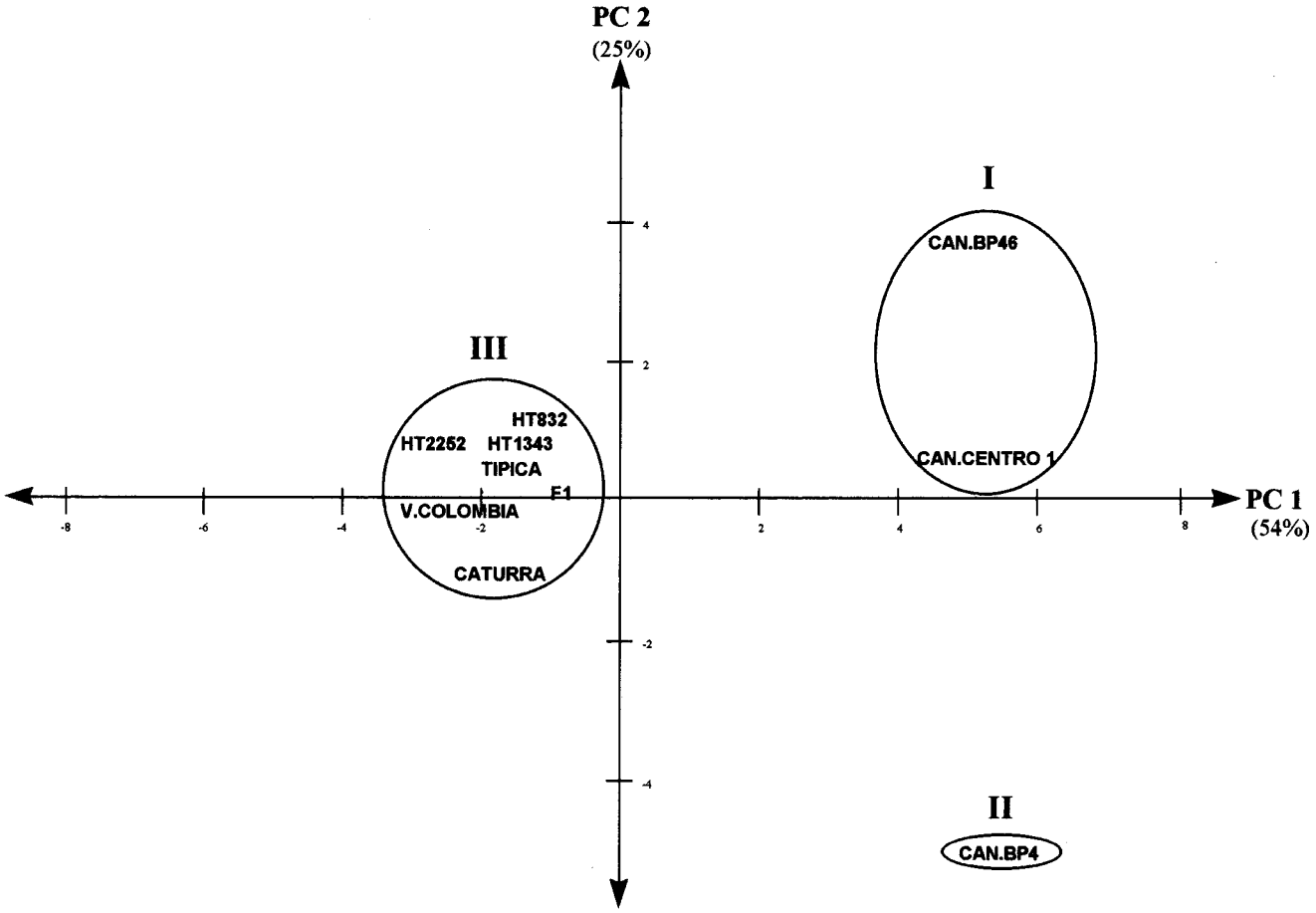

Figure 2. Analysis of the principal components for chlorogenic acids. Groupings are based on hierarchical ascending classification (I−III).

compounds above-mentioned but are different from accession 1343 with regard to two isomers: the compound with t_R = 2.13, absent in Híbrido de Timor accessions 832 and 2252 and present in accession 1343, and 3,4-CFQA, which shows the opposite behavior. The absence of this acid in the F1 offspring of Caturra × Híbrido de Timor 1343 and in Variedad Colombia, derived from this crossing, suggests the potential use of this compound for identifying the Híbrido de Timor accession from which the offspring was obtained.

Multivariate Analysis of Chlorogenic Acids. Because the biggest differences in the chromatographic profiles of the chlorogenic acids were found between Variedad Colombia and *C. canephora* Centro 1 genotypes, only their corresponding HPLC chromatograms are illustrated in Figure 1.

PCA summarized the data for the 16 peaks in three principal components or axes accounting for 89% of the variation. Figure 2 shows genotype distributions on the 1−2 plane. Axis 1 discriminates mainly on the basis of the content of chlorogenic acids 5-FQ, 3,4-diCQ, 4,5-diCQ, 3,4-CFQA, and 4,5-CFQ and the compounds with t_R = 2.13 and 2.21, all of which are positively associated with this axis. 4-FQA and the compounds of t_R = 1.44 and 2.04 participate with a negative relation in this same axis. 5-CQA and the compounds of t_R = 0.72 and 1.30 have an important participation in the makeup of axis 2 with a positive gradient. The compounds of t_R = 1.13 and 2.57 contribute with a negative gradient.

On the plane, along axis 1, *C. canephora* Bp46, Centro 1, and BP4 accessions are found at one end and Híbrido de Timor accessions, Caturra, Típica, and Colombia varieties, and Caturra × Híbrido de Timor F1 are found at the other end. Axis 2 principally separated the BP4 accessions of *C. canephora* from the other accessions.

Table 3. Relative Content of Chlorogenic Acids (Percent) from Each of the Groups Formed in the Multivariate Analysis (Data Correspond to Average Content)[a]

compound	t_R	group I	group II	group III
unknown	0.72	49.77	12.35	37.88
5-CQA	1.00	42.78	18.57	38.65
unknown	1.13	0.00	77.71	22.29
4-FQA	1.26	0.00	0.00	100
unknown	1.30	48.52	21.64	29.83
5-FQA	1.35	50.24	26.71	23.06
unknown	1.44	33.04	0.00	66.96
3,4-diCQA	1.73	50.64	32.47	16.89
3,5-diCQA	1.82	42.69	27.30	30.01
4,5-diCQA	1.92	50.04	31.03	18.93
3,4-CFQA	2.03	46.89	48.59	4.52
unknown	2.04	0.00	0.00	100
unknown	2.13	49.01	49.42	1.57
unknown	2.21	51.29	45.93	2.78
4,5-CFQA	2.30	52.12	44.04	3.85
unknown	2.57	28.29	71.71	0.00

[a] t_R = retention time relative to 5-caffeoylquinic acid.

Genotype classification was done by means of an HAC, and the groups are shown in Figure 2. This grouping was confirmed by means of a DFA, which indicated that the genotypes were classified correctly in 100% of cases.

Table 3 shows the content of chlorogenic acids in each of the formed groups in order to characterize them. Group I, consisting of Centro 1 and BP46 canephora accessions, was characterized by the highest content of

5-CQA, 5-FQA, 3,4-diCQA, 3,5-diCQA, 4,5-diCQA, 4,5-CFQA, and the compounds of $t_R = 0.72$, 1.30, and 2.21. 4-FQA and the compounds of $t_R = 1.13$ and 2.04 were not detected.

Group II (*C. canephora* BP4) had the highest content of the compounds of $t_R = 1.13$ and 2.57. 4-FQA and the compounds of $t_R = 1.44$ and 2.04 were not detected.

Finally, group III, formed by the Caturra and Típica varieties of *C. arabica*, Variedad Colombia, F1 of Caturra × Híbrido de Timor, and the three Híbrido de Timor accessions, had the highest content of 4-FQA and of the compounds of $t_R = 1.44$ and 2.04 and the lowest content of 3,4-diCQA, 4,5-diCQA, 3,4-CFQA, and 4,5-CFQA and of the compounds of $t_R = 2.13$ and 2.21. The compound with $t_R = 2.57$ was not detected.

In conclusion, it was possible to discriminate the genotypes studied by means of the chlorogenic acids, using total content, qualitative differences (presence or absence), and quantitative differences assessed by means of multivariate analysis techniques. This discrimination can be used in genotype selection because it allows the correlation of differences or similarities with the parental line.

ACKNOWLEDGMENT

We thank Dr. B. Guyot (CIRAD-CP) for technical assistance and Cenicafé for providing the laboratories and study materials.

LITERATURE CITED

(1) 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.; Avi Publishing: Westport, CT, 1985; Chapter 13.

(2) Ferreira, L.; Vilar, H.; Fragoso, M.; Aguiar, M.; Cruz, M.; Gonçalves, M. Subsidios paru a caracteriçacao do gran do café de Híbrido de Timor. *Proceedings of the 5th International Scientific Colloquium on Coffee*, Lisbon, Portugal; ASIC: Paris, France, 1971; pp 128−147.

(3) Clifford, M. N.; Williams, T.; Bridson, D. Chlorogenic acids and caffeine as possible taxonomic criteria in *Coffea* and *Psilanthus*. *Phytochemistry* **1989**, *28*, 829−838.

(4) Clifford, M. N.; Kazi, T. The influence of coffee bean maturity on the content of chlorogenic acids, caffeine and trigonelline. *Food Chem*. **1987**, *26*, 59−69.

(5) Clifford, M. N.; Ohiokpehai, O. The influence of extraction method and analytical method on the chlorogenic acid content of green coffee beans. *Proceedings of the 11th International Scientific Colloquium on Coffee*, Lomé, Togo; ASIC: Paris, France, 1985; pp 252−261.

(6) Clifford, M. N.; Jarvis, T. The chlorogenic acids content of green Robusta coffee beans as a possible index of geographic origin. *Food Chem*. **1988**, *29*, 291−298.

(7) Bicchi, C. P.; Binello, A. E.; Pellegrino, G. M.; Vianni, 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.

(8) Castillo, J.; Moreno, G. *La Variedad Colombia Selección de un Cultivar Compuesto Resistente a la Roya del Cafeto*; Federación Nacional de Cafeteros de Colombia: Cenicafé, Chinchiná, Colombia, 1988.

(9) Rakotomalala, J. Diversité biochemique des caféiers, particularités des caféiers sauvages de la région malgache. Ph.D. Dissertation, The University of Montpellier II, Montpellier, France, 1992.

(10) Lehman, H. G.; Martinod, P.; Luzuriaga, O. Methoden zur quantitativen der chlorogensaure. *Dtsch. Lebensm.-Rundsch*. **1967**, *63* (5), 144−151.

(11) Clifford, M. N.; Kellard, B.; Birch, G. Characterization of caffeoylferuloylquinic acids by simultaneous isomerisation and transesterification with tetramethylammonium hydroxide. *Food Chem*. **1989**, *34*, 81−88.

(12) Clifford, M. N.; Kellard, B.; Birch, G. Characterization of chlorogenic acids by simultaneous isomerisation and transesterification with tetramethylammonium hydroxide. *Food Chem*. **1989**, *34*, 115−123.

(13) National Bureau of Standards. *U.S. EPA/NIH, Mass Spectral Database*; U.S. GPO: Washington, DC, 1978.

(14) Macrae, R. Phenolic compounds. In *Coffee. Vol. 1, Chemistry*; Clarke, R. J., Macrae, R., Eds.; Elsevier Applied Science: London, U.K., 1985.

(15) Moreno, R. G. Etude du polimorphysme de l'hybride de Timor en vue de l'amélioration du cafeier arabica: variabilité enzymatique et agronomique dans les populations d'origine; résistance incompléte à *Hemileia vastatrix* dans les croisements avec *Coffea arabica*. Ph.D. Dissertation, The Superior National School of Agronomy of Montpellier, Montpellier, France, 1989.

Received for review October 24, 2000. Revised manuscript received February 2, 2001. Accepted February 6, 2001. We thank Colciencias for financial support.

JF001286U