COFFEE SCIENCE
ISSN: 1984-3909

eISSN: 1984-3909
Coffee Science, e172027, 2022

Doi: https://doi.org/10.25186/.v17i.2027

Bioactive compounds in blends of coffee defects originating from the harvesting

Rafael Carlos Eloy Dias[1], Sebastian Ed Wieland Opitz[1], Chahan Yeretzian[1]

[1]Zurich University of Applied Sciences, ZHAW, Campus Wädenswil, School of Life Sciences and Facility Management, LSFM, Institute of Chemistry and Biological Chemistry, ICBC, Wädenswil, Canton Zurich, Switzerland
Contact authors: rafael.rafaeltam@gmail.com; sebastian.opitz@zhaw.ch; yere@zhaw.ch
Received in March 17, 2022 and approved October 11, 2022

ABSTRACT

A coffee crop may consist of up to 1/5 of defective beans and finding a suitable destination for this material is economically interesting. Many coffee industries collect the selections - material containing coffee defects - and blend them with non-defective coffee fruits in specific proportions to obtain a marketable product. Studies on the composition of selections are scarce. Hydro- and liposoluble bioactive compounds were determined in five types of roasted and ground selections of coffee Arabica and in healthy Arabica and Robusta coffee species throughout an optimized HPLC-UV/Vis-MS-based method. Nicotinic acid and 4-CQA were not detected. Black and sour beans seem to increase the level of caffeine (variation from 1.3 to 2.4 g 100 g^{-1} of sample) in the selections. The occurrence of defects decreases the level of 5-CQA, the main representative chlorogenic acid. Trigonelline content is high in Arabica coffee, and the presence of defects does not promote a clear variation in its amount. Kahweol concentration (~74.6 to 76.9 mg g^{-1} of oil) was practically the same up the Arabica sample set; this diterpene was absent in Robusta coffee. Cafestol (variation between 12.4 and 16.4 mg g^{-1} of oil) is a good quality indicator. Kahweol and 16-*O*-methyl cafestol are species indicators, and caffeine can point out the species of coffee. PCA revealed that sour beans were associated with the presence of kahweol, while cafestol and trigonelline were correlated to the occurrence of coffee skin. The higher the proportion of black beans, the more balanced the contents of water-soluble and liposoluble compounds.

Key words: PVA; Coffee harnessing; HPLC-based method; Selection of coffee; Water- and fat-soluble compounds.

1 INTRODUCTION

The sensory attributes of roasted and ground coffee brews are heavily dependent on the chemical composition of coffee beans. The presence of defective beans during processing and roasting reduces overall cup quality by contributing to off flavours (Casas et al., 2017; Dias et al., 2010; Giacalone, et al. 2018). Two species of coffee are commercially relevant: *Coffea canephora* Pierre ex A. Froehener (Robusta coffee) and *Coffea arabica* L. (Arabica coffee). Arabica usually produces full-bodied coffees, with a more refined aroma, but more expensive; Robusta is widely used in the production of instant coffee because this species has a higher content of soluble solids. Robusta can be also added to Arabica coffee to produce blends of roasted and ground coffee (Dias and Benassi, 2015; Mendonça; França and Oliveira, 2009).

There are two main methods for harvesting coffee cherries: selective harvesting and stripping. In the first system, only the ripened fruit is harvested selectively from the tree, in several rounds. Visually chosen and handpicked, the harvested coffee meets the standards due to the uniformity of the cherries. In the second method, whole coffee cherries are harvested at most in three rounds, usually made by machinery. In both cases, when cherries are dropped to the ground (this is especially the case in the second technique), they have to be either raked up and collected manually or collected by machines (mostly based on a suction principle) and transported to the processing facilities. This coffee stubble may contain non-defective ripen beans, and important coffee defects, such as immature or overripe cherries, insect damaged, broken or brocade beans, black and sour beans, leaves, and strange matters, such as straw, wood, skin, stones, and clod (Haile and Kang, 2019; Wintgens, 2004). This mix of the matter is commonly named as “selection” in Brazilian coffee farms. Varied proportions of healthy coffee beans and defective beans and other undesirable material are found in different selections. The selection composition is dependent on a series of factors including the harvesting method chosen, the type of machinery employed, the procedures of selection, the uniformity of maturation, the yield per hectare, tree density, the gradient of the slope etc.

Defective coffee beans are present in a crop up to 20% (Dias et al., 2018) and separating them from non-defective beans is not an easy task - sometimes it is impossible, depending on the type of defect and the occurrence of foreign matter. In many roasting industries, selections are intentionally added to harvested non-defective beans in specific blends and proportions (from 5 to 20% of selections). This practice improves the use of raw materials and allows them to standardize the quality of industrialized roasted coffee (Dias et al., 2018; Brasil, 2003).

Before any coffee is commercialized, it is classified by the number of defects, screen size, and cup quality. The defect count is supposed to provide a general idea of the quality of the cup. There are international standards for the

presence of defects of coffee and the allowed use limit. Two green coffee classification methods have been described as more comprehensive and accurate. The SCAA Green Coffee Classification Method (SCAA, 2022) is more relevant when applied to specialty green coffee beans, but it leaves out a few of the important defects, typical for lower grade coffees. The Brazilian/New York Green Coffee Classification Method is more precise in terms of differentiation more along the whole spectrum of quality and complete in terms of encountered defects, but also more time consuming to determine. Yet, it allows achieving a good relationship between the defective coffee beans and the cup quality.

Roasting is the process step that most strongly changes the chemical composition of the bean when transforming the green into roasted beans (Diviš; Pořízka and Křikala, 2019; Vignoli et al., 2014). Almost all sucrose content of coffee is degraded (96-98%) after medium to dark roasting. Among chlorogenic acids, a class of antioxidant compounds of the coffee matrix, 65% (medium roasting degree), and 85% (dark roasting degree) of compounds contained in green coffee are degraded. The production of formic and lactic acids was related. The highest concentration of organic acids was observed in medium roasted coffee. Cafestol and kahweol, two diterpenes of the unsaponifiable matter of coffee, are degraded during roasting to dehydro derivatives (Dias et al., 2014). It is important to mention that cafestol is probably responsible for the effect on serum cholesterol levels and kahweol is mainly responsible for the effect on liver enzyme levels. During roasting, trigonelline partially degrades to produce two important compounds, pyridines and nicotinic acid (aka vitamin B3 and niacin), such that a dark roast will contain only a small fraction of its original trigonelline content. Trigonelline has therapeutic potential as a neuroprotective and hypoglycaemic agent, in addition to anticarcinogenic effects. Nicotinic acid may help lower cholesterol, ease arthritis, and propel brain function, among other benefits (Alves et al., 2006; Casal et al., 2000; Ashihara; Ludwig and Crozier, 2020; Jeszka-Skowron; Frankowsk and Zgoła-Grześkowiak, 2020).

The main reasons for caffeine being the most well-known and investigated coffee compound is that (i) the level of caffeine in healthy coffee beans is not markedly reduced during the process of obtaining RG coffee (Dias; Benassi, 2015), and (ii) its health effects (Ashihara; Ludwig and Crozier, 2020).

There are well-defined methods to grade green coffee according to the presence of coffee defects - it is important to highlight that it is applicable for unroasted beans; there is no official method for roasted samples in Brazil. For example, considering the SCAA Coffee Beans Classification, specialty green coffee beans cannot have more than 5 full defects in 300 grams of coffee. Specialty coffee must possess at least one distinctive attribute in the body, flavor, aroma, or acidity. It must be free of faults and taints and no quakers are permitted (SCA, 2022). All these classification definitions of the selections are done based on the green bean. Once roasted and ground, however, the assessment of the quality of coffee becomes exceedingly difficult based on visual clues. Besides, there is no information in the literature about the composition of selections, once roasted.

De Morais et al. (2007) studied the chemical composition of Arabica coffee and defective beans submitted to different degrees of roasting. Posteriorly, Franca and Oliveira (2008) reported a review on physical and chemical attributes of defective coffee beans in comparison to healthy ones is herein provided, for both green and roasted coffees. They highlighted that differences in chemical attributes include proximate composition, acidity, pH, sucrose levels, caffeine, trigonelline, chlorogenic acids, amines, and volatile substances, but, in the case of roasted coffee, only an evaluation of the volatile profile will effectively provide the means for differentiation. Casas et al. (2017) emphasized that the presence of defective beans during processing and roasting contributes to off flavors and reduces overall cup quality. The researchers listed a set of 35 compounds in roasted beans, dominated by volatile compounds, organic acids, sugars, and sugar alcohols, which were sufficient to discriminate the defective to non-defective fractions.

In a previously work, this research group reported the development of a method based on Infrared spectroscopy and photoacoustic detection (FTIR-PAS) associated with multivariate calibration potentially able for discriminating the coffee selections (Dias et al., 2018a). In the current study, water-soluble bioactive compounds and diterpenes of the lipid fraction of coffee were identified and their concentration was estimated. The selections have never been investigated before in terms of chemical composition and in this form (directly obtained from the harvest field and analysed). The selections are genuine samples and new to the field of investigation of the chemical composition of coffee.

2 MATERIAL AND METHODS

Samples were provided by Instituto Agronômico do Paraná (IAPAR): Latitude−23.29, Longitude−51.17; 23° 17′ 34″ S, 51° 10′ 24″ W, humid subtropical climate. Samples of healthy and whole Arabica coffee beans, healthy and whole Robusta coffee beans and five selections from the Arabica crop were evaluated (Table 1). The selections were chosen from a panel of 25 blends with different proportions of defects. They were described in a previous report (Dias et al., 2018b) and analysed by FTIR-PAS and chemometrics (Dias et al., 2018a). Those selections that presented the highest proportion of each pre-identified parameter were used (Table 1).

Table 1: Samples of selections investigated here. Each selection is defined by a Sample ID and the proportion (w/w) of six selected defects.

Sample ID[a]	Parameter (%)					
	whole	broken	sour	black	skin	wood
E1	10.6*	37.4*	37.4	13.2	1.30	0.100
E2	3.60	11.8	31.6	43.0*	9.90	0.100
E15	0.60	7.60	73.5*	15.3	1.70	0.300
E16	3.10	19.1	50.1	7.70	19.6*	0.200
E25	7.20	7.70	58.9	14.3	6.7	5.10*

[a]Identification of selections (E) according to the panel of 25 selections described in Dias et al. (2018b). *The highest value (in %) of the parameter found among the 25 selections. (E1 has the highest % of healthy whole beans and broken beans).

The seven samples (five selections and the two whole/non-defective samples of Arabica and Robusta coffees) were roasted to the same roasting degree: medium-dark (~17% weight loss; 22 to 26 lightness provided by a Konica Minolta® portable colorimeter BC-10). They were then ground (Ditting grinder, from Bachenbülach®, Switzerland, model KR805; setting/level 2) before going for analysis.

3 DETERMINATION OF BIOACTIVE COMPOUNDS

Qualitative and quantitative assessment of specific water- and lipo-soluble compounds was achieved through an Agilent 1200 series HPLC instrument coupled to a DAD UV-Vis detector and mass spectrometer (Agilent®, Santa Clara, USA; model 6130) with simple quadrupole/electrospray ionization (ESI).

Certified standards for caffeine, 3-CQA, 4-CQA, and 5-CQA (Sigma®, Steinheim, Germany) were used, while additional compounds were evaluated by their molar absorptivity. The calibration curves for 5-CQA and caffeine were obtained with nine dilution levels: 1, 2, 10, 50, 100, 200, 300, 400, and 500 ppm (triplicate). 3-CQA was quantified considering the molar absorptivity of 5-CQA (Shan et al., 2017).

The moisture content of the samples was determined in Halogen Moisture Analyzer equipment (Mettler® Toledo; Greifensee, Switzerland) with a heating gradient of 0 min (25 ºC) - 5 min (105 ºC) - 7 min (105 ºC) and used to calculate concentrations on a dry matter basis.

Extraction with water and HPLC-DAD-ESI-MS analysis were used to determine caffeine, chlorogenic acids, trigonelline and nicotinic acid. A previously developed and validated method was used (Alves et al., 2006), where 0.500 g of roasted and ground (RG) coffee was submitted to hot extraction in 30.0 mL of purified water (Milli-Q® Reference Water Purification System) (resistivity 18.2 MΩ.cm at 25 °C), dilution (factor 5), filtration (Millipore®, Billerica, USA; 0.45 µm), and injection into the HPLC. Start conditions were based on a method previously described (Alves et al., 2006; Dias and Benassi, 2015). The chromatographic conditions were adapted.

The lipid content was determined by extraction (10.0 g of ground coffee with 20.0 g of sodium sulphate - as dehydrating agent) via a Soxhlet apparatus (Büchi® B-811; Flawil, Switzerland) for 3.5 h with *tert*-butyl methyl ether, adapted from Wermelinger et al. (2011) (originally proposed a 5 h of extraction).

For liposoluble compounds determination, a previously developed extraction method (Dias et al., 2010) was employed, which consists of direct hot saponification of RG coffee sample (0.200 g) with potassium hydroxide saturated solution (Merck®, Darmstadt, Germany), extraction of unsaponifiable matter (which contains diterpenes) with *tert*-butyl methyl ether (Merck®, Darmstadt, Germany), cleaning with water, dilution, filtration (Millipore®, Billerica, USA; 0.45 µm) and injection into the chromatograph. Cafestol, kahweol, and 16-*O*-methyl cafestol (16-OMC) were determined.

The identification of compounds was based on the UV-Vis spectrum profile, maximum absorption wavelength ($\lambda_{máx}$), retention time (R_t), peak profile, and MS spectrum (more information is available in Table 2).

The chromatographic column (Phenomenex Kinetex® 2.6 µm, C18, 100 Å, 100 x 2.1 mm) was shorter and had smaller stationary phase particles compared to that used in the original method (Alves et al., 2006). After studying chromatography and MS conditions, minor changes were made. The best results for water-soluble compounds determination were obtained using mobile phase A (water), B (acetonitrile); 0 min, 70% A; 10 min, 55% A; 15 min, 40% A; 20 min (stop), 35% A (the analysis time was reduced to 2/3). For liposoluble compounds, isocratic elution (55:45 v/v acetonitrile: water solution) was used. Further suited parameters were an injector temperature of 10 ºC and column temperature of 25 ºC; for MS: drying gas flow (N_2) of 13 L min^{-1}, drying gas temperature of 350 °C, nebulizer pressure of 60 psig, the capillary voltage of 2500 V for positive mode and 3500 V for negative, 100-600 m/z range, and 150 V of shredder voltage.

The average (triplicate) of HPLC-DAD-ESI-MS measurements was used. Analysis of variance and the Tukey test ($p \leq 0.05$) were applied for group means comparison using a randomized split-plot design of Statistica 7.0 (Statistica for Windows-Computer program manual Version 7.0, Statsoft Inc.: Tulsa, OK, USA, 2005). Data was evaluated by Principal Components Analysis (PCA) using software Paleontological Statistics version 4.03 (Hammer et al., 2001).

Table 2: Summary of HPLC-DAD-MS analysis.

Compounds	Chromatographic data			Mass spectrometry data		Identification of compounds	
Water-soluble	λ_{max} (nm)	R_T (min)	Mode (MH^+)	MS (*m/z*): literature data (Perrone, Donangelo and Farah, 2008)	Experimental *m/z*	Chromatographic and MS peaks	Confirmation
Trigonelline	260 nm	1.63	Positive (138)	138	138 (MH^+)	Well-defined chromatographic peaks. MS peaks similar to those described in the literature (Alves et al., 2006; Dias et al., 2014;)	Confirmed
Nicotinic acid	260 nm	2.46	Positive (124)	124	124 (MH^+) and others less important	Chromatographic and MS peaks of low intensity (Alves et al., 2006; Dias et al., 2014; Narita and Inouye, 2015)	Confirmed
3-CQA	325 nm	10.13	Negative (353)	353, 191 (10%), 179 (5%)	353 (MH^+), 191 (10%), 179 (5%)	Well-defined chromatographic peaks. MS peaks similar to the standard and literature (Jeon et al., 2019; Narita and Inouye, 2015)	Confirmed
4- CQA	325 nm	-	Negative (353)	353, 191 (6%), 179 (5.5%), 173 (5%)	Characteristic peaks not found	No characteristic peaks	Not found
5- CQA	325 nm	11.45	Negative (353)	353, 191 (80%)	353 (MH^+), 191 (40 - 45%)[1]	Well-defined chromatographic peaks. MS peaks similar to the standard[1] (Alves et al., 2006; Dias et al., 2014; Jeon et al., 2019; Narita and Inouye, 2015)	Confirmed
Caffeine	275 nm	11.8	Positive (195)	195, 138 (11%)	195 (MH^+), 138 (11%)	Well-defined chromatographic peaks. MS peaks similar to the standard and literature (Alves et al., 2006; Dias et al., 2014; Jeon et al., 2019)	Confirmed
Liposoluble	λ_{max} (nm)	R_T (min)	Mode (MH^+)	MS (*m/z*): dados da literatura (Dias et al. 2014)	Experimental *m/z*	Chromatographic and MS peaks description	Confirmation
Kahweol	290	4.55	Positive (315)	315, 279, 297	279, 297 (base peak), 315	Well-defined chromatographic peaks. MS peaks similar to the literature (Dias et al., 2010; Speer and Kölling-Speer, 2006)	Confirmed
Cafestol	230	4.77	Positive (317)	317, 281, 299	281, 299 (base peak), 317	Well-defined chromatographic peaks. MS peaks similar to the literature (Dias et al., 2010; Speer and Kölling-Speer, 2006)	Confirmed
16-*O*-methyl cafestol	230	9.02	Positive (331)	331, 299, 147	331, 299 (base peak)	Well-defined chromatographic peaks. MS peaks similar to the literature (Gunning et al., 2018; Speer and Kölling-Speer, 2006)	Confirmed

[1] Peak m/z 191 does not have the same relative intensity of the MH^+ peak described in the literature. Therefore, there is a possibility that 4-CQA was eluted together with 5-ACQ. Abbreviations: R_T = retention time, MH^+ = molecular ion peak, λ_{max} = maximum absorption wavelength.

4 RESULTS

Caffeine, trigonelline, nicotinic acid, chlorogenic acids (3-ACQ, 4-ACQ and 5-ACQ), and the diterpenes caveol, cafestol and 16-*O*-methyl cafestol were assessed by HPLC-DAD-MS.

Figure 1 shows a typical chromatogram obtained from RG coffee. The chromatographic conditions produced a clean chromatogram even considering the high complexity of the coffee matrix.

Calibration curves for caffeine and 5-CQA showed R^2 of 0.99. Nicotinic acid and 4-CQA were not detected in the sample set. When found, these compounds are in low concentrations in roasted coffees, sometimes only traces (Jeon et al., 2019; Narita; Inouye, 2015; Nogueira and Trugo, 2003); furthermore, 4-CQA separation is difficult, since chlorogenic isomers have similar chemical structures and are often eluted together from the chromatographic column (Nogueira and Trugo, 2003). Nicotinic acid is a product of roasting, and its occurrence depends on the intensity of this process and the concentration of trigonelline, its precursor (Dias and Benassi, 2015; Seninde and Chambers, 2020).

Table 2 presents details of the analytical results.

Figure 2 reveals the discrepancy in the composition of RG coffee samples regarding some of the most studied water-soluble compounds of coffee.

The lipid content (w/w) of samples was 16.1% (A100), 15.7% (E2), 15.2% (E15), 14.5% (E1), 14.2% (E16), 13.2% (E25), and 10.7% (R100). These data indicate that the presence of coffee defects decreases the concentration of lipids.

To estimate the content of diterpenes, the previously determined molar absorptivity (Dias et al., 2014) was used (R^2 = 0.99 for the calibration curve of cafestol and kahweol). 16-OMC was determined using the absorptivity of cafestol. This research group reported the development and validation of the chromatographic methodology for the determination of diterpenes in coffee, where chromatograms and other related data can be found in detail (Dias et al. 2010; Dias et al., 2014). Figure 3 presents the diterpene contents of the coffee samples.

For an exploratory view of data, Principal Component Analysis (PCA) was applied considering water-soluble and liposoluble compounds (Figure 4).

5 DISCUSSION

The level of water-soluble compounds for Robusta and Arabica coffee species agrees with the range described in the literature (Agresti et al., 2008; Dias et al., 2014; Dias and Benassi, 2015; Feifei and Tanokura, 2015). Caffeine content was approximately 1.6 time higher in Robusta than in Arabica. E2, E15 and E16 selections showed higher levels of caffeine than E1 and E25. E2 is the sample with a higher level of black beans, and E15 has 73% sour beans, i.e., these defects seem to raise the level of caffeine. No significant differences in caffeine levels were reported among the defective beans (black, sour, and immature defects) (Mazzafera, 1999).

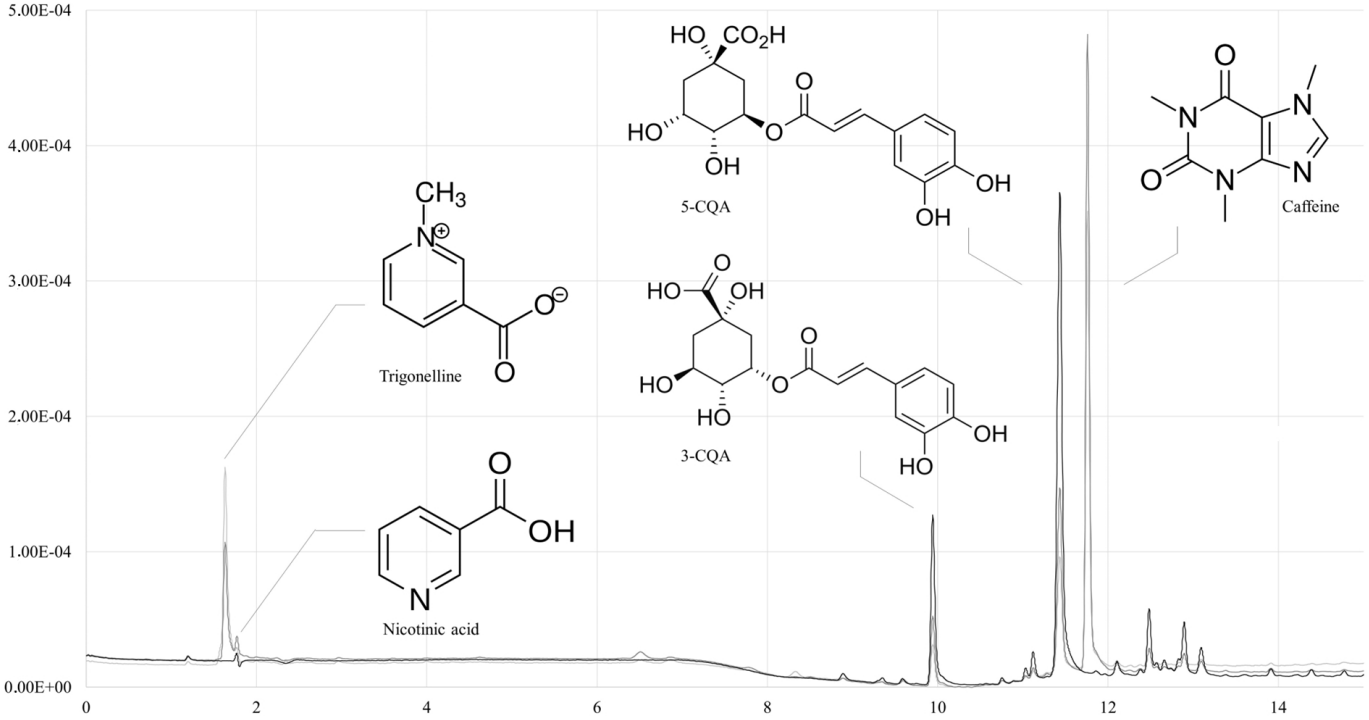

Figure 1: UV-Vis chromatogram of roasted and ground coffee (in this case, Arabica 100%). Wavelength of detection: 260 nm (trigonelline and nicotinic acid), 325 nm (CQAs) and 275 nm (caffeine). Further details are found in Table 2.

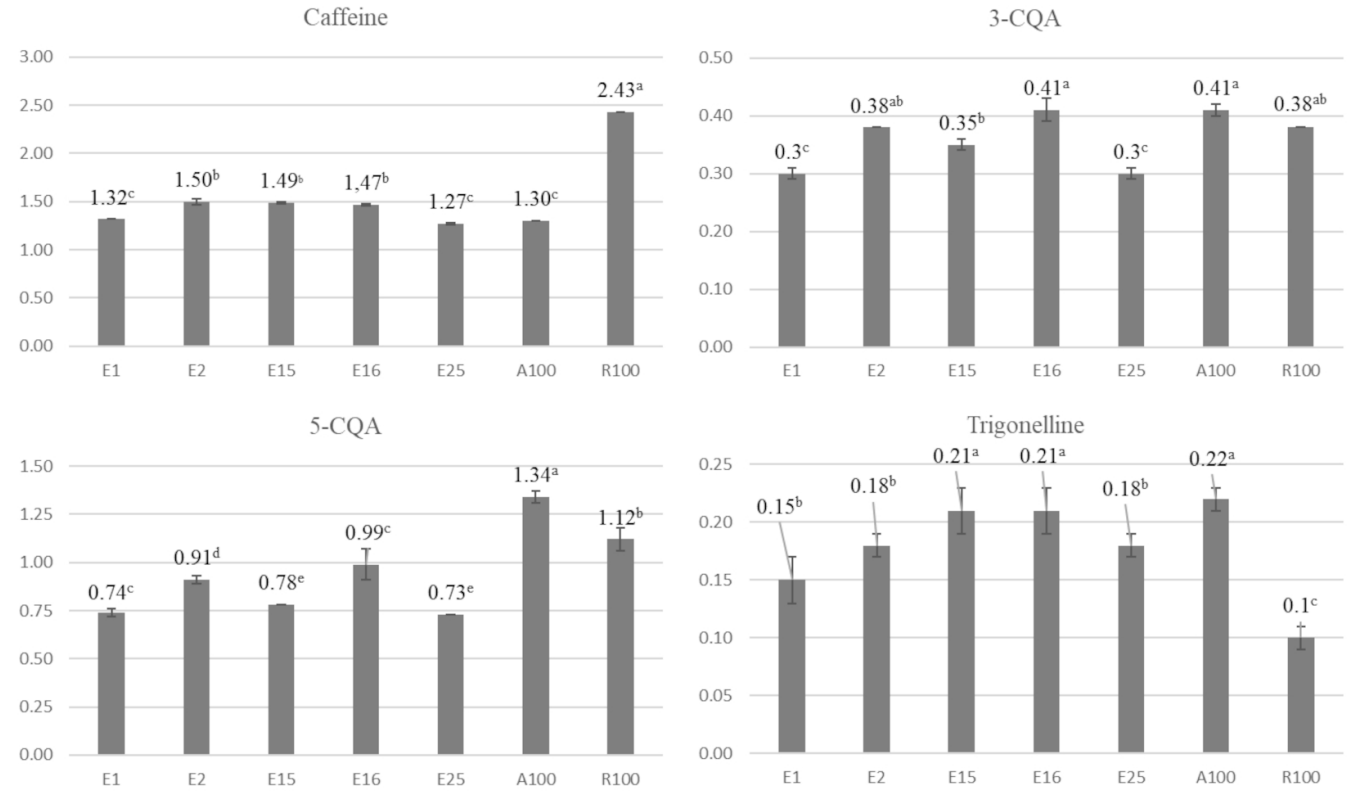

Figure 2: Levels[1] of water-soluble compounds of roasted coffee samples[2]. ([1]Means (g 100 g^{-1} db) followed by at least one same letter do not differ significantly ($p<0.05$; Tukey test). [2]Samples: Arabica 100% (A100), Robusta 100% (R100) and selections (E) (see Table 1)).

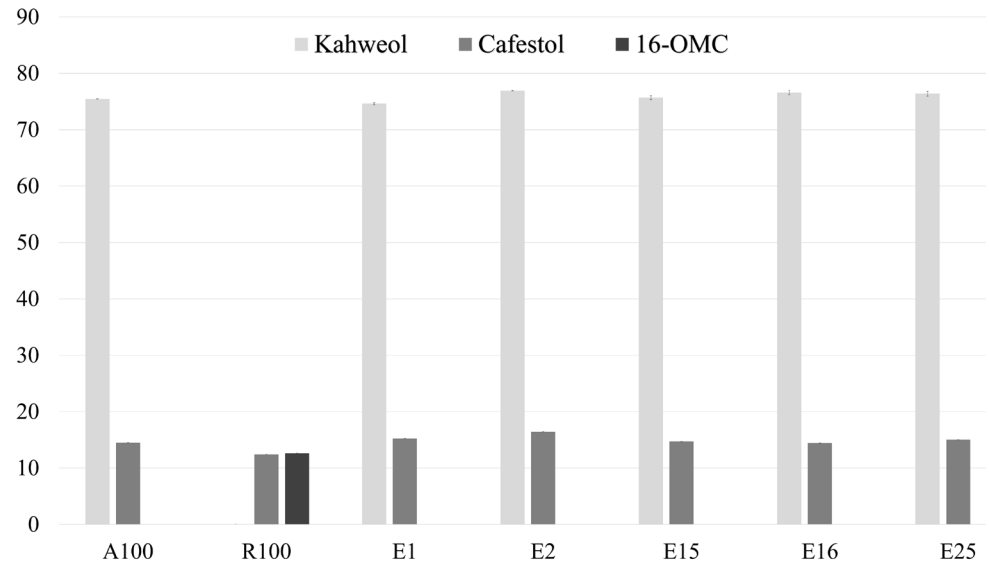

Figure 3: Level of diterpenes (mg g^{-1} of oil) in samples of roasted and ground coffee (Arabica 100%, A100; Robusta 100%, R100; and selections, E).

However, Franca et al. (2005a) found that non-defective beans presented lower caffeine contents (0.9%) than the defective ones. The concentration of 5-CQA for Arabica coffee was slightly higher than that found in Robusta coffee. Among the selections, E16 showed the greatest content of 5-CQA, and E15, E1 and E25 presented the lowest values. In general, the occurrence of defects decreases the level of this chlorogenic acid. For 3-CQA, concentrations did not differ between coffee species. However, among the selections, there was a significant discrepancy, with the highest content for E16 and the lowest

for E1 and E25, similar to that observed for 5-CQA (Figure 2). Mazzafera (1999) reported that the contents of soluble phenols and 5-CQA were approximately 35% higher in immature coffees compared to immature-black and black beans. Franca et al. (2005a) observed that black beans had significantly lower levels of 5-CQA (~ 2-time less) compared to the other samples (non-defective, black, immature, and sour coffee beans). There is no information in the literature regarding 3-CQA content and the quality of coffee beans in terms of defects.

Trigonelline content depends on the species and the quality of coffee (Figure 2). In a previous work (Dias and Benassi, 2015), the amount of trigonelline decreased while the proportion of Robusta coffee increased in blends with Arabica, as observed in the current study. Trigonelline contents were reported: 0.489 g 100 g^{-1} for Arabica coffee and 0.380 g 100 g^{-1} for Robusta (medium roasting degree, 17% weight loss) (Dias and Benassi, 2015). E15 and E16 presented higher content of trigonelline (Figure 2). There is no consensus in the literature about the levels of trigonelline among healthy coffees and coffee defects. Franca et al. (2005a) reported trigonelline levels of approximately 1% in non-defective, green (immature) and sour coffee beans and lower values (~0.8%) for black beans. However, Franca et al. (2005b) did not observe differences in trigonelline contents between high- and low-quality coffees, whereas Farah et al. (2006) found a decrease in trigonelline concentration as coffee quality is depreciated.

The literature describes approximately 14% of lipids (w/w) for green Arabica coffee beans and approximately 9.3% for green Robusta (Rubayiza and Meurens, 2005). However, roasting causes the oil concentration to increase due to water evaporation and volatiles, and oil is more thermostable than other fractions, such as carbohydrates, which are degraded by thermochemical processes such as the Maillard reaction (Dias et al., 2014; Seninde and Chambers, 2020).

The coffee species, genetic background, and technological parameters like roasting and brewing have a clear effect on coffee diterpene content (Moeenfard and Alves, 2020). However, among coffees with different quality it was not observed. The level of kahweol (~74.6 to 76.9 mg g^{-1} of oil) practically did not vary with the presence of coffee defects (Figure 3). Robusta was absent for kahweol. 16-OMC (16.2 mg g^{-1} of oil) was present only in Robusta species; cafestol content (12.4 to 16.4 mg g^{-1} of oil) subtly varied among samples and seems to be independent of species and the quality of coffee. Recently other studies confirmed that the diterpene profile is dependent on other parameters, such as genetics (Francisco et al., 2021), the roasting degree, the brew method, and the number of ingredients (Novaes et al., 2019).

Cafestol was more efficient in differentiating coffee quality. Kahweol and 16-OMC are species markers, result also found in the literature (Gunning et al., 2018). Samples of Arabica, Robusta, and of their blends, with different amounts of defects and degrees of roasting were studied in a previous work (Campanha; Dias and Benassi, 2010). Kahweol content varied between 661 and 923 mg 100 g^{-1} in the Arabica coffee, and its presence was not observed in the Conilon. Cafestol ranged from 360 to 478 mg in Arabica, and from 163 to 275 mg 100 g^{-1} in Conilon coffee. The study highlighted the potential of the parameters kahweol and cafestol for discriminating species of coffee (Campanha; Dias and Benassi, 2010).

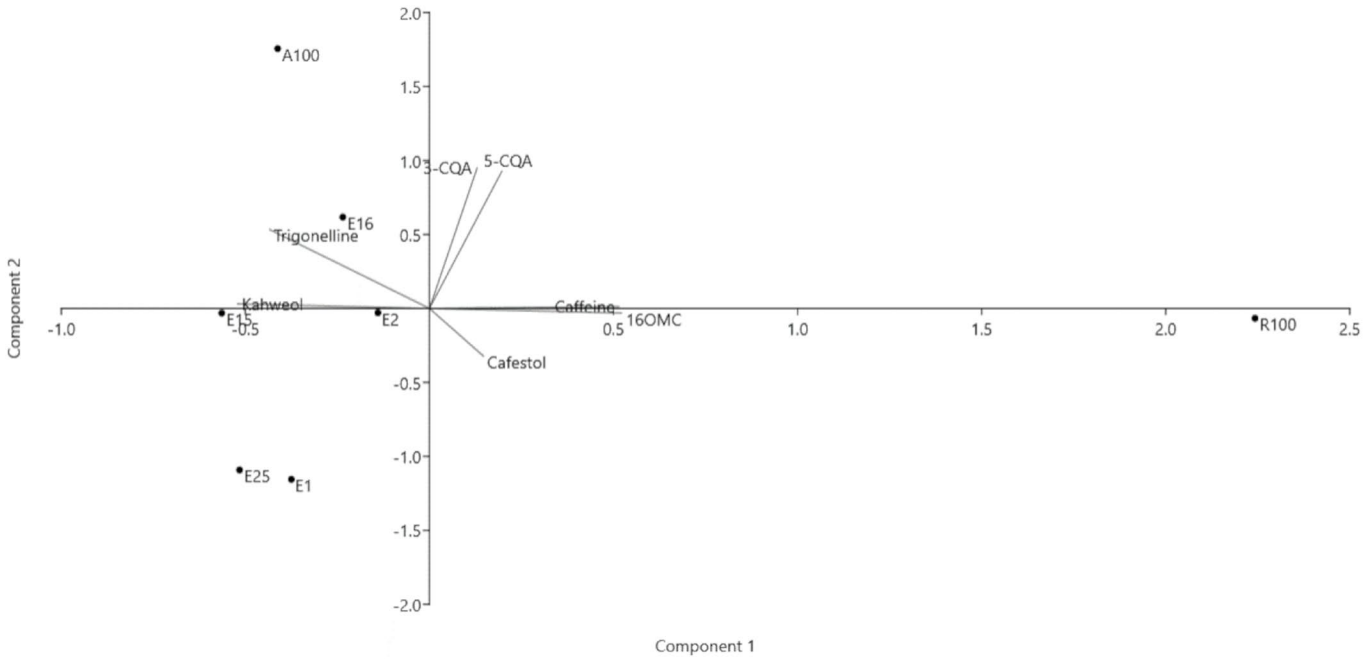

Figure 4: PCA (PC1 x PC2 plot; eigenvalue scale) for Arabica coffee (A100), Robusta coffee (R100) and selections (E).

In the Principal Component Analysis, the first two components accounted for 82,8 % of the explained variance. The water-soluble and liposoluble compounds evaluated by PCA were able to differentiate the coffee species well: A100 and R100 were plotted in opposite quadrants and far. Caffeine, 16-OMC, and kahweol were the variables with major relevance for PC 1. R100, located in the right side of the plot, presented higher values of caffeine and 16-OMC, and less content of kahweol. E15 and E2 were discriminated only by PC1. Chlorogenic acids were the compounds more important for PC 2, with positive correlation. A100 and E16 were located in the positive quadrant of PC 2, and for this reason have more 5-CQA, 3-CQA and trigonelline.

E15 has more sour beans (73,5%), being correlated to higher levels of kahweol. With almost 20% of skin, E16 was discriminated from the other samples mainly by its high levels of trigonelline and less content of cafestol. With similarly percentage of black beans (~ 14%), E1 and E25 formed a group with low concentration of both chlorogenic acids. E2 were located almost in the center of PCA. This selection has 3 to 4 times more black beans than the others, which indicates that this type of defect hinders the efficiency of the compounds in discriminating samples.

6 CONCLUSION

An optimized HPLC-DAD-MS-based method was used to determine some of the more important bioactive compounds of lipo- and water-soluble fractions of roasted coffees with different quality. Arabica and Robusta coffees, and the selections (blends of healthy Arabica beans with whole, broken, sour, black, skin, and wood) were investigated. There is a relevant discrepancy in the composition of the two species, such as in the content of caffeine, kahweol and 16-*O*-methyl cafestol. Among the selections, the main differences were detected in water-soluble compounds, such as caffeine, 5- and 3-CQA, and trigonelline. Considering all the compounds together, Principal Component Analysis promoted the discrimination of the roasted coffee samples. The selection with more sour beans (E15) was correlated to greater levels of kahweol. High contents of cafestol and trigonelline were associated with the presence of coffee skin. Low values of 5- and 3-CQA were found in the group of selections with a similar percentage of black beans (~14%) (E1 and E25). This type of defect seems to equilibrate the levels of compounds. The sample with more than 40% of black beans (E2) was practically centered on the PC1 x PC2 plotting. The assessed compounds have the potential to discriminate defects of coffee in roasted samples in selections found in coffee crops.

7 ACKNOWLEDGMENT

The authors thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the fellowship - Science without Borders Program of the Brazilian Federal Government (2015/2016).

8 AUTHORS' CONTRIBUTION

RCED wrote the manuscript, performed the experiment, and conducted all statistical analyses, SEWO supervised the experiment and co-work the manuscript, CY review and approved the final version of the work.

9 REFERENCES

AGRESTI, P. D. C. M. et al. Discrimination between defective and non-defective Brazilian coffee beans by their volatile profile. **Food Chemistry**, 106(2):787-796, 2008.

ALVES, S. T. et al. Metodologia para análise simultânea de ácido nicotínico, trigonelina, ácido clorogênico e cafeína em café torrado por cromatografia líquida de alta eficiência. **Química Nova**, 29(6):1164-1168, 2006.

ASHIHARA, H.; LUDWIG, I. A.; CROZIER, A. Bioavailability and potential impact on human health of caffeine, theobromine, and trigonelline. In: ASHIHARA, H.; LUDWIG, I. A.; CROZIER, A. **Plant nucleotide metabolism - biosynthesis, degradation, and alkaloid formation**. Weinheim: Wiley-VCH, v.1, p.397-414, 2020.

BRASIL, 2003. Instrução normativa nº 8, de 11 de junho de 2003. Ministério da Agricultura, Pecuária e Abastecimento Available in: <http://www.codapar.pr.gov.br /arquivos/File/pdf/cafebenef008_03.pdf>. Access in September, 2 2020.

CAMPANHA, F. G.; DIAS, R. C. E.; BENASSI, M. T. Discrimination of coffee species using kahweol and cafestol: effects of roasting and of defect. **Coffee Science**, 5(1):87-96, 2010.

CASAL, S. et al. Discriminate analysis of roasted coffee varieties for trigonelline, nicotinic acid, and caffeine content. **Journal of Agricultural and Food Chemistry**, 48(8):3420-3424, 2000.

CASAS, M. I. et al. Identification of biochemical features of defective *Coffea arabica* L. beans. **Food Research International**, 95(1):59-67, 2017.

DE MORAIS, S. A. L. et al. Chemical analysis of arabica coffee (*Coffea arabica* L.) and defective beans submitted to different degrees of roasting. **Coffee Science**, 2(2):97-111, 2007.

DIAS, R. C. E. et al. Data on roasted coffee with specific defects analyzed by infrared- photoacoustic spectroscopy and chemometrics. **Food Chemistry Data in Brief**, 20(1):242-249, 2018b.

DIAS, R. C. E. et al. Evaluation of kahweol and cafestol in coffee tissues and roasted coffee by a new high-performance liquid chromatography methodology. **Journal of Agricultural and Food Chemistry**, 58(1):88-93, 2010.

DIAS, R. C. E. et al. Quantitative assessment of specific defects in roasted ground coffee via infrared-photoacoustic spectroscopy. **Food Chemistry**, 255(1):132-138, 2018a.

DIAS, R. C. E. et al. Roasting process affects the profile of diterpenes in coffee. **European Food Research and Technology**, 239(6):961-970, 2014.

DIAS, R. C. E.; BENASSI, M. T. Discrimination between Arabica and Robusta coffees using hydrosoluble compounds: Is the efficiency of the parameters dependent on the roast degree? **Beverages**, 1(3):127-139, 2015.

DIVIŠ, P.; POŘÍZKA, J.; KŘÍKALA, J. The effect of coffee beans roasting on its chemical composition. **Potravinarstvo Slovak Journal of Food Sciences**, 13(1):344-350, 2019.

FARAH, A. et al. Correlation between cup quality and chemical attributes of Brazilian coffee. **Food chemistry**, 98(1):373-380, 2006.

FEIFEI, W.; TANOKURA, M. Chemical changes in the components of coffee beans during roasting. In: PREEDY, V. R. **Coffee in Health and Disease Prevention**. London: Academic Press-Elsevier. v.1, p.83-92, 2015.

FRANCA, A.; OLIVEIRA, L. Chemistry of defective coffee beans. **Food chemistry research developments**, 4(1):105-138, 2008.

FRANCA, A. S.; MENDONCA, J. C. F.; OLIVEIRA, S. D. Composition of green and roasted coffees of different cup qualities. **LWT - Food Science and Technology**, 38(1):709-715, 2005b.

FRANCA, A. S. et al. Physical and chemical attributes of defective crude and roasted coffee beans. **Food Chemistry**, 90(1):84-89, 2005a.

FRANCISCO, J. S. et al. Natural intervarietal hybrids of *Coffea canephora* have a high content of diterpenes. **Beverages**, 7(77):1-9, 2021.

GIACALONE, D. et al. Common roasting defects in coffee: Aroma composition, sensory characterization and consumer perception. **Food Quality and Preference**, 71(1):463-474, 2018.

GUNNING, Y. et al. 16-*O*-methylcafestol is present in ground roast Arabica coffees: Implications for authenticity testing. **Food Chemistry**, 248(1):52-60, 2018.

HAILE, M.; KANG, W. H. The harvest and post-harvest management practices' impact on coffee quality. In: CASTANHEIRA, D. T. **Coffee - production and research**. London: IntechOpen Limited. v.1, p.59-76, 2019.

HAMMER, Ø.; HARPER, D. A. T.; RYAN, P. D. 2001. PAST: Paleontological Statistics software package for education and data analysis. **Paleontologia Electronica**, 4(1):9, 2001.

JEON, J. S et al. Contents of chlorogenic acids and caffeine in various coffee-related products. **Journal of Advanced Research**, 17(1):85-94, 2019.

JESZKA-SKOWRON, M.; FRANKOWSKI, R.; ZGOŁA-GRZEŚKOWIAK, A. Comparison of methylxantines, trigonelline, nicotinic acid and nicotinamide contents in brews of green and processed Arabica and Robusta coffee beans - Influence of steaming, decaffeination and roasting processes on coffee beans. **LWT**, 125(1):1- 9, 2020.

MAZZAFERA, P. Chemical composition of defective coffee beans. **Food Chemistry**, 64(1):547-554, 1999.

MENDONÇA, J. C. F.; FRANCA, A. S.; OLIVEIRA, L. S. Physical characterization of non-defective and defective Arabica and Robusta coffees before and after roasting. **Journal of Food Engineering**, 92(4):474-479, 2009.

MOEENFARD, M.; ALVES, A. New trends in coffee diterpenes research from technological to health aspects. **Food Research International**, 134(1):109207, 2020.

NARITA, Y.; INOUYE, K. Chlorogenic acids from coffee. In: PREEDY, V. R. **Coffee in Health and Disease Prevention**. London: Academic Press-Elsevier. v.1, p.189-199, 2015.

NOGUEIRA, M.; TRUGO, L. C. Distribuição de isômeros de ácido clorogênico e teores de cafeína e trigonelina em cafés solúveis brasileiros. **Food Science and Technology**, 23(2):296-299, 2003.

NOVAES, F. J. M. et al. The occurrence of cafestol and kahweol diterpenes in different coffee brews. **Coffee Science**, 14(2):265-280, 2019.

PERRONE, D.; DONANGELO, C. M.; FARAH, A. Fast simultaneous analysis of caffeine, trigonelline, nicotinic acid and sucrose in coffee by liquid chromatography-mass spectrometry. **Food Chemistry**, 110(4):1030-1035, 2008.

RUBAYIZA, A. B.; MEURENS, M. Chemical discrimination of Arabica and Robusta coffees by Fourier transform Raman spectrometry. **Journal of Agricultural and Food Chemistry**, 53(12):4654-4659, 2005.

SCA - SPECIALTY COFFEE ASSOCIATION. Protocols and best practices. Available in: <https://sca.coffee/research/protocols-best-practices?page=resources&d=coffee-protocols>. Access in: September, 28 2022.

SCAA. **Specialty Coffee Association**. 2022. Available in: <http://www.coffeeresearch.org/coffee/scaaclass.htm>. Access in: September, 28 2022.

SENINDE, D. R.; CHAMBERS, E. Coffee flavor: A review. **Beverages**, 6(3):44-69, 2020.

SHAN, Y. et al. Simultaneous determination of chlorogenic acids in green coffee bean extracts with effective relative response factors. **International Journal of Food Properties**, 20(9):2028-2040, 2017.

SPEER, K.; KÖLLING-SPEER, I. The lipid fraction of the coffee bean. **Brazilian Journal of Plant Physiology**, 18(1):201-216, 2006.

VIGNOLI, J. A. et al. Roasting process affects differently the bioactive compounds and the antioxidant activity of arabica and robusta coffees. **Food Research International**, 61(1):279-285, 2014.

WERMELINGER, T. et al. Quantification of the Robusta fraction in a coffee blend via Raman spectroscopy: Proof of principle. **Journal of Agricultural and Food Chemistry**, 59(17):9074-9079, 2011.

WINTGENS, J. N. Factors influencing the quality of green coffee. In: WINTGENS, J.N. **Coffee: Growing, processing, sustainable production. A guidebook for growers, processors, traders and researchers**. Weinheim: Wiley-VCH. v.1, p.789-809, 2004.