ELSEVIER

Food Chemistry 75 (2001) 223–230

Food Chemistry

www.elsevier.com/locate/foodchem

Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild *Coffea arabica* L. and *C. canephora* P. accessions

C.-L. Ky[a], J. Louarn[b], S. Dussert[a], B. Guyot[c], S. Hamon[a], M. Noirot[a,*]

[a]*Centre IRD, 911 Avenue Agropolis, BP 5045, 34032 Montpellier Cedex 1, France*
[b]*Station IRD, BP 434, Man, Côte d'Ivoire, France*
[c]*CIRAD-CP, BP 5035, 34032 Montpellier Cedex 1, France*

Received 3 October 2000; received in revised form 1 April 2001; accepted 11 April 2001

Abstract

Numerous aroma precursor evaluations have been undertaken with green coffee beans of both species of worldwide economic importance: *Coffea arabica* L. and *Coffea canephora* P. Efforts have been made to characterise cultivars of these two species. The originality of this study is to present the biochemical diversity of wild accessions originating from Ethiopia and Kenya for *C. arabica* (38 genotypes) and from five African countries (Côte d'Ivoire, Guinea, Congo, Cameroon and Central African Republik) for *C. canephora* (38 genotypes). The biochemical aroma parameters assessed by HPLC analysis were: (1) the two alkaloids, caffeine and trigonelline, (2) chlorogenic acids and (3) sucrose. Results reveal that the two species showed significant accession differences for all compounds. Between-species-average-content comparison confirms that *C. arabica* showed more trigonelline and sucrose and that *C. canephora* presented more CGA and caffeine. *C. canephora* diversity was higher than that of *C. arabica*, except for trigonelline and sucrose. For *C. canephora*, results showed that: (1) no differences were highlighted between accessions for countries of origin for the alkaloids and sucrose, and (2) the 3-CQA content allowed to accessions to be pooled into two groups. Published by Elsevier Science Ltd.

Keywords: Coffeae; *Coffea arabica* L; *Coffea canephora* P.; Biochemical diversity; Caffeine; Chlorogenic acids; Sucrose; Trigonelline; HPLC

1. Introduction

The two coffee tree species of worldwide importance, i.e. *Coffea arabica* ($2n=44$) and *Coffea canephora* ($2n=22$) (commonly known as Robusta), are adapted to very different ecological environments: (i) tropical highlands (>600 m), mainly in Latin America, East Africa and India, for the former; and (ii) lowlands of West Africa, Indonesia, Vietnam and Brazil for the latter. Arabica coffee, with its lower bitterness and better flavour, is more appreciated by consumers and it costs twice as much as Robusta. Consequently, improving Robusta cup quality could increase lowland farmer incomes.

Aroma formation is a very complex process, including Maillard and Strecker's reactions and thermal degradation during roasting (De Maria, Trugo, Moreira, & Werneck, 1994). Some aroma precursors, such as sucrose and trigonelline, give rise to appreciated flavour products, including furans, pyrazine, alkyl-pyridines and pyrroles (Clifford, 1985a; Dart & Nursten, 1985; De Maria, Trugo, Aquino Neto, Moreira, & Alviano, 1996; Feldman, Ryder, & Kung, 1969; Thaler & Arneth, 1967). Other precursors, such as chlorogenic acids (CGA) and caffeine, increase bitterness, the former after degradation into phenol derivatives (Leloup, Louvrier, & Liardon, 1995) and the latter without any degradation (Voilley, Sauvageot, & Durand, 1977). Enhancing Robusta cup quality would thus imply increasing sucrose and trigonelline contents while decreasing CGA and caffeine contents. However, is the within-species diversity wide enough to permit such improvement?

Several past studies have been carried out to evaluate and compare sucrose, CGA, trigonelline and caffeine contents in the two cultivated species, but they were limited to cultivars. In the 1970s, the genetic resources of coffee trees were sampled in the centres of origin of the genus (Africa, Madagascar and Mascarene islands).

* Corresponding author. Fax: +33-4-6741-6222.
E-mail address: noirot@mpl.ird.fr (M. Noirot).

0308-8146/01/$ - see front matter Published by Elsevier Science Ltd.
PII: S0308-8146(01)00204-7

These wild *C. arabica* and *C. canephora* accessions have been evaluated for their agro-morphological traits (Bouharmont, 1978; Capot, 1978; Charrier, 1978; Berthaud, 1986; Montagnon, Leroy, & Yapo, 1991; Leroy, Montagnon, Charrier, & Eskes, 1993), and analysed with isozyme markers (Berthou & Trouslot, 1977; Berthaud, 1986; Montagnon et al., 1991) and molecular markers (Lashermes, Combes, Robert, Trouslot, D'Hont, Anthony & Charrier, 1999; Lashermes, Combes, Trouslot, Anthony, & Charrier, 1996). In particular, for *C. canephora*, RFLP (restriction fragment length polymorphism) analysis helped to distinguish five genetic groups (A–E) (Dussert, Lashermes, Anthony, Montagnon, Trouslot, Combes, Berthaud, Noirot, & Hamon, 1999). In contrast, the biochemical diversity of these coffee genetic resources has yet to be assessed.

In this paper, we present a biochemical evaluation of sucrose, CGA, trigonelline and caffeine variations in wild genetic resources of *C. arabica* and *C. canephora*. We also compare: (i) the two species in terms of their diversity range and mean values, and (ii) their biochemical diversity to their agro-morphological, isozymic and molecular diversity.

2. Material and methods

2.1. Plant material

2.1.1. C. arabica *genetic resources*

Thirty-eight of *C. arabica* accessions collected by IRD[1] were evaluated. All but two were native to Ethiopia (Guillaumet & Hallé, 1967); the others were from Marsabit mountain in Kenya (Berthaud, Anthony, & Lourd, 1983). This collection was maintained at the Tonkoui mountain Station (Côte d'Ivoire).

2.1.2. C. canephora *genetic resources*

For *C. canephora,* collecting trips were carried out in five African countries between 1975 and 1988: Côte d'Ivoire, Guinea, Congo, Cameroon and Central African Republic in collaboration with CIRAD,[2] IPGRI[3] and FAO[4] (Anthony, Couturon, & De Namur, 1985; Berthaud & Guillaumet, 1978; De Namur, Couturon, Sita, & Anthony, 1988; Le Pierres et al., 1989). Genetic resources are maintained in Côte d'Ivoire (CNRA, Divo station). Eight accessions per country were used for the biochemical evaluations, but only six from the Central African Republic were assessed.

[1] Institut de Recherche pour le Développement, formerly named ORSTOM

[2] Centre de coopération Internationale en Recherche Agronomique pour le Développement

[3] International Plant Genetic Ressources Institute, formerly IBPGR

[4] Food and Agricultural Organization

2.2. Sample preparation

Coffee cherries were harvested at full maturity and depulped using the wet processing method. After desiccation on silicagel, 50 green beans per tree were frozen in liquid nitrogen before crushing in a ball mill (Dangoumill) for 2 min. The fine powder was split into six samples, three to estimate the water content and three for extraction and analysis.

2.3. Extraction and purification

Different extraction and purification methods were used, according to the compounds:

i. Caffeine and trigonelline: slightly modified methods of Trugo, Macrae and Dick (1983) and Barre, Akaffou, Louarn, Charrier, Hamon, and Noirot (1998) were used: each 50-mg powder sample was placed in a 50 ml capped tube (Sarstedt) with 500 mg of magnesium oxide (Merck) and 25 ml of distilled water. Tubes were heated for 20 min at 105 °C in an autoclave. Extracts were filtered (0.2 μm) and directly analysed by HPLC;
ii. Chlorogenic acids: the method was described in Ky, Noirot, and Hamon (1997) and Ky, Louarn, Guyot, Charrier, Hamon, and Noirot (1999);
iii. Sucrose: the method was described in Ky et al. (2000).

2.4. Analytical HPLC

Three different elution programs were used to analyse alkaloids, chlorogenic acids and sucrose contents in green beans.

For caffeine, trigonelline and CGA, chromatography was carried out in a system consisting of (i) two Waters Associates Model 510 pumping units, (ii) an automated sample injector (Waters 717 plus autosampler), (iii) a variable-wavelength UV detector (Waters 996 Photodiode Array Detector), (iv) a C_{18} pre-column, and (v) a 250×4 mm Merck Superspher 100 RP 18 column (5 μm particle size). Sucrose was analysed using anion-exchange chromatography coupled to pulsed amperometric detection (Ky et al., 2000).

The elution programme for chlorogenic acids was described in Ky et al. (1997). Two solvents were used for co-elution of caffeine and trigonelline: (i) an aqueous solution containing triethylamine and acetic acid (pH 5.3) (solvent A); and (ii) methanol (HPLC grade) (solvent B). These two mobile phases were filtered (0.2 μm), degassed and sonicated (Ney, 300 ultrasonik) before use. Samples and standard (10 μl) were analysed at room temperature using the following elution programme: A–B mixture (70/30) of linear gradient for 15 min. The flow rate was 1.0 ml/min. UV detection was

carried out at 272.8 and 263.3 nm wavelength, corresponding to caffeine and trigonelline maximum absorption, respectively. Quantification was achieved by peak-area measurement and comparison with standards (Sigma Chemical Co). The calibration curve was generated using three replicate points of a caffeine and trigonelline mixture at 10, 20, 30 and 40 mg/l.

The processing order was fully randomized. Every 10 extracts, a control was used to verify the stability of the measurements. The content unit was the percentage dry matter basis (% dmb).

2.5. Statistical analysis

All results were analysed using the Statistica software package (version 5.1, 1997 for Microsoft Windows).

Within each species, between-tree differences were tested using a one-way ANOVA. For *C. canephora*, between-group comparisons were also carried out using a one-way ANOVA.

Between species variance comparisons were processed using the F test ($\sigma^2_{max}/\sigma^2_{min}$), whereas mean comparisons required the nonparametric Kruskall–Wallis test.

2.6. Variables

Chlorogenic acids are presented according to the IUPAC (1976) numbering system. Abbreviations were proposed by Clifford (Clifford, 1985a, 1985b, 1999, 2000). The variables are:

3-CQA: 3-caffeoylquinic acid;
4-CQA: 4-caffeoylquinic acid;
5-CQA: 5-caffeoylquinic acid;
3,4-diCQA: 3,4-dicaffeoylquinic acid;
3,5-diCQA: 3,5-dicaffeoylquinic acid;
4,5-diCQA: 4,5-dicaffeoylquinic acid;
3-FQA: 3-feruloylquinic acid;
4-FQA: 4-feruloylquinic acid;
5-FQA: 5-feruloylquinic acid;
CQA: caffeoylquinic acids (3-CQA, 4-CQA and 5-CQA);
diCQA: dicaffeoylquinic acids (3,4-diCQA, 3,5-diCQA and 4,5-diCQA);
FQA: feruloylquinic acids (3-FQA, 4-FQA and 5-FQA); and
CGA: chlorogenic acids (CQA, diCQA and FQA).

We added two abbreviations: (i) *CGAs.s.* (total CQA; Clifford 1985b) for quinic acids with a caffeate unit (CQA + diCQA); and (ii) 4-and5-CQA for the two CQA isomers (4-CQA and 5-CQA) for which chromatogram peaks are not separable.

N.B.: Other cinnamoylquinic acid, such as *p*-coumaoylquinic acids (*p*CoQA), caffeoylferuloylquinic acids (CFQA), feruloylcaffeoylquinic acids (FCQA), and CGA like components, where the caffeic unit is coupled with an amino acid, such as tyrosine, are also present in trace amounts (Clifford, Kellard, & Ah-Sing, 1989; Correia, Leitao, & Clifford, 1995; Morishita et al. 1987; Murata, Okada, & Homma, 1995, 1996) and were not assessed in the present study.

2.6.1. Relative variables

Their abbreviations begin with %. For each isomer, the relative part is the proportion of the isomer in CGA content (for example, %3-CQA = 3-CQA/CGA). For CQA, the relative part is the proportion of CQA content in CGA*s.s.*, while %*CGAs.s.* is the part of *CGAs.s.* in CGA.

2.6.2. Composite variable

This only concerns the CQA content/caffeine content ratio (CQA/CAF).

3. Results

3.1. C. arabica

3.1.1. Overall contents

In *C. arabica*, the sum of caffeine, trigonelline, CGA and sucrose contents represented 16% dmb on average (Table 1).

3.1.2. Sucrose. Sucrose was the main compound with 9.2% dmb on average. Between accession-differences

Table 1
Contents (% dmb) and ANOVA results for biochemical traits in *Coffea arabica*. All F tests were very highly significant ($P < 0.001$)

Trait	Mean	Min	Max	Max/Min	$F_{38,117}$
Sucrose	9.23	7.40	11.1	1.50	40.1
3-CQA	0.20	0.16	0.25	1.56	17.3
5- and 4-CQA	3.06	2.43	3.72	1.53	7.98
CQA	3.26	2.61	3.97	1.52	7.46
5-FQA	0.15	0.08	0.22	2.75	47.5
FQA	0.19	0.12	0.27	2.25	34.9
3,4-diCQA	0.12	0.08	0.17	2.12	30.3
3,5-diCQA	0.23	0.17	0.36	2.11	31.9
4,5-diCQA	0.25	0.18	0.38	2.11	29.5
diCQA	0.60	0.43	0.88	2.04	27.9
CGA*s.s.*	3.9	3.2	4.6	1.43	4.98
CGA	4.1	3.4	4.8	1.41	4.48
%CQA	84.3	76.7	88.1	1.14	247
%CGA*s.s.*	95.1	89.1	96.9	1.08	2.15
%3-CQA	5.0	4.0	6.4	1.60	111
%5-CQA	75.3	67.7	80.3	1.18	16.7
%5-FQA	3.7	2.1	5.4	2.57	234
%3,4-diCQA	2.9	1.9	4.6	2.42	171
%3,5-diCQA	5.7	4.0	8.6	2.15	164
%4,5-diCQA	6.2	4.6	9.3	2.02	155
Caffeine	1.22	0.96	1.62	1.68	10.5
Trigonelline	1.19	0.88	1.77	2.01	20.9
CAF/CQA	0.38	0.29	0.52	1.79	

were significant and contents ranged from 7.4 to 11.1% dmb. The max/min ratio was 1.5.

3.1.3. Chlorogenic acids. Chlorogenic acids constituted the second group of compounds with 4.1% dmb CQA. On average, diCQA and FQA corresponded to 80, 15 and 5% of the total CGA content, respectively. Thereafter, two isomers (3-FQA and 4-FQA) were not taken in account because of their low CGA content (< 2%). All other isomers showed significant differences between accessions for both absolute and relative (%CQA,%*CGAs.s.*, etc.) contents. The max/min ratio was 1.4 for CGA and 1.5, 2.0 and 2.3 for CQA, diCQA and FQA, respectively. It did not vary between isomers within a CGA class. For relative contents, the index ranged from 1.1 (%CGA*s.s*) to 2.6 (%5-FQA).

Table 2
Contents (% dmb) and ANOVA results for biochemical traits in *Coffea canephora*

Trait	Mean	Min	Max	Max/Min	$F_{36,111}$
3-CQA	0.81	0.43	1.33	3.09	66.1
5- and 4-CQA	6.85	4.68	8.96	1.91	9.49
CQA	7.66	5.12	9.50	1.86	7.88
5-FQA	1.17	0.62	2.04	3.29	37.2
FQA	1.43	0.77	2.23	2.90	33.8
3,4-diCQA	0.77	0.54	0.95	1.76	11.9
3,5-diCQA	0.71	0.38	1.13	2.97	39.8
4,5-diCQA	0.83	0.41	1.24	3.02	22.6
diCQA	2.31	1.57	3.03	1.93	15.2
CGA*s.s.*	9.96	6.81	12.21	1.79	6.90
CGA	11.3	7.88	14.4	1.83	7.61
%CQA	76.8	68.5	84.2	1.23	192
%CGA*s.s.*	87.5	82.6	93.1	1.13	187
%3-CQA	7.11	3.71	11.4	3.07	2720
%5- and 4-CQA	60.1	54.0	66.8	1.24	304
%5-FQA	10.2	6.06	14.7	2.43	166
%3,4-diCQA	6.75	4.66	9.19	1.97	330
%3,5-diCQA	6.26	3.35	8.86	2.64	1273
%4,5-diCQA	7.30	4.67	11.9	2.55	58.3
Sucrose	5.45	4.05	7.05	1.74	25.6
Caffeine	2.54	1.51	3.33	2.21	16.7
Trigonelline	1.01	0.75	1.24	1.65	9.90
CAF/CQA	0.34	0.21	0.59	2.81	

3.1.4. Alkaloids. Caffeine and trigonelline contents were similar with 1.2% dmb on average. Between-accession differences were significant for both alkaloids. Contents ranged from 0.96 to 1.62% for caffeine and from 0.88 and 1.77% dmb for trigonelline. The CAF/CQA ratio ranged from 0.29 to 0.52 (mean 0.38). The max/min ratios were 1.7, 1.8 and 2.0 for the caffeine content, the CAF/CQA ratio and the trigonelline content, respectively.

3.2. C. canephora

3.2.1. Overall contents

The sum of caffeine, trigonelline, CGA and sucrose contents represented 20% dmb on average (Table 2).

3.2.2. Chlorogenic acids

In this species, chlorogenic acids constituted the main group with 11.3% dmb, on average. CQA, diCQA and FQA represented 67, 20 and 13% of the total CGA content, respectively. As in *C. arabica*, the 4-FQA and 3-FQA isomers, constituting less than 2% CGA content, were subsequently not taken in account. All other isomers showed significant between-accession differences for both absolute and relative contents. Relative contents were more discriminating than absolute contents, especially for %3-CQA and %3,5-diCQA. The Max/ Min ratio index was 1.8 for CGA and 1.9, 1.9 and 2.9 for CQA, diCQA and FQA, respectively. In contrast to *C. arabica,* it varied between isomers within each CGA class: for example, it was 1.8 and 3.0 for 3,4-diCQA and 4,5-diCQA, respectively. For relative contents, the index ranged from 1.1 (%*CGAs.s.*) to 3.1 (%3-CQA).

Between-accession variations were analysed in relation to: (i) their geographic origin, (ii) their genetic groups (determined by RFLP markers), and (iii) their 3-CQA content.

The country of origin was a source of intraspecific variation for 3-CQA, 5-FQA,%*CGAs.s.*, %3-CQA, %5-FQA and %4,5-diCQA (Table 3). The relative importance of this source of variation ranged from 29 to 39%, showing that within-country variations were still

Table 3
Between-origin ANOVA results in *Coffea canephora*[a]

Trait	GUI	CI	CAM	CAR	CON	$F_{4,32}$	P
3-CQA	0.69a	0.68a	1.05b	0.90ab	0.70a	3.74	0.013
5-FQA	1.40a	1.33ab	1.08ab	0.88b	1.10ab	3.16	0.027
CGA*s.s*	86.3a	86.1a	87.2ab	89.8b	88.6ab	3.10	0.029
%3-CQA	5.81a	6.25ab	9.18b	8.20ab	6.29a	3.65	0.015
%5-FQA	11.6a	11.9a	9.52ab	7.97b	9.70ab	4.37	0.006
%4,5-diCQA	6.93a	6.45a	6.83a	9.09b	7.57ab	3.14	0.027

[a] Contents for biochemical traits are expressed in % dmb. The country codes are: GUI, Guinea; CI, Côte d'Ivoire; CAM, Cameroon; CAR, Central African Republic and CON, Congo.

very high. The first five traits showing between-country variations (3–CQA-%5-FQA) differed in the western (Guinea and Côte-d'Ivoire) and eastern parts of the distribution area (Central Africa Republic and Cameroon): western countries showed lower values for 3-CQA,%3-CQA and %CGA*s.s* and higher values for 5-FQA and %5-FQA. Central African Republic differed from the other countries in terms of the latter trait %4,5-diCQA.

Genetic groups, based on RFLP analysis (Dussert et al., 1999), showed differences only for 4,5-diCQA ($F_{4,24}=4.84$; $P=0.005$) and %4,5-diCQA ($F_{4,24}=5.20$; $P=0.004$), with the B group differing from the C and D groups. The relative importance of the between-group variance was close to 45%. For other traits, the within-group variance represented 100% of the total variance.

The 3-CQA content allowed us to pool accessions into two groups: CGA-1 and CGA-2 (Fig. 1). CGA-1 and CGA-2 groups mainly differed in terms of their 3-CQA content, with 0.61 and 1.11% dmb, respectively (Table 4). The CGA-1 group was also characterized by lower %CQA and %3-CQA, and higher 3,5-diCQA, %5CQA and %3,5-diCQA.

3.2.3. Sucrose

Sucrose content was 5.5% dmb on average and showed significant between-accession differences (Table 2). There were no sucrose content differences between: (i) countries of origin ($F_{4,32}=1.16$; $P=0.35$); (ii) genetic groups ($F_{4,24}=0.86$; $P=0.50$) and (3) CGA-1 and CGA-2 groups ($F_{1,35}=2.72$; $P=0.11$). The max/min ratio was 1.74.

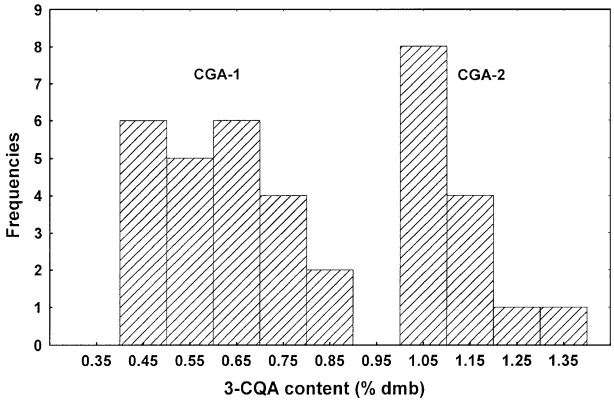

Fig. 1. Bimodal distribution of 3-CQA contents.

Table 4
Significant differences between CGA-1 and CGA-2 groups in *Coffea canephora*, defined by their 3-CQA contents[a]

Trait	Group CGA-1	Group CGA-2	$F_{1,35}$	P
3-CQA	0.61	1.11	169	<0.001
3,5-diCQA	0.80	0.56	20.9	<0.001
%CQA	76.0	78.0	4.20	0.048
%3-CQA	5.54	9.69	89.1	<0.001
%5-CQA	61.2	58.2	7.70	0.009
%3,5-diCQA	7.12	4.86	46.3	<0.001

[a] Degrees of freedom were d.f.$_1=1$ and d.f.$_2=36$ for CGA and d.f.$_1=1$ and d.f.$_2=34$ for amino-acids.

3.2.4. Alkaloids

Caffeine and trigonelline contents were 2.5% dmb and 1.0% dmb on average, respectively. The two alkaloids showed significant between-accession differences (Table 5). As for sucrose content, no differences were highlighted between: (i) countries of origin (caffeine $F_{4,32}=0.42$; $P=0.79$; trigonelline $F_{4,32}=2.57$; $P=0.06$); (ii) genetic groups (caffeine $F_{4,24}=0.60$; $P=0.67$; trigonelline $F_{4,24}=1.29$; $P=0.30$) and (iii) CGA-1 and CGA-2 groups (caffeine: $F_{1,35}=2.09$; $P=0.15$; trigonelline: $F_{1,35}=0.64$; $P=0.43$). The CAF/CQA ratio ranged from 0.21 to 0.59 (mean 0.34).

The max/min ratios were 2.2, 2.8 and 1.6 for the caffeine content, CAF/CQA ratio and trigonelline content, respectively.

3.3. Between-species comparison

3.3.1. Variance comparison

C. canephora diversity was higher for all absolute contents, except trigonelline and sucrose (*C. arabica* sucrose diversity was higher). It was also higher for all relative contents, except %CQA and %4- and 5-CQA. The variance ratio was especially high for 3-CQA and 5-FQA (>100×; Table 5) and was 10–23-fold higher in *C. canephora* for the other compounds. In all cases, the variance ratio was higher for contents than for relative contents.

Table 5
Between-species differences in diversity

Trait	Variance *Coffea arabica*	Variance *Coffea canephora*	$F_{37,37}$	P
3-CQA	0.000603	0.073506	122	<0.001
4-and 5-CQA	0.059795	0.784545	13.1	<0.001
5-FQA	0.001182	0.118766	100	<0.001
3,4-diCQA	0.000475	0.012741	26.8	<0.001
3,5-diCQA	0.001734	0.038369	22.1	<0.001
4,5-diCQA	0.002262	0.041491	18.3	<0.001
CQA	0.063703	0.802612	12.6	<0.001
FQA	0.001479	0.142241	96.1	<0.001
diCQA	0.011025	0.151556	13.7	<0.001
CGA	0.072575	1.704695	23.5	<0.001
CGA*s.s*	0.060681	1.190674	19.6	<0.001
%CQA	7.776089	9.281518	1.19	0.29
%CGA*s.s*	1.795817	6.777209	3.77	<0.001
%3-CQA	0.319205	5.804806	18.2	<0.001
%4-and5-CQA	7.868435	12.405502	1.58	0.08
%5-FQA	0.671280	5.559940	8.28	<0.001
%4,5-diCQA	0.324901	0.893529	2.75	0.001
%3,5-diCQA	1.083151	2.186997	2.02	0.017
%4,5-diCQA	1.484140	2.734469	1.84	0.032
Sucrose	0.946919	0.665679	1.42	0.015
Caffeine	0.019332	0.198145	10.2	<0.001
Trigonelline	0.025995	0.022062	1.18	0.31

3.3.2. Mean comparison

Differences were noted between species for all traits, except %3,5-diCQA (Table 6). Sucrose and trigonelline contents were 1.69- and 1.18-fold higher in *C. arabica* than in *C. canephora,* respectively. The relative parts of 5-CQA in CQA, CQA in CGA and CGA*s.s* in CGA were also higher in *C. arabica*. By contrast, CGA isomer contents and caffeine contents were higher in *C. canephora*. The "*C. canephora*/*C. arabica*" ratio ranged from 2.1 to 7.8 for caffeine and 5-FQA contents, respectively.

4. Discussion

4.1. Overall

This study is the first evaluation of biochemical compounds involving wild accessions of *C. arabica* and *C. canephora*. The discussion focuses on: (i) the respective importance of their biochemical polymorphism, and (ii) comparison with other types of diversity (isozymes, molecular markers, morphological traits). Finally, the breeding implications are discussed.

4.2. C. arabica

C. arabica is characterised by a very limited diversity in terms of isozymes and molecular markers, which could be explained by its recent allotetraploid origin (Lashermes, et al., 1999). In contrast, high polymorphism has been observed for agro-morphological traits (Charrier, 1978) and, in the present study, for biochemical compounds. The phenotypic diversity (morphological), which includes environmental variance, was obviously higher than the genotypic diversity (molecular). Nevertheless, the high heritabilities *s.l.* of most biochemical compounds (Barre et al., 1998; Ky et al., 1999, 2000) clearly indicates that environmental effects cannot explain the discrepancy between molecular markers and biochemical traits. We suggest that the contrast could result from few gene mutations with marked phenotypic effects.

Between-variable differences were recorded for their diversity when comparing %CGA*s.s*. with 5-FQA in terms of their max/min ratios. Since part of the range could be due to variations in environmental effects, a comparison of biochemical traits showing high heritabilities (h^2*s.l.* > 90%) would be more valid. The max/min ratio range was thus still high, ranging from 1.1 (%CGA*s.s*) to 2.0 (trigonelline). The outstandingly low max/min ratio for %CGA*s.s*, %CQA and %5-CQA could mean that the regulation processes were less variable than the accumulation.

The interest in wild *C. arabica* accessions for breeding, especially in terms of agronomical traits (Charrier, 1978), was therefore confirmed for biochemical traits. Since quantitative inheritance parameters for caffeine, CGA, sucrose and trigonelline contents (additivity and heritability *s.l.*) are known in coffee trees (Barre et al., 1998; Ky et al., 1999, 2000; Ky, Guyot, Louarn, Hamon, & Noirot, 2001), there is considerable potential for biochemical improvement in *C. arabica*.

Table 6
Between-species mean comparisons using the Kruskall–Wallis test

Trait	*Coffea arabica*	*Coffea canephora*	*P*
Sucrose	9.23	5.45	<0.001
3-CQA	0.20	0.81	<0.001
5- and 4-CQA	3.06	6.85	<0.001
CQA	3.26	7.66	<0.001
5-FQA	0.15	1.17	<0.001
FQA	0.19	1.43	<0.001
3,4-diCQA	0.12	0.77	<0.001
3,5-diCQA	0.23	0.71	<0.001
4,5-diCQA	0.25	0.83	<0.001
diCQA	0.6	2.31	<0.001
CGA*s.s*	3.9	9.96	<0.001
CGA	4.1	11.3	<0.001
%CQA	84.3	76.8	<0.001
%CGA*s.s*	95.1	87.5	<0.001
%3-CQA	5	7.11	<0.001
%5-CQA	75.3	60.1	<0.001
%5-FQA	3.7	10.2	<0.001
%3,4-diCQA	2.9	6.75	<0.001
%3,5-diCQA	5.7	6.26	0.25
%4,5-diCQA	6.2	7.3	<0.001
Caffeine	1.22	2.54	<0.001
Trigonelline	1.19	1.01	0.006

4.3. C. canephora

4.3.1. Biochemical and molecular diversity

For this species, there is high isozyme and RFLP marker diversity: the nine isozyme loci studied by Berthaud (1986) were polymorphic with 3.5 alleles per locus on average and 77% of RFLP loci studied were also polymorphic, with five alleles per locus on average (Dussert et al., 1999). This high diversity could be explained by the ancient origin of *C. canephora* (−20 MaBP), its reproductive allogamy and the absence of strong genetic bottlenecks (Cros, 1994). There is also high polymorphism in agro-morphological traits (Leroy et al., 1993) and in biochemical compounds, as confirmed here. Note that caffeine content diversity of another wild accession sample was formerly evaluated by Leroy et al. (1993) and the results were very similar to those obtained in this study.

As for *C arabica,* between-variable differences were recorded for biochemical diversity. For highly heritable biochemical traits (h^2*s.l.* > 90%), the max/min ratio ranged from 1.1 (%CGA*s.s*) to 2.2 (caffeine). Again we observed a lower max/min ratio for relative (%CGA*s.s*, %CQA and %5-CQA) than for absolute contents, showing that regulation processes were less variable than the accumulation.

4.3.2. Within-species diversity structure

C. canephora diversity is also characterised by the presence of genetic groups. In particular, two main groups—the Guinean and the Congolese groups—were formerly distinguished using isozyme markers (Berthaud, 1986). The Congolese group was then split into four sub-groups on the basis of RFLP marker analyses (Dussert et al., 1999): A, B, C and E (the D group corresponds to the Guinean group). This group division could result from a recent population isolation during the Pleistocene (Berthaud, 1986; Dussert et al., 1999).

Guinean and Congolese groups differ in rust susceptibility, bean weight, drought tolerance, branching, internode length, leaf shape and size, growth habit and earliness (Berthaud, 1986; Leroy et al., 1993). The relationships were not as clearcut for biochemical traits and concerned only differences for the 4,5-diCQA and %4,5-diCQA, thus contrasting the Guinean group from the C Congolese sub-group. Surprisingly, the C subgroup is the Congolese sub-group which is genetically the closest to the Guinean group (Dussert et al. 1999).

In contrast, many biochemical traits were influenced by the geographical origin of accessions. This clearly shows that the adaptation of populations, i.e. the fitness, to local environments, led to biochemical differentiation that was independent of the neutral RFLP marker evolution. This also explains why the Guinean group differs markedly from the C Congolese sub-group for 4,5-diCQA, despite its relative genetic RFLP similarity.

The existence of two groups with respect to 3-CQA isomers, i.e. genotypes native to Côte d'Ivoire contrasting with those from Cameroon, emphasises the biological role of CGA isomers in relation to the fitness in different ecological zones.

5. Conclusions

The high heterosis in Guinean×Congolese hybrids was the first major result that highlighted the interest of wild *C. canephora* accessions for breeding (Berthaud, 1986) and was the starting point of a reciprocal recurrent selection (Leroy et al., 1993). The current evaluation of biochemical polymorphism and previous inheritance results (Barre et al., 1998; Ky et al., 1999, 2000, 2001) clearly show that parents could be bred to decrease CGA and caffeine contents and increase sucrose contents. Nevertheless, the high *C. canephora* polymorphism excludes the possibility of obtaining parents with characteristics similar to those of *C. arabica* accessions.

The use of *C. arabica* as gene donor was suggested as an alternative. Unfortunately, this breeding scheme—called the Arabusta programme (Capot, 1972)—failed, due to partial sterility of tetraploid hybrids linked to the presence of tetravalents at meiosis (Kammacher & Capot, 1972; Grassias, 1980). Nevertheless, the presence of such quality genes in the amphiploid genome of *C. arabica* strongly suggests the presence of these genes in the diploid genepool. A biochemical evaluation of all *Coffea* genetic resources should pinpoint suitable parental species and enable the development of a breeding programme based on introgression at the diploid level.

References

Anthony, F., Couturon, E., & De Namur, C. (1985). Les caféiers sauvages du Cameroun: résultats d'une mission de prospection effectuée par l'ORSTOM en 1983. *Compte-rendus du Onzième Colloque de l'ASIC (Association Scientifique International sur le Café) 11*, 495–501.

Barre, P., Akaffou, S., Louarn, J., Charrier, A., Hamon, S., & Noirot, M. (1998). Inheritance of caffeine and heteroside contents in an interspecific cross between a cultivated coffee species *Coffea liberica* var '*dewevrei*' and a wild species caffeine free *C. pseudozanguebariae*. *Theoretical and Applied Genetics*, *96*, 306–311.

Berthaud, J. (1986). *Les ressources génétiques pour l'amélioration des caféiers africains diploïdes*. Paris: Coll. Trav. & Doc., ORSTOM.

Berthaud, J., Anthony, F., & Lourd, M. (1983). Les caféiers sauvages de Tanzanie. Résultats d'une mission de prospection effectuée du 5 mars au 11 avril 1982. *Café Cacao Thé*, *27*, 245–258.

Berthaud, J., & Guillaumet, J. L. (1978). Les caféiers sauvages en Centrafrique: résultats d'une mission de prospection (janvier–février 1975). *Café Cacao Thé*, *3*, 171–186.

Berthou, E., & Trouslot, P. (1977). L'analyse du polymorphisme enzymatique dans le genre Coffea:adaptation d'une méthode d'électrophorèse en série. Premiers résultats. *ASIC (Colloque Scientifique International sur le Café)*, *8*, 373–383.

Bouharmont, P. (1978). Comportement des collections de *Coffea arabica* implantées au Cameroun. In: Etude de la structure et de la variabilité génétique des caféiers. *IFCC*, 44–54.

Capot, J. (1972). L'amélioration du caféier en Côte d'Ivoire. Les hybrides Arabusta. *Café Cacao Thé*, *14*, 3–18.

Capot, J. (1978). Comportement des collections de *Coffea arabica* en Côte d'Ivoire. A-En basse altitude à Bingerville et à Abengourou. In: *Etude de la structure et de la variabilité génétique des caféiers. IFCC*, 33-43.

Charrier, A. (1978). Analyse de la variabilité phénotypique de la collection de *Coffea arabica* à Madagascar. In: *Etude de la structure et de la variabilité génétique des caféiers. IFCC*, 22–32.

Clifford, M. N. (1985a). Chemical and physical aspects of green coffee and coffee products. In X. Clifford, & X. Wilson (Eds.), *Coffee, botany, biochemistry and production of beans and beverage* (pp. 305–374). Westport, CT: Avi publishing company.

Clifford, M. N. (1985b). Chlorogenic acids. In: Clarke & Macrae (Eds), *Coffee I Chemistry* (pp. 153–202). London, UK: Elsevier Applied Science.

Clifford, M. N. (1999). Chlorogenic acids and other cinnamates—nature, occurrence, dietary burden. *Journal of the Science of Food and Agriculture*, *79*, 362–372.

Clifford, M. N. (2000). Chlorogenic acids and other cinnamates—nature, occurrence, dietary burden, absorption and metabolism. *Journal of Agricultural and Food Chemistry*, *80*, 1033–1042.

Clifford, M. N., Kellard, B., & Ah-Sing, E. (1989). Caffeoyl-tyrosine from green robusta coffee beans. *Phytochemistry*, *28*, 1989–1990.

Correia, M. N. G., Leitao, M. C. A., & Clifford, M. N. (1995). Caffeoyl-tyrosine and Angola II as characteristic markers for Angolan robusta coffees. *Food Chemistry*, *53*, 309–313.

Cros, J. (1994). *Implications phylogénétiques des variations de l'ADN chloroplastique chez les caféiers (genre* Coffea *L. et* Psilanthus *Hook. f.*). Thèse Doct., Univ. Montpellier II. 160.

Dart, S. K., & Nursten, H. E. (1985). Volatile components. In R. J. Clarke, & R. Macrae (Eds.), *Coffee: chemistry (Vol. I)* (pp. 223–265). London, UK: Elsevier Applied Science.

De Maria, C. A. B., Trugo, L. C., Aquino Neto, F. R., Moreira, R. F. A., & Alviano, C. S. (1996). Composition of green coffee water-soluble fractions and identification of volatiles formed during roasting. *Food Chemistry*, *55*, 203–207.

De Maria, C. A. B., Trugo, L. C., Moreira, R. F. A., & Werneck, C. C. (1994). Composition of green coffee fractions and their contribution to the volatile profile formed during roasting. *Food Chemistry*, *50*, 141–145.

De Namur, C., Couturon, E., Sita, P. & Anthony, F. (1988). Résultats d'une mission de prospection des caféiers sauvages du Congo. *ASIC (Colloque Scientifique International sur le Café) 7*, 397-404.

Dussert, S., Lashermes, P., Anthony, F., Montagnon, C., Trouslot, P., Combes, M.-C., Berthaud, J., Noirot, M., & Hamon, S. (1999). Le caféiers, *Coffea canephora*. In P. Hamon, M. Seguin, X. Perrier, & C. Glaszmann (Eds.), *Diversité génétiques des plantes tropicales cultivées* (pp. 175–194) CIRAD, Montpellier.

Feldman, J. R., Ryder, W. S., & Kung, J. T. (1969). Importance of nonvolatile compounds to the flavour of coffee. *Journal of Agricultural and Food Chemistry*, *17*, 733–739.

Grassias, M. (1980). *Etude de la fertilité et du comportement méiotique des hybrides interspécifiques tétraploïdes arabusta* Coffea arabica×C. canephora. Thèse Doc. Univ. Paris-Sud (Centre d'Orsay).

Guillaumet, J.-L., & Hallé F. (1967). *Etude de la variability du* Coffea arabica *dans son aire d'origine*. Rapport de mission ORSTOM dans le Sud-Ouest de l'Ethiopie. Mult. ORSTOM (Adiopodoumé), Abidjan.

International Union of Pure and Applied Chemistry (1976). *Biochemical Journal*, 153, 23–31.

Kammacher, P., & Capot, J. (1972). Sur les relations caryologiques entre *Coffea arabica* et *C. canephora*. *Café Cacao Thé*, *16*, 289–294.

Ky, C.-L., Doulbeau, S., Guyot, B., Charrier, A., Hamon, S., Louarn, J., & Noirot, M. (2000). Inheritance of sucrose content in the interspecific cross: *Coffea pseudozanguebariae*×*Coffea liberica* '*dewevrei*'. *Plant Breeding*, *119*, 165–168.

Ky, C.-L., Guyot, B., Louarn, J., Hamon, S., & Noirot, M. (2001). Trigonelline inheritance in the interspecific cross between *Coffea pseudozanguebariae* × *C. liberica* var. '*dewevrei*'. *Theoretical and Applied Genetics*, *102*, 630–634.

Ky, C.-L., Louarn, J., Guyot, B., Charrier, A., Hamon, S., & Noirot, M. (1999). Relations between and inheritance of chlorogenic acid content in an interspecific cross between *Coffea pseudozanguebariae* and *Coffea liberica* var '*dewevrei*'. *Theoretical and Applied Genetics*, *98*, 628–637.

Ky, C.-L., Noirot, M., & Hamon, S. (1997). Comparison of five purification methods for chlorogenic acids in green coffee beans (*Coffea* sp.). *Journal of Agricultural and Food Chemistry*, *45*, 786–790.

Lashermes, P., Combes, M.-C., Robert, J., Trouslot, P., D'Hont, A., Anthony, F., & Charrier, A. (1999). Molecular characterisation and origin of the *Coffea arabica* L. genome. *Mol. Gen. Genet.*, *261*, 259–266.

Lashermes, P., Combes, M. C., Trouslot, P., Anthony, F., & Charrier, A. (1996). Molecular analysis of the origin and genetic diversity of *Coffea arabica* L.: implications for coffee improvement. *Compte-Rendus de colloque Eucarpia sur les Plantes Tropicales* (Montpellier), pp. 23-29.

Leloup, V., Louvrier, A., & Liardon, R. (1995). Degradation mechanisms of chlorogenic acids during roasting. *Compte-rendus du Seizième Colloque de l'ASIC (Association Scientifique Internationale sur le Café*, 192–198.

Le Pierres, D., Charmetant, P., Yapo, A., Leroy, T., Couturon, T., Bontems, S., & Tehe, H. (1989). Les caféiers sauvages de Côte d'Ivoire et de Guinée:bilan des missions de propection effectuées de 1984 à 1987. *Compte-rendus du Huitième Colloque de l'ASIC (Assocaition Scientifique International sur le Café)*, *8*, 420–428.

Leroy, T., Montagnon, C., Charrier, A., & Eskes, A. B. (1993). Reciprocal reccurent selection applied to *Coffea canephora* Pierre in Côte d'Ivoire. I. Characterization and evaluation of breeding populations and value of intergroup hybrids. *Euphytica*, *67*, 113–125.

Montagnon, C., Leroy, T., & Yapo, A. (1991). Etude complémentaire de la diversité génotypique et phénotypique des caféiers de l'espèce *Coffea canephora* en collection en Côte d'Ivoire. *Compte-Rendus du Quatorzième Colloque de l'ASIC (Association Scientifique International sur le Café*), 14–19.

Morishita, H., Takai, Y., Yamada, H., Fukuda, F., Sawada, M., Iwahashi, H., & Kido, R. (1987). Caffeoyltryptophan from green robusta coffee beans. *Phytochemistry*, *26*, 1195–1196.

Murata, M., Okada, H., & Homma, S. (1995) A novel hydroxycinnamic acid derivative of coffee bean. *Compte-rendus du Seizième Colloque de l'ASIC (Association Scientifique International sur le Café*), p 199-207.

Murata, M., Okada, H., & Homma, S. (1995). Hydroxycinnamic acid derivative, from coffee beans. *Bioscience, Biotechnology and Biochemistry*, *59*, 1887–1890.

Thaler, H., & Arneth, W. (1967) Hochpolymere Kolhenhydrate des rohen end gerösteten Kaffees. *Compte-Rendus du Troisième Colloque International sur la Chimie des Cafès*, 127–136.

Trugo, L. C., Macrae, R., & Dick, J. (1983). Determination of purine alkaloids and trigonelline in instant coffee and other beverages using high performance liquid chromatography. *Journal of the Science of Food Agriculture*, *34*, 300–306.

Voilley, A., Sauvageot, F., & Durand, D. (1977). Influence sur l'amertume d'un café boisson de quelques paramètres d'extraction. *Compte-Rendus du Huitième Colloque de l'ASIC (Association Scientifique International sur le Café*, 251–259.