ORIGINAL PAPER

Identification of chemical clusters discriminators of the roast degree in Arabica and Robusta coffee beans

Natalina Cavaco Bicho · António Eduardo Leitão ·
José Cochicho Ramalho · Fernando Cebola Lidon

Abstract To identify chemical parameters that might be used as discriminators, pH, soluble solids, caffeine, trigonelline, total caffeoylquinic acids, 3-*O*-caffeoylquinic acid, 4-*O*-caffeoylquinic acid, 5-*O*-caffeoylquinic acid, total dicaffeoylquinic acids, 3,4-*O*-dicaffeoylquinic acid, 3,5-*O*-dicaffeoylquinic acid, 4,5-*O*-dicaffeoylquinic acid, total feruloylquinic acids, 3-*O*-feruloylquinic acid, and 5-*O*-feruloylquinic acid were measured in Arabica and Robusta coffees submitted to three roasting levels. It was found that the fraction of soluble solids increased with roasting level, being slightly higher in Robusta roasted coffee. The contents of caffeine did not vary significantly between roasting degrees within the Arabica and Robusta samples, respectively, revealing a considerable stability during browning. The contents of trigonelline in Arabica and Robusta coffee decreased significantly with browning intensification. Overall, the levels of chlorogenic acids remained higher in Robusta roasted coffee beans but decreased sharply with roast increase. With roasting intensification, the ratio of total caffeoylquinic acids, total dicaffeoylquinic acids, and total feruloylquinic acids varied markedly in both species, with the proportion of total caffeoylquinic acids and total feruloylquinic acids increasing significantly, whereas the opposite occurred with dicaffeoylquinic acids. One can conclude, through the application of a multivariate analysis, that these chemicals form four clusters, constituting caffeine, trigonelline, total dicaffeoylquinic acids, and total feruloylquinic acids a relevant group for T_3 roasting level discrimination, in both coffee species. Additionally, detailing discriminators for roasting intensity in Arabica coffee might be caffeine, trigonelline, 3-*O*-caffeoylquinic acid, and 4-*O*-caffeoylquinic acid, whereas in Robusta roasted coffee are trigonelline, 3-*O*-caffeoylquinic acid, 4-*O*-caffeoylquinic acid, 5-*O*-caffeoylquinic acid, 3,4-*O*-dicaffeoylquinic acid, 3,5-*O*-dicaffeoylquinic acid, 4,5-*O*-dicaffeoylquinic acid, 3-*O*-feruloylquinic acid, and 5-*O*-feruloylquinic acid.

Abbreviations

CGA	Chlorogenic acid
CQA_{total}	Total caffeoylquinic acids
3-CQA	3-*O*-caffeoylquinic acid
4-CQA	4-*O*-caffeoylquinic acid
5-CQA	5-*O*-caffeoylquinic acid
$diCQA_{total}$	Total dicaffeoylquinic acid
3,4-diCQA	3,4-*O*-dicaffeoylquinic acid
3,5-diCQA	3,5-*O*-dicaffeoylquinic acid
4,5-diCQA	4,5-*O*-dicaffeoylquinic acid
FQA_{total}	Total feruloylquinic acid
3-FQA	3-*O*-feruloylquinic acid
5-FQA	5-*O*-feruloylquinic acid
ICO	International coffee organization

N. C. Bicho · F. C. Lidon (✉)
Departmento de Ciências e Tecnologia da Biomassa, Campus da Caparica, Faculdade de Ciências e Tecnologia, New University of Lisbon, 2829-516 Monte de Caparica, Portugal
e-mail: fjl@fct.unl.pt

N. C. Bicho · A. E. Leitão · J. C. Ramalho
Centro de Ecofisiologia, Bioquímica e Biotecnologia
Vegetal/Instituto de Investigação Científica Tropical, I.P., Quinta do Marquês, 2784-505 Oeiras, Portugal

Introduction

In green beans of *C. arabica* and *C. canephora* (Arabica and Robusta coffees, respectively), pH may vary between

5.26 and 6.13 [1], while soluble solids represent 23.85–27.31 and 26.07–30.60% of biomass, respectively [2]. In coffee beans, caffeine prevails in the cytoplasm of parenchyma cells, associated with chlorogenic acids, forming a potassium chlorogenate complex. Caffeine usually varies between 0.8 and 1.4% in *C. arabica* seeds and 1.7–4.0% in *C. canephora* (considering both Robusta and Conilon cultivars) [3]. Plants that produce seeds almost completely free of caffeine (0.08%) were also reported for *C. arabica* [4], while exceptionally high content values (5.5%) were detected in *C. canephora* [5].

Moreover, trigonelline content ranges between 0.3 and 1.3% d.b. [6]. Trigonelline has little direct influence (weak bitter taste) but its thermal degradation products contribute for final coffee aroma [7, 8]. The chlorogenic acids further accumulate in the cytoplasm of epidermal cells, but larger quantities can also be found in the periplasm. In green coffee beans, caffeoylquinic acids, dicaffeoylquinic acids, and feruloylquinic acids account for about 98% of total chlorogenic acids [9].

After roasting, the amount of caffeine in the seed slightly decreases, while chlorogenic acids and trigonelline are particularly degraded [3, 10]. Indeed, a large number of aromatic compounds result from the thermal decomposition of chlorogenic acids, namely phenolic esters, carbonyl compounds, esters, and polycyclic compounds [11], which contribute to flavor, acidity, and astringency of coffee drink [8]. In this context, hydroxycinnamic acids, which are particularly common in coffee beans, especially as chlorogenic acid in the double form of caffeine and potassium chlorogenate [10, 12], are significantly destroyed during roasting, with the release of alkaloids [13]. Additionally, 50–80% of trigonelline can be degraded during roasting [7, 14], with the formation of other compounds, namely nicotinic acid, pyridine, 3-methyl-pyridine, methyl ester of nicotinic acid, and the decarboxylated derivative 1-methylpyridinium as a major reaction product of trigonelline subjected to pyrolytic conditions [15].

It has long been known that the chemical composition of roasted coffee beans mostly depends on the temperature and duration of the roasting procedure [13]. Accordingly, this work aims to identify chemical discriminators that might be applied to differentiate the roasting intensity of commercial Arabica and Robusta coffee beans. The chemical composition of Arabica (from Brazil) and Robusta (from India) coffee beans after applying three levels of roasting levels was analyzed. For that was used a multivariate analysis in order to identify chemical clusters that might be usefully used as discriminators of roasting intensity and quality.

Materials and methods

Materials

Sampling from green coffee bags of *Coffea arabica* (from Brazil) and *Coffea canephora* (from India), selected because they are the two major trading species, was carried out according to the *Instrução Normativa* No 8 [16], NP 1666 [17], and ISO 4072 [18] and as recommended by the ICO (International Coffee Organization). The sampling process began with the selection of green coffee bags [19], at random (a minimum of 10% of the lot). The selected bags were separated from the lot, and ca. 30-g samples of coffee were collected in triplicate from three different points (top, middle, and bottom) of each bag. After extraction and homogenization, the portions were joined, for an overall take of green coffee, with a minimum mass of 1.5 kg. Arabica and Robusta green coffee samples were then submitted to three levels of roasting, T_1, T_2, and T_3, at 220 ± 10 °C, for 7, 9, and 11 min, respectively. These samples were then used for all subsequent studies.

Chemical characterization

Soluble solids and pH were measured according to AOAC [20, 21], at 25 °C. Ground roasted coffee (10 g $\pm$ 0.1 mg) mixed with water (200 mL) was boiled for 5 min and cooled at room temperature and the weight adjusted by adding water. After filtration (Whatman no 1 filter), the pH was measured at ca. 25 °C. For the quantification of soluble solids, 25 mL of the filtrate was evaporated to dryness and the residue placed in an oven at 105 °C and finally in a desiccator for cooling and weighed. Data are the average of triplicates for each sample of roasted coffee.
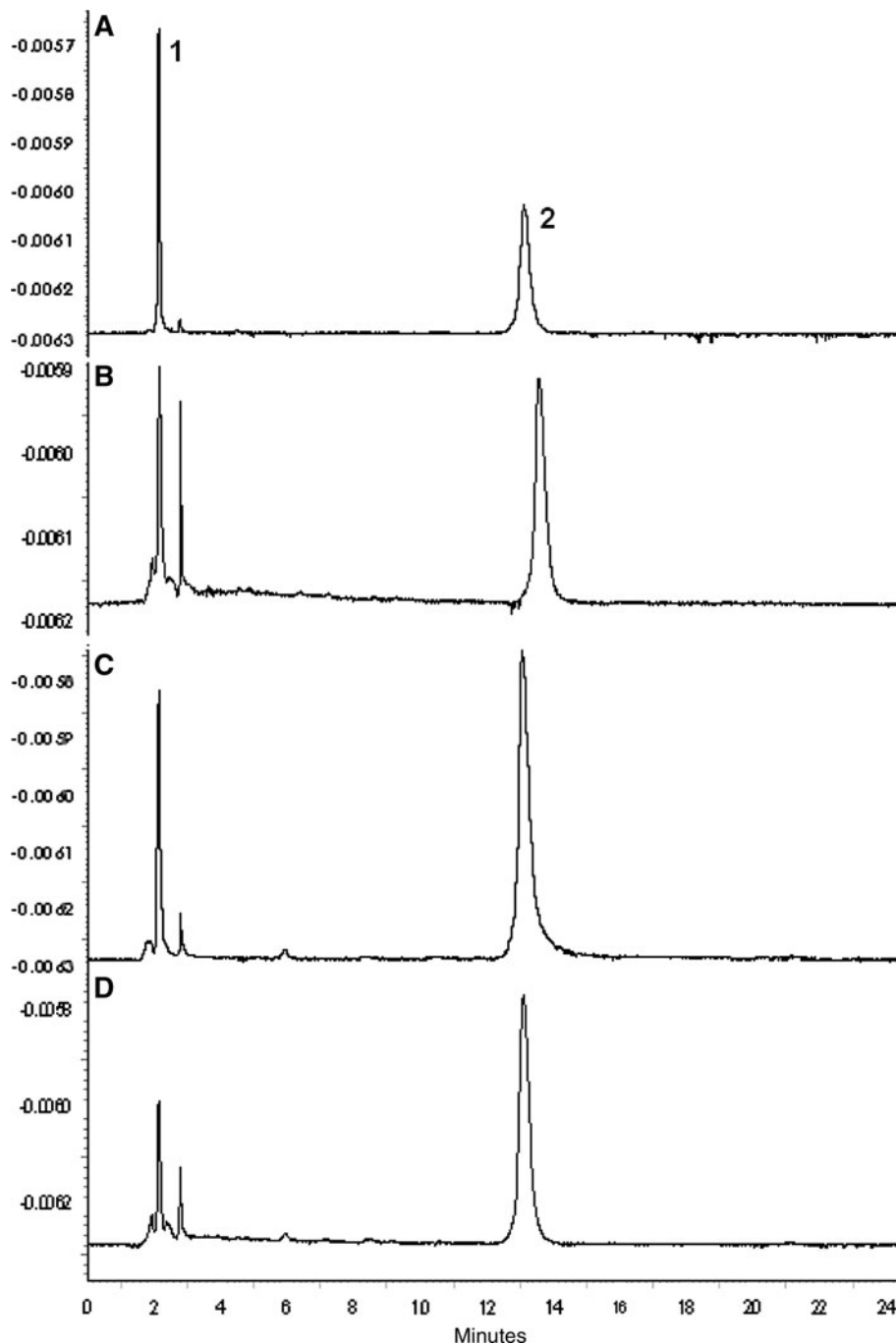
Caffeine and trigonelline contents were measured according to ISO 10095 [22]. Ground roasted coffee (1 g $\pm$ 0.1 mg) and magnesium oxide (4.5 $\pm$ 0.5 g) were homogenized in water (100 mL) and placed in a water bath at 90 °C for 20 min, with continuous stirring. After cooling, the volume was restored to the original level and the mixture was filtered (Whatman no 1 filter) without washing the solid residue. A 2-mL aliquot of the mixture was added to 10 mL of distilled water and filtered through a 0.45-μm filter. Caffeine and trigonelline were quantified in an integrated HPLC system (Waters, equipped with UV–VIS detector, model 440, column Lichrosorb 100 RP-18 from Merck, 5 μm particle size, 4 mm × 250 mm), using the 32 Karat Software (version 7.0, Beckman Coulter, Inc). The elution of 20-μL aliquots was performed with 20 mM phosphate buffer (pH 4.3) and acetonitrile solution (9:1), using a 1 mL min^{-1} isocratic flow rate and a column temperature of 25 °C. Detection was performed at 254 nm,

and the retention times for trigonelline and caffeine were 2 and 13 min, respectively (Fig. 1). For identification and quantification, standard curves were built with concentrations ranging between 8 and 1,000 µg mL^{-1} for caffeine and between 8 and 500 µg mL^{-1} for trigonelline. All extractions and chromatographic analysis were performed in triplicate.

Chlorogenic acids were analyzed as described elsewhere [20]. Briefly, after mixing 2 g ± 0.1 mg of ground roasted coffee with 10 mL of methanol/water (40:60), and mechanical agitation for 30 min, the mixture was

centrifuged (9,400×g, 5 min, 25 °C). The supernatant was decanted, 1 mL of Carrez solutions I (aqueous solution of zinc acetate dihydrate and glacial acetic acid, 10.95 g and 1.5 mL, respectively, to a final volume of 50 mL) and II (aqueous solution of 5.3 g of potassium hexacyanoferrate II trihydrate in a final volume of 50 mL) was added for clearing, and thereafter, a methanol/ water (40:60) solution was added to a final volume of 100 mL. After 15 min, the mixture was filtered (Whatman filter no 1) and an aliquot of 10 mL was taken and filtered (PVDF filter, 0.45 µm) prior

Fig. 1 Chromatograms of trigonelline and caffeine in Arabica and Robusta coffees. *A* green Arabica, *B* roasted Arabica (T$_3$), *C* green Robusta, *D* roasted Robusta (T$_3$). *1* trigonelline, *2* caffeine

to a reversed-phase HPLC analysis (Beckman System Gold system) using an end-capped, C18, 5-μm Spherisorb S5 ODS-2 column (250 × 4.6 mm) and the 32 Karat Software (version 7.0, Beckman Coulter, Inc). The elution of 20-μL aliquots was performed at ca. 25 °C, over 45 min with a 1 mL min^{-1} flow rate, using a linear gradient of 20–70% methanol in 100% tripotassium citrate buffer solution (10 mM, pH 2.5). The column was allowed to re-equilibrate for 10 min. Detection was carried out at 325 and 330 nm using a diode-array detector (mod.168, Beckman). Chromatograms showing the retention times of chlorogenic acids are shown in Fig. 2.

For the isomerization of chlorogenic acid (5-CQA), in order to obtain 3-CQA and 4-CQA standards, 200 mg of 5-CQA standard was solubilized in 20 mL of distilled water and the pH was adjusted to 8 with ammonium hydroxide (4 M). Then, the solution was boiled for 30 min in a water bath and cooled and the pH adjusted to 2.5 with HCl (4 M). After filtration, samples were used for quantification. The identification of chromatographic peaks and quantification of results were carried out using standard solutions of 5-CQA. To identify the isomers 3-CQA and 4-CQA, the standard 5-CQA isomer was subjected to isomerization, as already described. The peaks appeared with the following sequence: 3-CQA, 3-FQA, 4-CQA, 5-CQA, 4-FQA, 5-FQA, 3,4-diCQA, 3,5-diCQA, and 4,5-diCQA. The calibration curve was obtained from 5-CQA with readings at 325 and 330 nm. The quantification was done assuming the peak areas as a reference and comparing them with the standard 5-CQA. To quantify each compound, the following equation [23, 24] was used: $c = \frac{Fr \times \varepsilon_1 \times Mr_2 \times A}{\varepsilon_2 \times Mr_1}$ (being: c, the concentration of the isomer to quantify, in mg L^{-1}; Fr, the response factor of the standard 5-CQA in mg L^{-1} per unit area; ε_1, the molar absorption coefficient of the standard 5-CQA in L mol^{-1} cm^{-1}; ε_2, the molar absorption coefficient of the isomer to quantify, in L mol^{-1} cm^{-1}; Mr_2, the molecular weight on the isomer under study— CQA = 354.31 g mol^{-1}, FQA = 368.28 g mol^{-1}, diC-QA = 516.44 g mol^{-1}; Mr_1, the molar mass of acid 5-CQA; A, the peak area of the isomer to quantify). The molar absorption coefficients (3-CQA = 18,400, 4-CQA = 18,000, 5-CQA = 19,500, 3,4-diCQA = 31,800, 3,5-diCQA = 31,600, 4,5-diCQA = 33,200, with $\lambda = 330$ nm; 3-FQA = 19,000, 4-FQA = 19,500, 5-FQA = 19,300, with $\lambda = 325$ nm) indicated in [23, 24], in L mol^{-1} cm^{-1}, were used. Data were within the detection limits of the method. All extractions and chromatographic analysis were performed in triplicate.

Statistical analysis

Data were statistically analyzed using a two-way ANOVA ($p \leq 0.05$). Based on the ANOVA results, a Tukey's test was performed for mean comparison, for a 95% confidence level. Different letters in the tables indicate significant differences for 95% confidence level.

The multivariate analysis was carried out with STAT-ISTICA 6.0 software Copyright StatSoft, Inc., following several authors [25–29].

Results and discussion

The generation, volatilization, and degradation of acids occur simultaneously during roasting [11]. However, depending on the amount of acids formed during the pyrolysis of the coffee beans, the pH of different coffee genotypes can vary in a similar roasting [2]. In this context, it was found that pH reached a minimum value both for Arabica and for Robusta coffees at T$_2$ roasting level, whereas the highest value was developed at T$_3$ roasting level (Table 1). Such highest values would eventually result from the lower proportion between total aliphatic
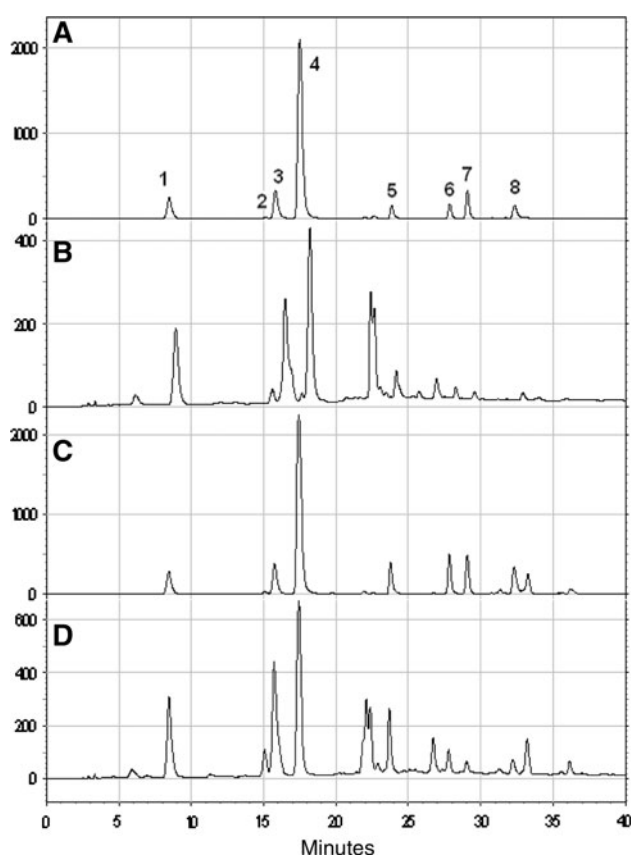

Fig. 2 Chromatograms of chlorogenic acids in Arabica and Robusta coffees (325 nm). *A* green Arabica, *B* roasted Arabica (T$_3$), *C* green Robusta, *D* roasted Robusta (T$_3$). *1* 3-CQA, *2* 3-FQA, *3* 4-CQA, *4* 5-CQA, *5* 5-FQA, *6* 3,4-diCQA, *7* 3,5-diCQA, *8* 4,5-diCQA

Table 1 Levels of pH, soluble solids, caffeine, and trigonelline (the last three in % of dry weight) of Arabica and Robusta coffee submitted to three (T_1, T_2, and T_3) roasting levels

Roasted samples	Parameters			
	pH	Soluble solids (%)	Caffeine (%)	Trigonelline (%)
Arabica				
T_1	5.12 ± 0.00[b,s]	26.2 ± 0.2[b,r]	1.38 ± 0.01[ab,s]	1.29 ± 0.02[a,r]
T_2	5.02 ± 0.00[c,s]	27.0 ± 0.1[b,s]	1.23 ± 0.04[b,s]	1.15 ± 0.00[b,r]
T_3	5.45 ± 0.01[a,r]	30.6 ± 0.0[a,r]	1.30 ± 0.01[ab,s]	0.553 ± 0.01[c,r]
Robusta				
T_1	5.32 ± 0.00[b,r]	26.99 ± 0.08[c,r]	2.16 ± 0.03[b,r]	0.902 ± 0.006[a,s]
T_2	5.29 ± 0.00[b,r]	28.66 ± 0.07[b,r]	2.28 ± 0.02[ab,r]	0.873 ± 0.014[a,s]
T_3	5.49 ± 0.01[a,r]	31.41 ± 0.22[a,r]	2.33 ± 0.02[a,b,r]	0.571 ± 0.005[b,r]

Each value is the mean $\pm$ SE ($n = 3$). *Different letters* indicate significant differences in a multiple range analysis, for 95% confidence level, among roasting levels for the same species (a, b, c) or between coffee types for the same roasting level (r, s)

acids and bean mass, as inferred from [30]. Independently of the roasting level applied, all pH values fall in the characteristic range observed among different genotypes [2, 11, 31].

The fraction of soluble solids also increased with roasting level, with a tendency to display values slightly higher in Robusta roasted coffee (Table 1). Although organic acids losses and chemicals volatilization during pyrolysis might trigger soluble solids losses, the raise of the later might prevailed after roasting due to the increased cell breakage of coffee beans, which implicates re-solubilization of celluloses and carbohydrates as well as protein denaturation [2].

During coffee roasting, the bean temperature rises up to about 200 °C, mostly due to a combination of external with internal heating (as temperature changes inside the roaster between 170 and 230 °C). Therefore, as temperature value surpasses the temperature of sublimation of caffeine (178 °C), considerable losses of this compound would be expected [32]. However, during roasting, the sublimation point of caffeine also increases due to the internal pressure rise, triggered by carbon dioxide synthesis and low vapor diffusion produced through the outer layers of the grain [32]. In this context, it was found that the contents of caffeine in Arabica and Robusta roasted coffees, although significantly different among them (Table 1), revealed a considerable stability during browning, close to values observed in other beans from different geographical origins [32–34]. Nevertheless, the caffeine levels were above the minimum (0.7%) required for marketing [35]. Caffeine contents of roasted coffees are similar to others [35, 36].

The contents of trigonelline found in the present work in Arabica and Robusta beans were similar to what was previously found in other coffees from different origins [34, 37]. These values strongly decreased with browning

intensification (Table 1), showing 43% (Arabica) and 63% (Robusta) in T_3 as compared to T_1. That would lead to pyridines, *N*-methylpyrrole and, due to demethylation, niacin synthesis [32]. The degradation of trigonelline, which was more pronounced in roasted Arabica coffee (except roasting level T_1), would result from the physical and chemical characteristics of grains of each type of coffee [32]. The values obtained for roasted coffees are similar to others [35].

The roasting also promoted severe decreases in CGA contents, as a result from various reactions, such as acyl migration, hydrolysis, oxidation, fragmentation, polymerization, and association with denatured proteins [38]. That led to lactone synthesis due to the dehydration of quinic acid residue, with the consequent formation of a ring, resulting from the participation of acid substituent groups at position C_1 and alcohol in position C_5. In this context, following the increase in roasting intensity, 3-CQA, 4-CQA, and 5-CQA decreased gradually to 17–20% in Arabica and to 27–37% in Robusta beans, thereby determining the sharp decrease in CQA$_{total}$. These losses were stronger in Arabica, which showed simultaneously the lowest values (Table 2). 5-CQA contents are similar for Arabica, but lower for Robusta than those obtained by others [35].

Concerning 3,4-diCQA, 3,5-diCQA, and 4,5-diCQA, a even higher susceptibility to degradation was observed. In fact, in T_3 treatments, the values ranged from 5 to 8% in Arabica and 10–14% in Robusta, when compared with T_1, and also remained significantly higher in Robusta beans (Table 3).

Although less affected than CQAs and diCQAs, the FQA values were significantly reduced under higher roasting conditions, with 5-FQA being more reduced than 3-FQA, while 4-FQA can be considered irrelevant, as also found by others [10, 37–39]. As for the CQAs and diCQAs,

Table 2 Levels of 3-CQA, 4-CQA, 5-CQA, and CQA$_{total}$ (in % of dry weight), of Arabica and Robusta coffee submitted to three (T$_1$, T$_2$, and T$_3$) roasting levels

Roasted samples	CQA (%)			
	3-CQA	4-CQA	5-CQA	CQA$_{total}$
Arabica				
T$_1$	0.987 ± 0.003[a,s]	1.23 ± 0.01[a,s]	2.21 ± 0.01[a,r]	4.43 ± 0.00[a,s]
T$_2$	0.594 ± 0.002[b,s]	0.814 ± 0.001[b,s]	1.28 ± 0.00[b,s]	2.68 ± 0.01[b,s]
T$_3$	0.203 ± 0.001[c,s]	0.258 ± 0.003[c,s]	0.379 ± 0.002[c,s]	0.840 ± 0.004[c,s]
Robusta				
T$_1$	1.06 ± 0.01[a,r]	1.31 ± 0.01[a,r]	2.19 ± 0.02[a,r]	4.56 ± 0.04[a,r]
T$_2$	0.777 ± 0.005[b,r]	1.03 ± 0.00[b,r]	1.58 ± 0.01[b,r]	3.38 ± 0.02[b,r]
T$_3$	0.310 ± 0.002[c,r]	0.488 ± 0.004[c,r]	0.589 ± 0.004[c,r]	1.39 ± 0.01[c,r]

Each value is the mean ± SE (*n* = 3). *Different letters* indicate significant differences in a multiple range analysis, for 95% confidence level, among roasting levels for the same species (a, b, c) or between coffee types for the same roasting level (r, s)

Table 3 Levels of 3,4-diCQA, 3,5-diCQA, 4,5-diCQA, and diCQA$_{total}$ (in % of dry weight) of Arabica and Robusta coffee submitted to three (T$_1$, T$_2$, and T$_3$) roasting levels

Roasted samples	diCQA (%)			
	3,4-diCQA	3,5-diCQA	4,5-diCQA	diCQA$_{total}$
Arabica				
T$_1$	0.166 ± 0.000[a,s]	0.130 ± 0.000[a,s]	0.175 ± 0.002[a,s]	0.471 ± 0.002[a,s]
T$_2$	0.070 ± 0.001[b,s]	0.050 ± 0.001[b,s]	0.065 ± 0.000[b,s]	0.185 ± 0.002[b,s]
T$_3$	0.014 ± 0.000[c,s]	0.009 ± 0.000[c,s]	0.009 ± 0.000[c,s]	0.033 ± 0.000[c,s]
Robusta				
T$_1$	0.315 ± 0.005[a,r]	0.202 ± 0.007[a,r]	0.276 ± 0.012[a,r]	0.794 ± 0.024[a,r]
T$_2$	0.180 ± 0.003[b,r]	0.116 ± 0.002[b,r]	0.160 ± 0.001[b,r]	0.456 ± 0.006[b,r]
T$_3$	0.043 ± 0.000[c,r]	0.021 ± 0.000[c,r]	0.033 ± 0.000[c,r]	0.097 ± 0.000[c,r]

Each value is the mean ± SE (*n* = 3). *Different letters* indicate significant differences in a multiple range analysis, for 95% confidence level, among roasting levels for the same species (a, b, c) or between coffee types for the same roasting level (r, s)

Table 4 Levels of 3-FQA, 5-FQA, and FQA$_{total}$ (in % of dry weight) of Arabica and Robusta coffee submitted to three (T$_1$, T$_2$, and T$_3$) roasting levels

Roasted samples	FQA (%)		
	3-FQA	5-FQA	FQA$_{total}$
Arabica			
T$_1$	0.070 ± 0.001[a,s]	0.162 ± 0.004[a,s]	0.232 ± 0.006[a,s]
T$_2$	0.055 ± 0.000[b,s]	0.107 ± 0.003[b,s]	0.162 ± 0.002[b,s]
T$_3$	0.028 ± 0.000[c,s]	0.040 ± 0.001[c,s]	0.068 ± 0.001[c,s]
Robusta			
T$_1$	0.165 ± 0.001[a,r]	0.413 ± 0.002[a,r]	0.578 ± 0.004[a,r]
T$_2$	0.142 ± 0.001[b,r]	0.305 ± 0.005[b,r]	0.447 ± 0.006[b,r]
T$_3$	0.084 ± 0.001[c,r]	0.157 ± 0.001[c,r]	0.241 ± 0.000[c,r]

Each value is the mean ± SE (*n* = 3). *Different letters* indicate significant differences in a multiple range analysis, for 95% confidence level, among roasting levels for the same species (a, b, c) or between coffee types for the same roasting level (r, s)

the FQA$_{total}$ of Robusta was less affected and a significantly higher value persisted when compared with Arabica beans (Table 4).

Our data showed that among CGA compounds, the diCQAs are the most sensitive, being severely degraded with more intense roasting, while FQAs, although strongly
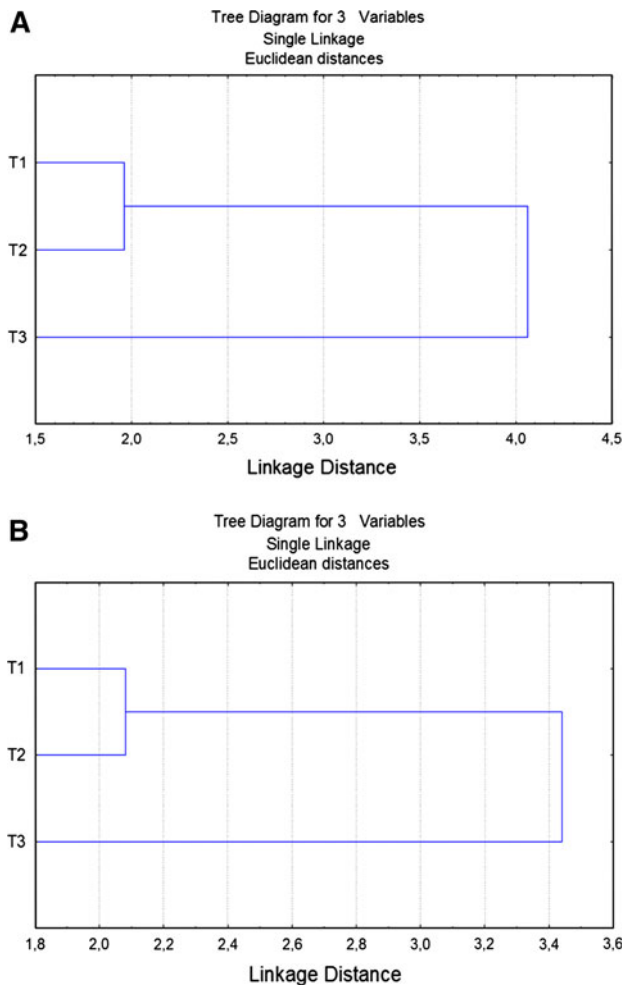
Fig. 3 Samples ($n = 3$) dendogram of roasting treatments (T_1, T_2, and T_3) of Arabica (**a**) and Robusta (**b**) coffees considering the parameters pH, soluble solids, caffeine, trigonelline, CQA_{total}, di-CQA_{total}, and FQA_{total} based on Euclidean distances between them

Fig. 4 Arabica coffees dendogram of roasting treatments (T_1, T_2, and T_3) of **a** pH, soluble solids, caffeine, trigonelline, CQA_{total}, diCQA_{total}, and FQA_{total} based on Euclidean distances between them; **b** pH, soluble solids, caffeine, trigonelline, the 3-CQA, 4-CQA, 5-CQA, 3,4-diCQA, 3.5-diCQA, 4.5-DICQ, 3-FQA, and 5-FQA, based on the Euclidean distances between them

affected, suffered a lower reduction, thus increasing its relative weight (Tables 2, 3, 4). Furthermore, in the overall CGAs of roasted coffees, the CQA fraction maintained the highest relative weight, prevailing 5-CQA.

To reduce the large number of variables (pH, soluble solids, caffeine, trigonelline, CQA_{total}, di-CQA_{total}, and FQA_{total}) to a smaller number of factors for data modeling, a cluster analysis using a linkage distance around 1.97 and 2.08 (Fig. 3a, b) allowed a discrimination of roasting T_3 from T_1 and T_2, in Arabica and Robusta samples. In this context, considering this critical design and the correlation across the hierarchical levels, in both coffee species, the differentiation of roasting T_3 was related to higher values for pH and soluble solids (Table 1), whereas lower values were detected for trigonelline, CQA_{total}, diCQA_{total}, and FQA_{total} (Tables 1, 2, 3, 4). The differentiation between these species was further linked to higher values of pH, soluble solids, caffeine, CQA_{total}, diCQA_{total}, and FQA_{total}
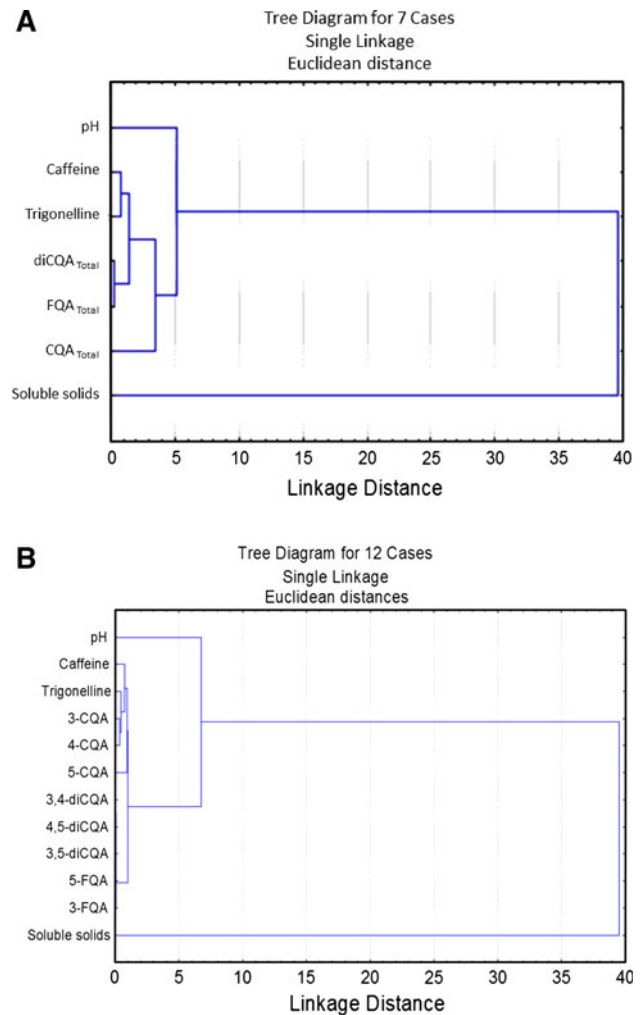
in Robusta roasted coffee, while the opposite occurred with trigonelline. To select a subset of variables from all variables set, (*i.e.*, considering the highest correlations with the principal component factors), data suggest that, following the patterns in each roasting degree, the parameters can be subdivided into four groups in both coffee species, according to their importance for discrimination (Figs. 3, 4). Following this hierarchical clustering, in both species, although the relevance of pH, soluble solids, and CQA_{total} content might be minimized (Figs. 4, 5), caffeine, trigonelline, di-CQA_{total}, and FQA_{total} constitute a relevant group for roasting T_3 discrimination (Fig. 4). Moreover, further assessing the significant variations (Tables 1, 2, 3, 4), T_3 roasting treatment can be differentiated in both species, as the detailing discriminators in Arabica roasted coffee are caffeine, trigonelline, 3, CQA, and 4-CQA
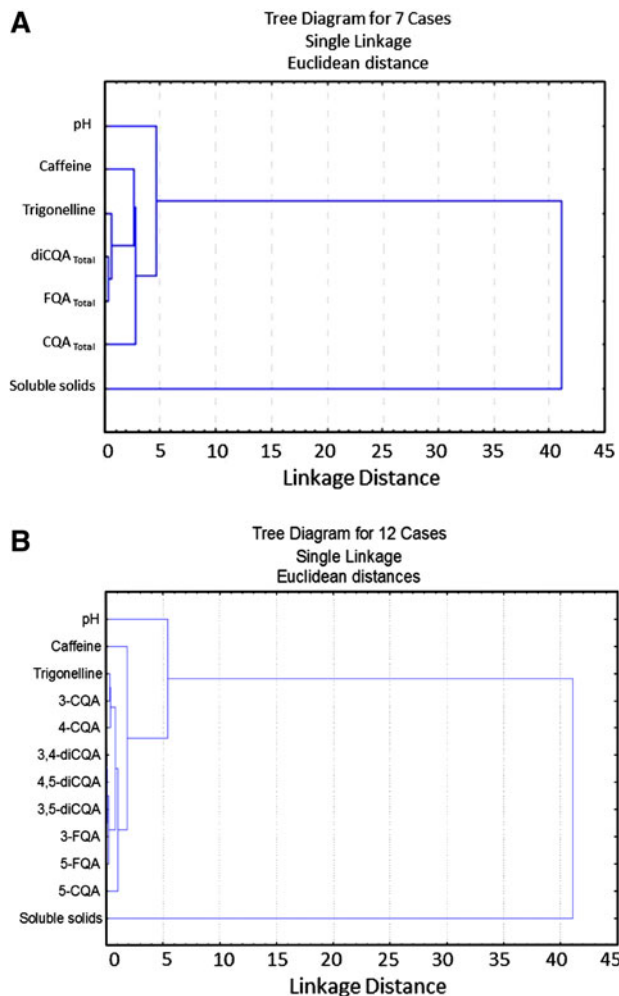
Fig. 5 Robusta coffees dendogram of roasting treatments (T$_1$, T$_2$, and T$_3$) of **a** pH, soluble solids, caffeine, trigonelline, CQA$_{total}$, diCQA$_{total}$, and FQA$_{total}$ based on Euclidean distances between them; **b** pH, soluble solids, caffeine, trigonelline, the 3-CQA, 4-CQA, 5-CQA, 3,4-diCQA, 3.5-diCQA, 4.5-DICQ, 3-FQA, and 5-FQA, based on the Euclidean distances between them

(Fig. 4b), whereas in Robusta roasted coffee are trigonelline, 3-CQA, 4-CQA, 5-CQA, 3,4-diCQA, 3,5-diCQA, 4,5-diCQA, 3-FQA, and 5-FQA (Fig. 3b).

Considering that factor analysis might reduce space from a larger number of variables to a smaller number of factors and following the chemical characterization of Arabica and Robusta roasted coffee provided in this study (that in general closely follows what is observed worldwide for these parameters in these coffee species), we suggest that the fourth cluster (pH, caffeine, trigonelline, diCQA$_{total}$, FQA$_{total}$, CQA$_{total}$, and soluble solids) might be used as a common discriminator for roasting intensity (Fig. 4a), further considering coffee species (Fig. 4b).

References

1. Leroy T, Ribeyre F, Bertrand B, Charmetant P, Dufour M, Montagnon C, Marraccini P, Pot D (2006) Braz J Plant Physiol 18:229–242
2. Mendonça LMVL, Pereira RGFA, Mendes ANG (2005) Ciências Tecnologia Alimentar 25:239–243
3. Nestlé (1991) Coffee. Documentation Centre, Corporate Affairs Department, Nestec Ltd, Vevey, Switzerland, pp 1–60
4. Silvarolla MB, Mazzafera P, Fazuoli LC (2004) Nature 429:826
5. Ferrão RG, Ferrão MAG, Fonseca AFA, Pacova BEV (2007) Melhoramento genético de *Coffea canephora*. In: Ferrão RG et al. (ed) Café Conilon, pp 121–173
6. Stennert A, Maier HG (1994) Z Lebensm Unters Forsch 199:198–200
7. Macrae R (1985) Nitrogenous components. In: Clarke RJ, Macrae R (eds) Coffee I: Chemistry. Elsevier Applied Science Publishers Co Ltd, London & New York, pp 115–152
8. Morais SAL, Aquino FJT, Nascimento PM, Nascimento EA, Chang R (2009) Quim Nova 32:327–331
9. Clifford MN, Staniforth PS (1977) A critical comparison of six spectrophotometric methods for measuring chlorogenic acids in green coffee beans. In: Huitième Colloque Scientifique International sur le Café. ASIC, Paris, pp 109–114
10. Viani R (1993) The composition of coffee. In: Caffeine, coffee, and health. S Garatin Raven Press Ltd, New York, pp 17–41
11. Clifford MN (1987) Chemical and physical aspects of green coffee products. In: Clifford MN, Wilson KC (eds) Coffee, botany, biochemistry and production of beans and beverage. Croom Helm & Methuen Inc, New York, pp 305–374
12. Clifford MN (1999) J Sci Food Agric 79:362–372
13. Correia MNG (1995) Manual da tecnologia do café. Cultivar, Associação dos Técnicos de Culturas Tropicais, Porto, p 117
14. Martin MJ, Pablos F, Bello MA, González AG (1997) Fresenius J Anal Chem 357:357–358
15. Stadler RH, Varga N, Hau J, Vera FA, Welti DH (2002) J Agric Food Chem 50:1192–1199
16. Instrução Normativa no 8 (2003) Regulamento técnico de identidade e de qualidade para a classificação do café beneficiado grão cru. Ministério de Estado da Agricultura, Pecuária e Abastecimento, Brasil
17. NP 1666 (1980) Café—Colheita de amostras de café verde em sacos. Direcção Geral de Qualidade, Lisboa
18. ISO 4072 (1982) Green coffee in bags-sampling. International Organization for Standardization
19. PSCB no 36/02 (2002) Code of practice: enhancement of coffee quality through prevention of mould formation. International Coffee Organization
20. AOAC (1996a) Coffee and tea. AOAC International—Association of Official Analytical Chemists, vol. 2, chapter 30
21. AOAC (1996b) Definition of terms and explanatory notes. AOAC International—Official methods of analysis, vol. 1, note 11
22. ISO 10095 (1992) Coffee. Determination of caffeine content. Method using high-performance liquid chromatography. International Organization for Standardization
23. Trugo LC, Macrae R (1984) Analyst 109:263–266
24. Farah A, De Paulis T, Trugo LC, Martin PR (2005) J Agric Food Chem 53:1505–1513
25. Neto JMM, Moita GC (1998) Quim Nova 21:467–469
26. Alvarenga N, Silva P, Rodriguez-Garcia J, Sousa I (2008) J Dairy Res 75:233–239

27. Chapman KW, Lawless HT, Boor KJ (2001) J Dairy Sci 84:12–20

28. Maeztu L, Andueza S, Ibañez C, de Peña MP, Bello J, Cid C (2001) J Agric Food Chem 49:4743–4747

29. Da Silva AFS, Minim VPR, Chaves JBP, Stringheta PC, Ribeiro MM (2004) Ciências Tecnologia Alimentar 24:468–472

30. Woodman JS (1985) Carboxylic acids. In: Clarke RJ, Macrae R (eds) Coffee I: Chemistry. Elsevier Applied Science Publishers Co Ltd, London and New York, pp 266–289

31. Siqueira HH, Abreu CMP (2006) Ciências Agrotecnológicas 30:112–117

32. Toci A, Farah A, Trugo LC (2006) Quim Nova 29:965–971

33. Daglia M, Cuzzoni MT, Dacarro C (1994) J Agric Food Chem 42:2273–2277

34. IC AB (2007) Revisão 15—Norma de qualidade recomendável e boas práticas de fabricação de cafés torrados em grão e cafés torrados moídos. Associação Brasileira da Indústria do Café, Brasil, pp 1–20

35. Alves ST, Dias RCE, Benassi MT, Scholz MBS (2006) Quim Nova 29:1164–1168

36. Mori EEM, Bragagnolo N, Morgano MA, Anjos VDA, Yotsu-yanagi K, Faria EV, Iyomasa JM (2003) Food Food Ingred J Jpn 208:416–423

37. Clifford MN (1985) Chlorogenic acids. In: Clarke RJ, Macrae R (eds) Coffee I: Chemistry. Elsevier Applied Science Publishers Co Ltd, London and New York, pp 153–202

38. Ky LC, Louarn J, Dussert S, Guyot B, Hamon S, Noirot M (2001) Food Chem 75:223–230

39. Viani R (1986) Coffee. In: Ullmann's Encyclopedia of industrial chemistry, vol A7. VCH Verlagsgesellschaft, Weinheim, pp 315–339