NMR QUANTIFICATION OF 16-O-METHYLCAFESTOL AND KAHWEOL IN *Coffea canephora* var. robusta BEANS FROM DIFFERENT GEOGRAPHICAL ORIGINS

C. Finotello[1], C. Forzato[2], A. Gasparini[2], S. Mammi[1], L. Navarini[3], E. Schievano[1*]

[1]*Department of Chemical Sciences, University of Padova, via Marzolo 1, 35131 Padova, Italy*

[2]*Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, via L. Giorgieri 1, 34127 Trieste, Italy*

[3]*illycaffè S.p.A., via Flavia 110, 34147 Trieste, Italy*

**Corresponding author:* Department of Chemical Sciences, University of Padova, via Marzolo 1, 35131 Padova, Italy. Tel.: +39 049 8275742; fax: +39 049 8275829. E-mail address: elisabetta.schievano@unipd.it (E. Schievano)

Abstract

Diterpenes have recently received a great deal of interest as tools to investigate the botanical origin of coffee. Specifically, kahweol has been proposed as a marker of *Coffea arabica* while 16-O-methylcafestol (16-OMC) is a *Coffea canephora* specific marker and its detection and quantification allow the authenticity of pure *C. arabica* roasted coffee blends to be assessed. In this study, we evaluated the possibility of the industrial use of the quantification of these diterpenes to assess the relative amounts of the two coffee species in blends. The content of 16-OMC and kahweol was determined in 78 samples (i.e., 39 green and the corresponding 39 roasted beans) of *C. canephora* from different geographical origins using a recently published NMR approach. Our results show a small natural variability in 16-OMC content for the Asian samples (average content = 1837 ± 113 mg/kg) while a much larger spread was found for the African samples (average content = 1744 ± 322 mg/kg). This large variability prevents the use of 16-OMC to quantify *C. canephora* in unknown roasted coffee blends. We also show that kahweol cannot be considered a specific *C. arabica* marker since it was detected almost all coffees and quantified in about 30% of the *C. canephora* samples.

Keywords: 16-O-methylcafestol, 16-OMC; robusta coffee, cafestol, kahweol, qNMR

Abbreviations used: 16-O-methylcafestol, 16-OMC; robusta coffee, cafestol, kahweol, qNMR

1. Introduction

The genus *Coffea* L. includes the three coffee species used to produce one of the world's most popular beverages: *C. arabica* (Arabica coffee), *C. canephora* (Robusta coffee), and *C. liberica* (Liberica coffee, or Excelsa coffee) (Davis, Govaerts, Bridson, & Stoffelen, 2006). *C. arabica* is by far the most important commercial species and the only one grown up to the beginning of the 20th century, whereas *C. liberica* represents less than 1% of the marketed coffee. *C. canephora*, described by the French botanist Pierre in 1879, was first introduced into Indonesia (Java) from Congo in 1900 (via Belgium) because of its resistance to the disease known as coffee leaf rust (Waller, J.M. , Bigger, M., Hillocks, 2007). This coffee disease, starting from 1869 in Sri Lanka, in 20 years had virtually wiped out the cultivation of coffee (*C. arabica* only at that time) in Asia (Illy & Viani, 2005). In 1950, Robusta coffee accounted for about 13% of the world's coffee production; in 1989, it reached about 30% and in 2015, it had increased to 45% (ICO, 2015). In spite of this growth, mainly due to lower production costs and higher yields, the much richer and smoother aroma and flavor of beverages derived from Arabica coffee are more appreciated than those from Robusta coffee, which, for this reason, is still characterized by significantly lower prices. This fact opens the possibility of commercial frauds aimed at tainting the authenticity of 100% Arabica blends by deliberate and undeclared addition of *C. canephora*. The detection and quantification of such fraudulent blending in commercial samples is therefore important to protect consumers. To this end, sensory analysis may be insufficient, particularly at low (<20%) Robusta addition levels (Wermelinger, D'Ambrosio, Klopprogge, & Yeretzian, 2011). DNA-based techniques have been proposed to differentiate between Arabica and Robusta although roasted coffee still represents a challenging matrix when compared to the green raw material (Spaniolas, Tsachaki, Bennett, & Tucker, 2008; Trantakis et al., 2012).

Due to compositional differences between Arabica and Robusta, several approaches have been proposed and applied to discriminate these two coffee species chemically and compounds such as

caffeine and trigonelline (Casal, Oliveira, Alves, & Ferreira, 2000; Ky et al., 2001), amino acid enantiomers (Casal, Alves, Mendes, Oliveira, & Ferreira, 2003), volatile compounds (Hovell, Pereira, Arruda, & Rezende, 2010), homostachydrine (Servillo et al., 2016) as well as metal content (Grembecka, Malinowska, & Szefer, 2007; Martín, Pablos, & González, 1999) have been suggested as reliable discriminants. In this regard, the coffee lipid fraction has been the subject of abundant studies, because many of its components, such as fatty acids (María J. Martín, Pablos, González, Valdenebro, & León-Camacho, 2001; Romano et al., 2014; Rui Alves, Casal, Oliveira, & Ferreira, 2003), tocopherols (Cizkova, Soukupova, Voldrich, & Sevcik, 2007; Mariani & Fedeli, 1991; Pablos, González, Martín, Valdenebro, & León-Camacho, 1999), triglycerides (Alves, Casal, Alves, & Oliveira, 2009; González, Pablos, Martín, León-Camacho, & Valdenebro, 2001) and diterpenes (Pacetti, Boselli, Balzano, & Frega, 2012; Rubayiza & Meurens, 2005), could be used to differentiate the two coffee species. Several compounds proposed as markers investigated so far are present in both coffee species (Mariani & Fedeli, 1991; Servillo et al., 2016) with obvious limitations in their use to accurately determine blend composition. Within coffee sterols, for instance, when Δ^5-avenasterol and 24-methylene-cholesterol (higher level in Robusta) were used to quantify Arabica and Robusta amount in roasted coffee blends, the limit of detection for Robusta addition was reported as 30% and 15%, respectively (Mariani & Fedeli, 1991). For authenticity purposes, much more interesting is the potential offered by the chemical class of coffee diterpenes (Speer & Kolling-Speer, 2001), the most important of which are kahweol, cafestol, and 16-O-methylcafestol (16-OMC). These compounds are mostly esterified with various fatty acids and only a small amount is present in the free form (De Angelis et al., 2014). Cafestol is present in both species while 16-OMC is present exclusively in Robusta, according to the available literature (Speer & Kolling-Speer, 2001). The exclusive presence of 16-OMC in Robusta and its thermal stability make it an excellent marker to quantify the composition of blends and to detect possible frauds, but the reported variability of this compound in Robusta samples might limit its industrial use (Speer & Kolling-Speer, 2001).

On the other hand, kahweol has been proposed as a specific marker of *C. arabica*, since several studies indicated that it is almost absent in *C. canephora* (Campanha, Dias, & Benassi, 2010; Keidel, von Stetten, Rodrigues, Máguas, & Hildebrandt, 2010; Rubayiza & Meurens, 2005; Souza & Benassi, 2012; Wermelinger et al., 2011). This point, however, is still a subject of debate, spurred by the limited screening studies to ascertain the presence and the content of kahweol in *C. canephora*.

We recently developed a high-resolution ^{1}H NMR method (Schievano, Finotello, De Angelis, Mammi, & Navarini, 2014) that provides a quantitative determination of both free and esterified 16-OMC directly in coffee extracts and proved to be reliable to detect cafestol and kahweol as well. The major advantages of the NMR method with respect to the official methods till now adopted to determine 16-OMC ("Deutsches Institut für Normung e.V. (DIN). Standard DIN 10779, Analysis of Coffee and Coffee Products, Determination of 16-OMethyl Cafestol Content of Roasted Coffee HPLC Method; DIN: Berlin, Germany, 2011," n.d.; Pizarro, Esteban-Díez, & González-Sáiz, 2007; Wermelinger et al., 2011) are: minimal sample preparation required, the absence of any derivatization step, very short time of analysis, reproducibility, and provision of quantitative and structural information. The official methods are in fact so laborious and scarcely reproducible that a wide and systematic investigation of the amounts of 16-OMC and kahweol in both green and roasted Robusta samples has never been undertaken and literature data are scarce.

In the present study, the NMR method was applied to 39 samples of commercial lots of *C. canephora* from different geographical origins to investigate the variability of 16-OMC. The goal is to evaluate the possibility to apply the quantification of 16-OMC for industrial purposes. Cafestol and kahweol were also quantified.

2. Materials and Methods

2.1 Chemicals

Deuterated chloroform ($CDCl_3$, 99.96%D), stabilized with silver, and deuterated water (≥99.96%D) were purchased from Euriso-Top (Gif sur Yvette, France). N,N-dimethylformamide (DMF, ≥99.99% (GC)), 16-OMC (≥90%), palmitic acid (≥99%), N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC), 4-(dimethylamino)pyridine (DMAP), triethylamine (Et_3N), anhydrous CH_2Cl_2, 5 mm precision glass NMR tubes (535-pp, Wilmad) and coaxial inserts (wgs-5bl, Wilmad) were purchased from Sigma-Aldrich (Milan, Italy). Cafestol for the synthesis of the palmitate ester was purchased from Vinci-Biochem s.r.l. (Florence, Italy) while 16-OMC for the synthesis of the palmitate ester, due to its very high cost, was obtained by extraction from Robusta coffee beans using a modified DIN method (Guercia, Berti, Navarini, Demitri, & Forzato, 2016).

2.2 Coffee samples

Thirty-nine *C. canephora* green coffee samples from commercial lots, in whole beans, from Africa and Asia, were kindly provided by Sandalj Trading Company S.p.A., Trieste (Italy) and used as received. All the samples were harvested between 2014 and 2015. Details on the coffee origin are reported in Table 1. An aliquot of the green coffee beans (100 g) was roasted to a medium roasting degree (total weight loss of 15.9 ± 3.9 g 100 g^{-1}) in a Probat (Germany) lab roaster.

2.3 Synthesis and characterization of Cafestol palmitate and 16-OMC palmitate standards

2.3.1 Cafestol palmitate

To a solution of palmitic acid (16 mg, $6.3•10^{-5}$ mol, 1 eq) in 1.5 mL of anhydrous CH_2Cl_2, the following reagents were added under stirring: cafestol (20 mg, $6.5•10^{-5}$ mol, 1 eq), N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC•HCl)EDC•HCl (36 mg, $1.88•10^{-4}$ mol, 3 eq), Et_3N (0.018 mL, $1.3•10^{-4}$ mol, 2 eq) and 4-(Dimethylamino)pyridine (DMAP) (12 mg, $9.48•10^{-5}$ mol, 1.5 eq). The reaction was left under stirring overnight. After the addition of few milliliters of CH_2Cl_2, the organic phase was washed with 5% $KHSO_4$, distilled water, and 5% $NaHCO_3$. The organic phase was dried over anhydrous Na_2SO_4. Evaporation of the solvent gave the

crude product which was further purified by flash chromatography (Merck silica gel 60, 0.040-0.063 mm, 230-400 mesh ASTM) using petroleum ether/ethyl acetate from 90:10 v/v to 50:50 v/v as eluent. Cafestol palmitate (0.014 g) was obtained as a pale yellow solid in 48% yield.

R_f = 0.54 (petroleum ether/ethyl acetate 80:20 v/v, TLC stained with aqueous permanganate solution - Merck 60 F_{254} silica gel plates).

One dimensional ^1H NMR, ^{13}C NMR, and 2D-NMR (^1H-^1H COSY, ^1H-^{13}C HSQC) spectra were obtained on a Varian 500 spectrometer operating at 500MHz. One dimensional ^1H NMR, ^{13}C NMR were obtained with a relaxation delay of 1s. ^1H-^1H COSY and ^1H-^{13}C HSQC were obtained with a relaxation delay of 1s, acquisition time 0.150s and an observe pulse of 6.50μs at a power of 56. Chemical shifts are reported in ppm using the solvent residual signal as the internal reference ($CDCl_3$: ^1H = 7.27 ppm, ^{13}C = 77.0 ppm).

The ^1H NMR, ^{13}C NMR, and 2D-NMR spectra were processed using the MestReNova 10.0 software.

Electrospray Ionization mass spectrometry measurements ($ESI^{+/-}$-MS) were performed with an Esquire 4000 (Bruker-Daltonics) spectrometer.

^1H NMR (500 MHz, $CDCl_3$, ppm) δ = 7.25 (1H, d, J=1.9, H_{19}), 6.22 (1H, d, J=1.9, H_{18}), 4.28 (2H, AB system, H_{17}), 2.63 (2H, m, H_2), 2.37 (3H2H, t, J=7.4, $H_{23}CH_2CO$), 2.27 (1H, m, H_5), 2.09-2.03 (3H, m, H_1+H_{13}+H_{14a}), 1.82 (2H1H, dq, J_1=3.1, J_2=12.9, H_{6a}) 1.75-1.50 (11H12H, m, H_{6b}+H_7+H_{11}+H_{12}+H_{14}+H_{15}+2Hchain), 1.36-1.18 23 (26H, m, H-chain 26H), 1.18 (1H, m, H_9), 0.89 (3H, t, J=6.9, CH_3 chain), 0.84 (3H, s, H_{20}).

^{13}C NMR (125 MHz, $CDCl_3$, ppm) δ = 174.22 (s, C=O), 148.72 (s, C_3), 140.56 (d, C_{19}), 120.07 (s, C_4), 108.27 (d, C_{18}), 80.11 (s, C_{16}), 68.27 (t, C_{17}), 53.23 (t, C_{15}), 52.04 (d, C_9), 46.04 (d, C_{13}), 44.75 (s, C_8), 44.23 (d, C_5), 40.71 (t, C_7), 38.62 (s, C_{10}), 38.04 (t, C_{14}), 35.71 (t, C_1), 34.30 (t, CH_2C=O), 31.90 (t, chain), 29.68 (2t, chain), 29.67 (t, chain), 29.64 (2t, chain), 29.58 (t, chain), 29.46 (t,

chain), 29.35 (t, chain), 29.24 (t, chain), 29.14 (t, chain), 26.05 (t, C_{12}), 25.03 (t, chain), 23.08 (t, C_6), 22.68 (t, chain), 20.61 (t, C_2), 18.92 (t, C_{11}), 14.11 (q, CH_3 chain), 13.31 (q, C_{20}).

ESI^+-MS: $[M+Na]^+$ = 577.6 m/z.

2.3.2 16-OMC palmitate

To a solution of palmitic acid (16 mg, $6.3 \cdot 10^{-5}$ mol, 1 eq) in 1.5 mL of anhydrous CH_2Cl_2, the following reagents were added under stirring: 16-O-methylcafestol (20 mg, $6.1 \cdot 10^{-5}$ mol, 1 eq), EDC•HCl (36 mg, $1.82 \cdot 10^{-4}$ mol, 3 eq), Et_3N (0.017 mL, $1.21 \cdot 10^{-4}$ mol, 2 eq) and DMAP (11 mg, $9.08 \cdot 10^{-5}$ mol, 1.5 eq). The reaction was left under stirring overnight. After the addition of few milliliters of CH_2Cl_2, the organic phase was washed with 5% $KHSO_4$, distilled water, and 5% $NaHCO_3$. The organic phase was dried over anhydrous Na_2SO_4. Evaporation of the solvent gave the crude product which was further purified by flash chromatography using petroleum ether/ethyl acetate from 90:10 v/v to 50:50 v/v as eluent. 16-OMC palmitate (0.016 g) was obtained as a light brown oil in 46% yield.

R_f = 0.90 (petroleum ether/ethyl acetate 80:20 v/v, TLC stained with aqueous permanganate solution).

^1H NMR (500 MHz, $CDCl_3$, ppm) δ = 7.25 (1H, d, J=1.8, H_{19}), 6.21 (1H, d, J=1.8 Hz, H_{18}), 4.45 (1H, d, J=12.5 Hz, H_{17}), 4.24 (1H, d, J=12.5, H_{17}), 3.17 (3H, s, H_{21}), 2.63 (2H, m, H_2), 2.36 (2H, t, J=7.5 Hz, CH_2C=O), 2.28 (2H, m, H_5+H_{13}), 2.07 (1H, m, H_{1a}), 1.99 (1H, d, J=11.6, H_{14a}), 1.81 (2H1H, dq, J_1=3.1, J_2=12.9, H_{6a}), 1.76-1.47 (1312H, m, H_{6b}+H_7+H_{11}+H_{12}+H_{14b}+H_{15}+2Hchain), 1.36-1.19 (24H26H, m, H_{1b}+H_9+Hchain), 0.89 (3H, t, J=6.9, CH_3chain), 0.84 (3H, s, H_{20}).

^{13}C NMR (125 MHz, $CDCl_3$, ppm) δ = 174.12 (s, $C_{22}C$=O), 148.74 (s, C_3), 140.55 (d, C_{19}), 120.10 (s, C_4), 108.28 (d, C_{18}), 84.69 (s, C_{16}), 62.56 (t, C_{17}), 52.16 (d, C_9), 49.70 (t, C_{15}), 49.56 (q, C_{21}), 44.42 (s, C_8), 44.24 (d, C_5), 42.12 (d, C_{13}), 40.91 (t, C_7), 38.66 (s, C_{10}), 37.84 (t, C_{14}), 35.75 (t, C_1), 34.37 (t, $C_{23}CH_2CO$), 31.91 (t, chain), 29.69 (2t, chain), 29.68 (t, chain), 29.65 (2t, chain), 29.59 (t, chain), 29.47 (t, chain), 29.35 (t, chain), 29.25 (t, chain), 29.14 (t, chain), 26.00 (t, C_{12}), 25.06 (t,

chain), 23.10 (t, C_6), 22.68 (t, chain), 20.62 (t, C_2), 19.09 (t, C_{11}), 14.11 (q, C32CH_3 chain), 13.29 (q, C_{20}).

ESI^+-MS: $[M+Na]^+$ = 591.7 m/z, $[M+H]^+$ = 569.6 m/z.

2.4 Sample Coffee extraction

About 15 g of coffee beans, green or roasted, were ground in liquid nitrogen for 1 minute using a commercial household coffee grinder. The powder obtained was extracted using the method previously described (Schievano et al., 2014): 0.1530 g (± 0.0015) of powder and 1.5 mL of accurately weighted $CDCl_3$ were combined. The mixture was vortexed (vortex Heidolph, Multi Reax) for 15 min (nome vortex) and then quickly filtered through cotton wool directly in the NMR tube.

2.5 NMR Spectroscopy for diterpenes quantification

^1H NMR spectra were acquired with a Bruker (Rheinstetten, Germany) Avance DMX600 spectrometer operating at 599.90 MHz for ^1H and equipped with a 5 mm TXI xyz-triple gradient probe. A standard pulse-acquire experiment was used with a spectral width of 6000 Hz and 32768 data points; the number of acquired scans varied from 16 to 48 depending on the analyte concentration, for total measurement times ranging from 13 to 40 minutes. To produce quantitative data, a relaxation delay of 43 s was used, corresponding to five times the longest T_1 longitudinal relaxation time (for DMF methyls in D_2O) (Schievano et al., 2014).

The ^1H spectra were processed using the ACD software (ACD labs 12.0). Fourier transformation was performed after exponential line-broadening of 0.3 Hz, and the spectra were calibrated on the residual signal of $CHCl_3$ set to 7.27 ppm. Integrations were manually obtained after careful manual phase and baseline correction.

Absolute concentrations were determined as described in our previous work (Schievano et al., 2014): a coaxial insert filled with a $8.15 \cdot 10^{-4}$ M standard solution of DMF in D_2O, was placed in the

NMR tube as external reference. The molar concentration of 16-OMC esters was calculated by taking the integral ratio of the methyl signal (H_{21}, 3.17 ppm) and of the low-field methyl signal of DMF. The total molar concentration of the esters of cafestol and of 16-OMC was determined by integrating the signal at 6.21 ppm where protons H_{18} for both metabolites resonate. The cafestol content was estimated by calculating the difference of the total amount and the 16-OMC amount. Final concentrations of 16-OMC esters and of cafestol are reported in mg of analyte per kg of powder, considering the MW of the analytes in the free form (i.e., not esterified). In a few cases, kahweol signals were also detected in the spectra, but they were not considered because of their low intensity. To estimate kahweol content, ^{1}H NMR experiments were performed incrementing the number of scans to 256. The concentration was determined by integration of the H_{18} doublet signal at 6.25 ppm.

Repeatability of the entire analytical procedure, in terms of RSD, was tested on one sample of roasted Robusta coffee. Three different NMR samples were prepared, one spectrum was acquired per preparation, and each spectrum was processed three times. The standard deviation is about 0.2% taking three different integrations on the same sample while considering the integration of three different preparations, the standard deviation increases to 0.4%.

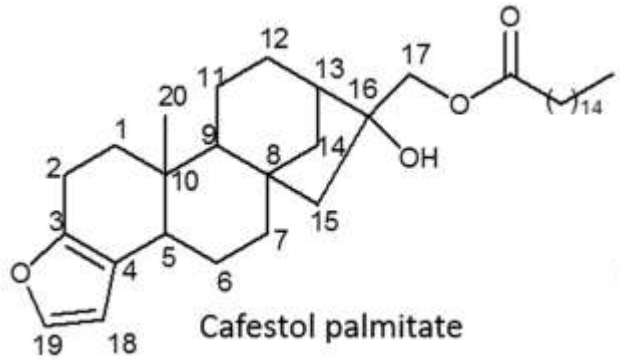
3. 3. Results and Discussion

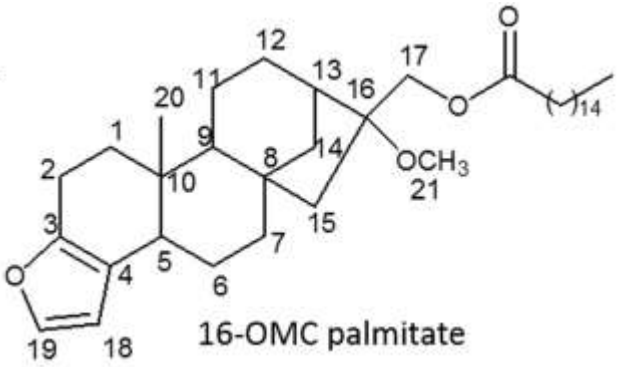
3.1 Analyte assignment confirmation in the extracts

To confirm the previously assignments of ^{1}H-NMR signals of the palmitate esters of both cafestol and 16-OMC in coffee extracts (Schievano et al., 2014), the two compounds have been synthetized using a literature procedure for the esterification of paraconic acids (Berti et al., 2006) by a condensation reaction between palmitic acid and cafestol (or 16-OMC) using EDC·HCl, DMAP and Et_3N in CH_2Cl_2 at room temperature. The two esters, represented in Fig. 1, were fully

characterized spectroscopically using a 500MHz NMR but ^{1}H-NMR with a 600MHz spectrometer were also registered in order to compare their spectra with those obtained on the extracts.displays the chemical structure of cafestol and 16-OMC esters with the carbon atom numbering. The spectra of cafestol palmitate, 16-OMC palmitate, 16-OMC, and of the chloroform extract of a roasted Robusta coffee are shown in Fig. 2. The singlet of H_{21} falls at 3.17 ppm both in the 16-OMC palmitate (Fig. 2b) and in the extract whereas a slight shift is observed in free 16-OMC (3.18 ppm, Fig. 2c). Protons H_{17} form an AB system in cafestol palmitate at 4.29 ppm (figure 2a) whereas in 16-OMC palmitate, protons H_{17} form an AX system with two doublets at 4.44 ppm and 4.23 ppm (figure 2b). In the extract, protons H_{17} of esterified cafestol partially overlap with the signals of the glycerol moiety of triglycerides.

As shown in Fig. 2, the resonance of methyl H_{21} at 3.17 ppm used for quantification and the two H_{17} protons lines at 4.45 and 4.24 ppm are diagnostic signals of a Robusta-containing extract. These resonances were previously assigned in investigating coffee lipophilic extracts by 2D experiments (Schievano et al., 2014). This assignment is now fully confirmed by comparison with the ^{1}H spectrum of the synthetized 16-OMC palmitate used as a standard (see paragraph 2.3).





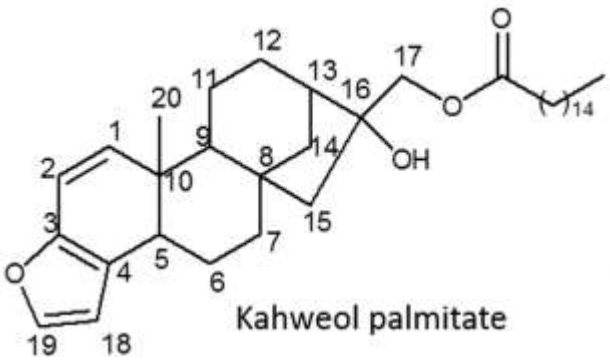


Figure 1 Left: Cafestol and right: 16-OMC esterified with palmitate

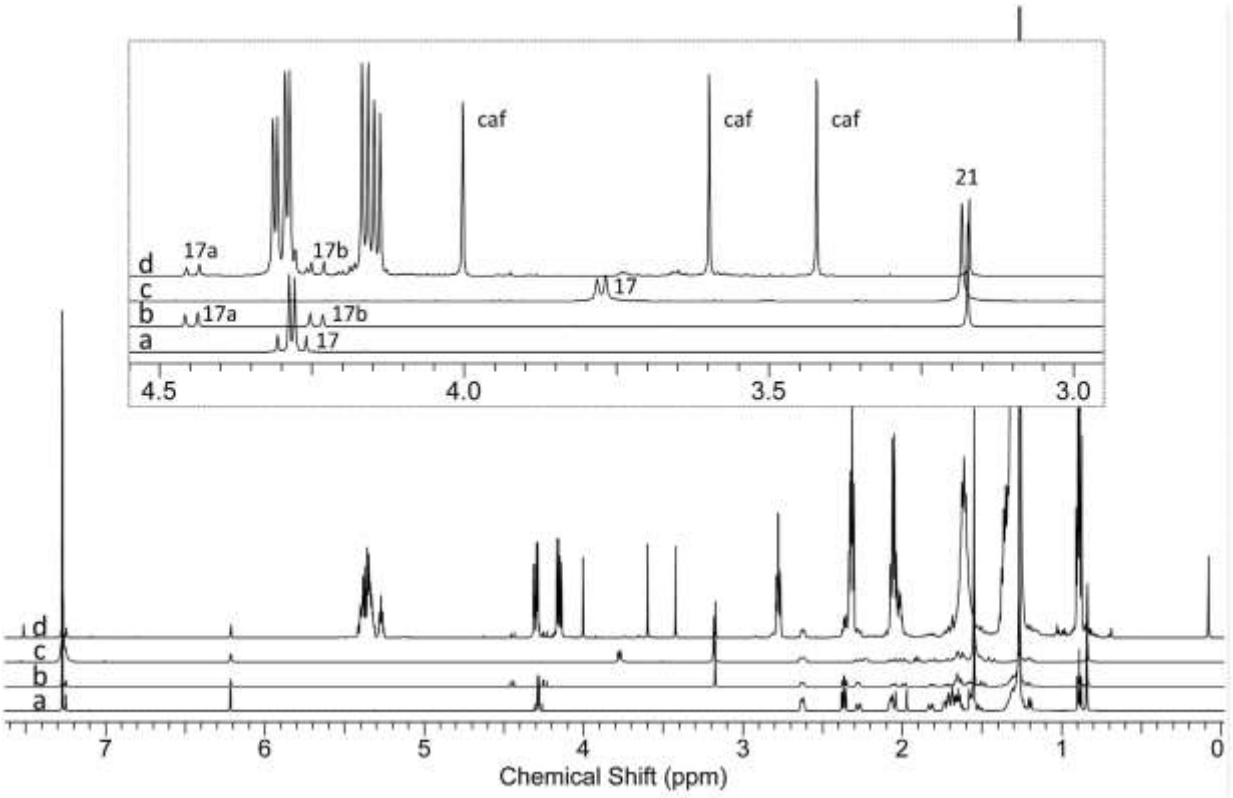


Figure 2 ^{1}H NMR spectra of cafestol palmitate (a), 16-OMC palmitate (b), 16-OMC (c) and of the chloroform extract of a roasted Robusta coffee (d). The spectral region between 2.9 ppm and 4.6 ppm is enlarged in the insert

3.2 Quantification

NMR spectra of coffee extracts were acquired on freshly prepared samples to avoid partial degradation of the analyte in chloroform, as typically revealed by the appearance of signals close to the 16-OMC palmitate resonances, both for H_{21} and H_{17} (see Fig. 1S). To slow down this degradation process, $CDCl_3$ stabilized with silver was used and the samples were stored at 4 °C in the dark. Under these conditions, the first signals of degradation products appeared only one week after sample preparation.

Quantification of 16-OMC esters in the roasted coffee extracts was carried out by integrating the methyl signal at 3.17 ppm (Table 1). The content of 16-OMC esters was between 2236 ± 16 mg/kg and 1204 ± 6 mg/kg for the set of 39 samples, with an average content of 1806 ± 208 mg/kg.

Table 1. Contents of 16-O-methylcafestol, Cafestol and Kahweol in *Coffea canephora* from different geographical origins

Origin	Roasted					
	16-OMC	SD	Cafestol	SD	Kahweol	SD
	mg/kg					
AFRICA	1793	20	1879	36		
	1964	12	2256	16		
	1204	6	1876	4	197	4
	1390	7	2145	52		
	1649	13	1875	19		
	1844	20	1817	51		
	1257	4	1791	23		
	1622	5	2058	33		
	2236	16	2816	48		
	1706	14	2445	44	34.6	0.5
	1906	4	1698	27		
	2208	4	2725	35		
	1889	13	2607	29		
INDIA	1804	2	1805	18	76	1
	1749	5	1494	6	143.4	1.6
	1774	5	1710	50		
	1807	7	1666	23		
	1849	13	2077	31		
	1774	9	1781	34		
	1867	20	1577	39		
	1844	4	2116	9	62.6	0.8
	2007	11	2557	13		
	1719	5	1702	19		
INDONESIA	2123	15	2053	16	99.2	0.8
	1731	9	1562	14	65.9	0.9
	1837	20	1297	10		
	1979	4	1939	21		
	1873	11	1616	44	25.4	0.4
	1921	8	1811	24	62.9	0.8
VIETNAM	1892	8	1683	44		
	1848	7	1830	11		
	1933	12	2233	11		
	1863	4	1623	44		
	1907	23	2510	53	200.4	1.4
	1895	15	1484	31		
	1820	24	2025	41		
	1714	17	1693	9		
LAOS	1606	19	1526	20	49.5	0.6
	1619	16	2141	30		

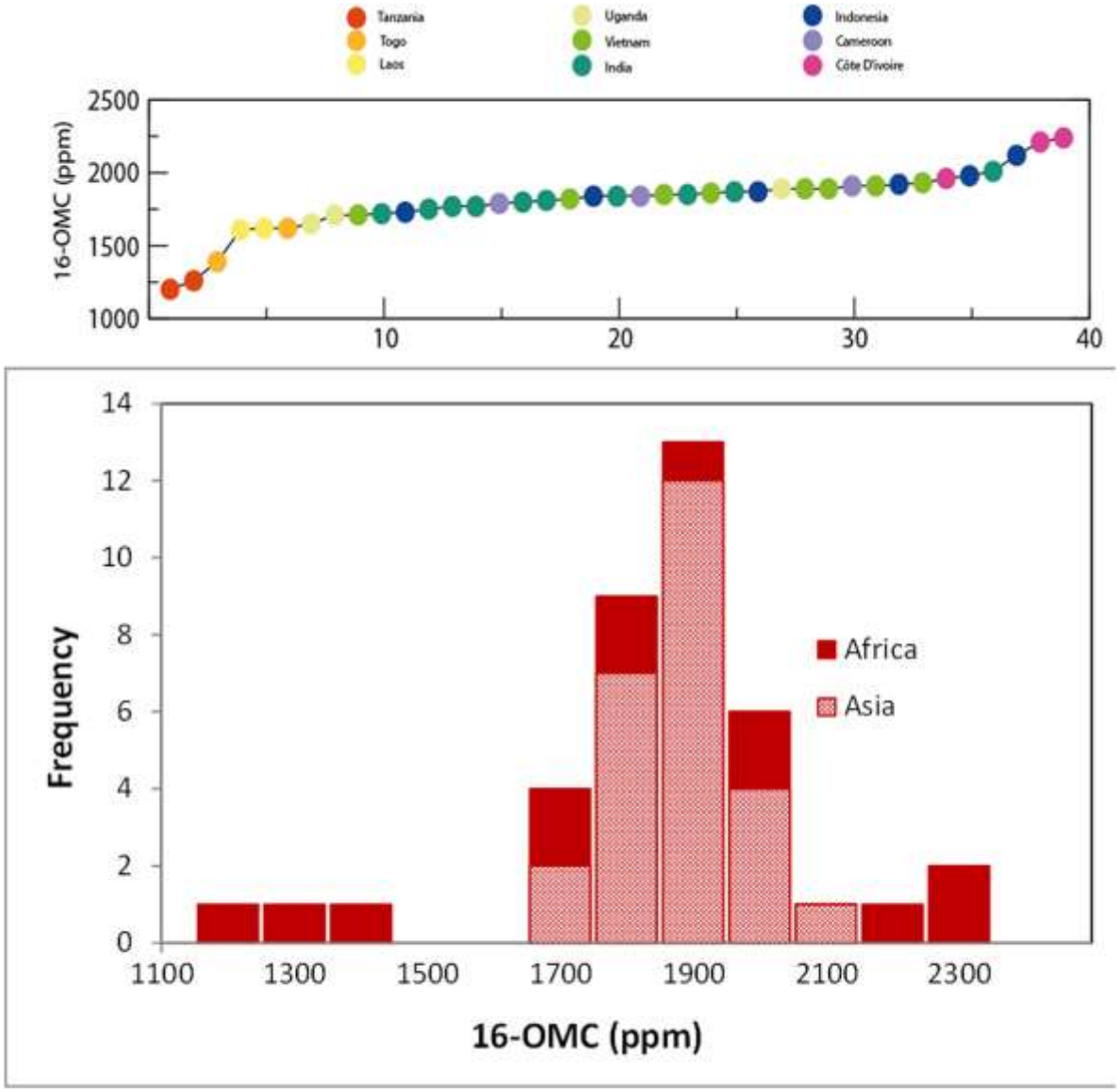


Figure 3 (top) Content of 16-OMC esters (express as 16-OMC free form) as determined by NMR for the 39 samples of roasted Robusta coffee. The geographical origin is indicated. (bottom) Histogram showing the frequency distribution of 16-OMC esters content for the Africa Robusta coffee samples (red) and for the subset originating from Asia (light red).

The values of 16-OMC esters contents are displayed in Fig. 3 in increasing order: the variation among the 39 samples considered in this study is much smaller than that reported in the literature (800-2500 mg/kg) (Kölling-Speer, Kurzrock, & Speer, 2006). From the histograms of Fig. 3, it is clear that most of the observed variability, in the presently investigated commercial lots, derives

from Robusta coffee samples originating from Africa while much smaller variations in the content of 16-OMC esters are found within the Asian samples. (average content = 1837 ± 113 mg/kg, standard deviation of 6%, min = 1606 mg/kg, max = 2123 mg/kg).

Considering the LoQ of 20 mg/kg of the NMR method (Schievano et al., 2014) and the range of values of 16-OMC esters found, adulterations of Arabica blends of 0.9 to 1.7% can be quantified. These values are slightly smaller than the 2% declared for the DIN method. In addition, it is possible to detect adulteration down to 0.2% considering that the LoD is 5 mg/kg**.**

Quantification of 16-OMC esters was also carried out on green coffee extracts for the sake of comparison. Different from the case of roasted coffee extracts, degradation products were always observed, irrespective of extraction timing and sample treatment (see Fig. 1S-c). We hypothesize that degradation of 16-OMC esters in the case of green coffee occurs as a consequence of cell disruption in the grinding step, possibly due to the high humidity content of the green beans and to the acidic conditions present in coffee (Defernez et al., 2017). Based on NMR spectra and $ESI^{+/-}$-MS measurements (Fig. 2S) of 16-OMC and of the same compound left in chloroform for a week, a molecular structure of the main degradation product is proposed (Table 1S). To take into account the effect of degradation, we quantified 16-OMC esters by integrating the singlet at 3.17 ppm, together with degradation signals (see Fig. 1S-c), as previously reported by (Monakhova et al., 2015).

The content of 16-OMC measured in the 39 samples, green and roasted, are reported in Fig. 4. The average value of 16-OMC esters over the 39 samples of green coffee was 1641 ± 16 mg/kg.

Considering the loss of water, of volatiles, the carbohydrates degradation, and a relative increase in lipid concentration resulting from the roasting process, the average values of 16-OMC esters in green and roasted coffee are consistent, confirming the substantial thermal stability of this compound (Dias, de Faria-Machado, Mercadante, Bragagnolo, & Benassi, 2014).

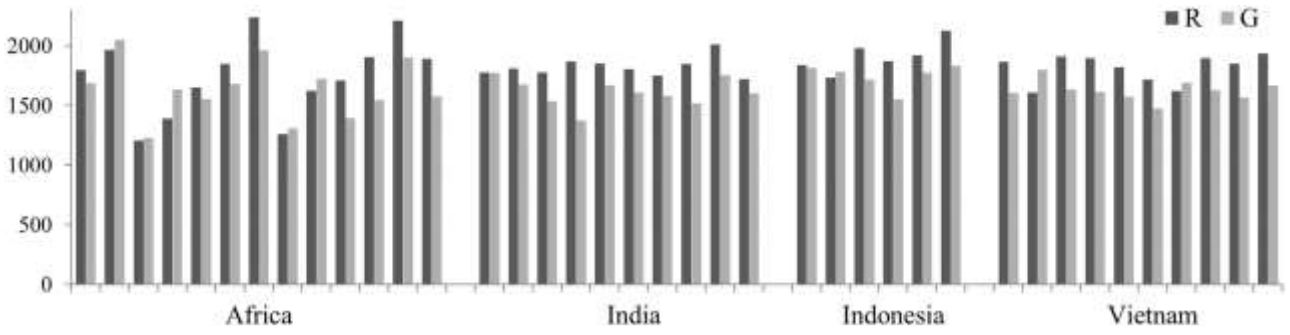


Figure 4 Content of 16-OMC (mg/kg) in roasted (R) and green (G) coffee samples.

Cafestol content was also measured in roasted coffee samples (Table 1); the values range between 1297 (± 10) and 2816 (± 48) mg/kg, again indicating a much smaller variation than that previously reported by Kölling-Speer et al. (2006).

Kahweol detection (S/N > 3), was possible in the majority of roasted coffee samples albeit by increasing the number of acquired transients to 256. Fig. 5 shows the kahweol signals in a Robusta and in an Arabica extract. However, kahweol quantification was possible only in 11 cases for which the estimated values were higher than the LoQ (S/N ≥ 10). The maximum estimated content of kahweol was about 200 mg/kg (Table 1). These data further confirm that kahweol cannot be considered a marker for Arabica given that some Arabica samples may contain as little as 0.1% (D'Amelio, De Angelis, Navarini, Schievano, & Mammi, 2013).

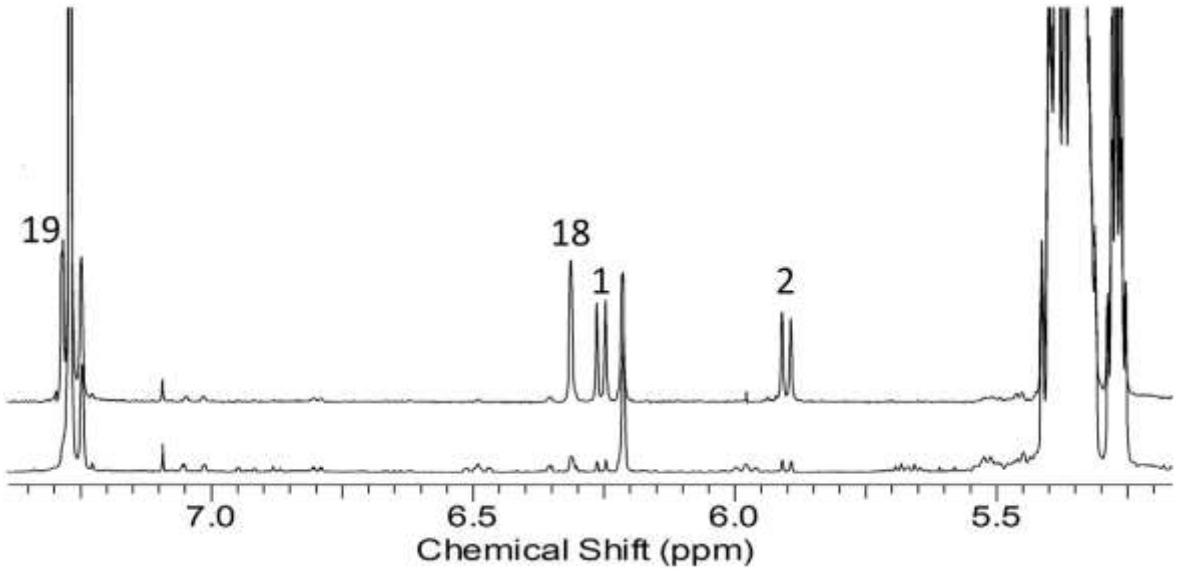


Figure 5 Expanded region (5.1 - 7.3 ppm) of the spectra of a Robusta (bottom) and an Arabica (top) extract in $CDCl_3$. The positions of the kahweol signals are marked with the corresponding number.

4. Conclusions

An analysis of 78 samples of Robusta coffee from different geographical origins was carried out to quantitatively determine the variability in the content of 16-OMC esters.

The observed range of 16-OMC esters in roasted beans was between 2236 ± 16 mg/kg and 1204 ± 6 mg/kg for the presently investigated commercial lots, with an average content of 1806 ± 208 mg/kg. This large natural variability precludes an accurate quantification of Robusta in unknown roasted coffee blends. Although much smaller variations in the content of 16-OMC esters were found within the lots of Asian samples (average content = 1837 ± 113 mg/kg) this finding cannot be of practical help to determine the composition of Arabica/Robusta roasted coffee blends because the geographical origin of the ingredients is generally unknown. However, the high amounts of such esters found in all samples confirm that 16-OMC is an excellent marker to detect fraudulent Robusta adulterations of pure Arabica blends, down to 0.2% using the NMR method.

Finally, our data confirm that kahweol cannot be considered a specific marker of *C. arabica* as it was detected in almost all Robusta samples in this study.

Figure captions

Fig.. 1 Left: Cafestol and right: 16-OMC esterified with palmitate.

Fig. 2. ^{1}H NMR spectra of cafestol palmitate (a), 16-OMC palmitate (b), 16-OMC (c) and of the chloroform extract of a roasted Robusta coffee (d). The spectral region between 2.9 ppm and 4.6 ppm is enlarged in the insert.

Fig. 3. (top) Content of 16-OMC esters (express as 16-OMC free form) as determined by NMR for the 39 samples of roasted Robusta coffee. The geographical origin is indicated. (bottom) Histogram showing the frequency distribution of 16-OMC esters content for the Africa Robusta coffee samples (red) and for the subset originating from Asia (light red).

Fig. 4. Content of 16-OMC (mg/kg) in roasted (R) and green (G) coffee samples.

Fig. 5. Expanded region (5.1 - 7.3 ppm) of the spectra of a Robusta (bottom) and an Arabica (top) extract in $CDCl_3$. The positions of the kahweol signals are marked with the corresponding number.

References

Alves, R. C., Casal, S., Alves, M. R., & Oliveira, M. B. (2009). Discrimination between arabica and robusta coffee species on the basis of their tocopherol profiles. *Food Chemistry*, *114*(1), 295–299.

Berti, F., Felluga, F., Forzato, C., Furlan, G., Nitti, P., Pitacco, G., Valentin, E. (2006). Chemoenzymatic synthesis of diastereomeric ethyl γ-benzyl paraconates and determination of the absolute configurations of their acids. *Tetrahedron: Asymmetry*, *17*, 2344-2353.

Campanha, F. G., Dias, R. C. E., & Benassi, M. de T. (2010). Discrimination of coffee species using kahweol and cafestol: effects of roasting and of defects. *Coffee Science*. Retrieved from http://www.coffeescience.ufla.br/index.php/Coffeescience/article/view/265

Casal, S., Alves, M. R., Mendes, E., Oliveira, M. B. P. P., & Ferreira, M. A. (2003). Discrimination between arabica and robusta coffee species on the basis of their amino acid enantiomers. *Journal of Agricultural and Food Chemistry*, *51*(22), 6495–501.

Casal, S., Oliveira, M. B. P. P., Alves, M. R., & Ferreira, M. A. (2000). Discriminate Analysis of Roasted Coffee Varieties for Trigonelline, Nicotinic Acid, and Caffeine Content. *Journal of Agricultural and Food Chemistry*, *48*(8), 3420–3424.

Cizkova, H., Soukupova, V., Voldrich, M., & Sevcik, R. (2007). Differentiation of coffee varieties according to their sterolic profile. *Journal of Food and Nutrition Research*, *46*, 28–34.

D'Amelio, N., De Angelis, E., Navarini, L., Schievano, E., & Mammi, S. (2013). Green coffee oil analysis by high-resolution nuclear magnetic resonance spectroscopy. *Talanta*, *110*, 118–127.

Davis, A. P., Govaerts, R., Bridson, D. M., & Stoffelen, P. (2006). An annotated taxonomic conspectus of the genus Coffea (Rubiaceae). *Botanical Journal of the Linnean Society*, *152*(4), 465–512.

De Angelis, E., Colomban, S., Lonzarich, V., Chetta, S., Navarini, L., Finotello, C., … Mammi, S. (2014). Coffee diterpenes: from green beans to espresso coffee. In *Proceedings of 25th International Scientific Colloquium on Coffee, ASIC* (pp. 42–48). Armenia, Colombia.

Defernez, M., Wren, E., Watson, A. D., Gunning, Y., Colquhoun, I. J., Le Gall, G., … Kemsley, E. K. (2017). Low-field 1H NMR spectroscopy for distinguishing between arabica and robusta ground roast coffees. *Food Chemistry*, *216*, 106–113.

Deutsches Institut für Normung e.V. (DIN). Standard DIN 10779, Analysis of Coffee and Coffee Products, Determination of 16-OMethyl Cafestol Content of Roasted Coffee HPLC Method; DIN: Berlin, Germany, 2011. (n.d.).

Dias, R. C. E., de Faria-Machado, A. F., Mercadante, A. Z., Bragagnolo, N., & Benassi, M. de T. (2014). Roasting process affects the profile of diterpenes in coffee. *European Food Research and Technology*, *239*(6), 961–970.

González, A. G., Pablos, F., Martín, M. J., León-Camacho, M., & Valdenebro, M. S. (2001). HPLC analysis of tocopherols and triglycerides in coffee and their use as authentication parameters. *Food Chemistry*, *73*(1), 93–101.

Grembecka, M., Malinowska, E., & Szefer, P. (2007). Differentiation of market coffee and its infusions in view of their mineral composition. *The Science of the Total Environment*, *383*(1-3), 59–69.

Guercia, E., Berti, F., Navarini, L., Demitri, N., & Forzato, C. (2016). Isolation and characterization of major diterpenes from C. canephora roasted coffee oil. *Tetrahedron: Asymmetry*, *27*(14-15), 649-656.

Hovell, A. M. C., Pereira, E. J., Arruda, N. P., & Rezende, C. M. (2010). Evaluation of alignment methods and data pretreatments on the determination of the most important peaks for the discrimination of coffee varieties Arabica and Robusta using gas chromatography-mass

spectroscopy. *Analytica Chimica Acta*, *678*(2), 160–8.
Illy, A., & Viani, R. (2005). *Espresso Coffee: The Science of Quality, 2nd edition*. London: Elsevier Academic Press.
Keidel, A., von Stetten, D., Rodrigues, C., Máguas, C., & Hildebrandt, P. (2010). Discrimination of green arabica and Robusta coffee beans by Raman spectroscopy. *Journal of Agricultural and Food Chemistry*, *58*(21), 11187–92.
Kölling-Speer, I., Kurzrock, T., & Speer, K. (2006). Diterpenes in green coffees harvested in different years. In *Proceedings of the 21st ASIC Colloquium on Coffee* (pp. 197–199). Montpellier, France.
Ky, C.-L., Louarn, J., Dussert, S., Guyot, B., Hamon, S., & Noirot, M. (2001). Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild Coffea arabica L. and C. canephora P. accessions. *Food Chemistry*, *75*(2), 223–230.
Mariani, C., & Fedeli, E. (1991). Sterols of coffee grain of Arabica and Robusta species. *La Rivista Delle Sostanze Grasse*, *68*, 111–115.
Martín, M. ., Pablos, F., & González, A. . (1999). Characterization of arabica and robusta roasted coffee varieties and mixture resolution according to their metal content. *Food Chemistry*, *66*(3), 365–370.
Martín, M. J., Pablos, F., González, A. G., Valdenebro, M. S., & León-Camacho, M. (2001). Fatty acid profiles as discriminant parameters for coffee varieties differentiation. *Talanta*, *54*(2), 291–297.
Monakhova, Y. B., Ruge, W., Kuballa, T., Ilse, M., Winkelmann, O., Diehl, B., … Lachenmeier, D. W. (2015). Rapid approach to identify the presence of Arabica and Robusta species in coffee using 1H NMR spectroscopy. *Food Chemistry*, *182*, 178–84.
Neumann, M. R. (n.d.). Global Coffee Forum, 30 September - 1 October 2015, Milano, Italy. Retrieved March 9, 2016, from http://www.globalcoffeeforum.com/presentation.php
Pablos, F., González, A. G., Martín, M. J., Valdenebro, M. S., & León-Camacho, M. (1999). Determination of the arabica/robusta composition of roasted coffee according to their sterolic content. *The Analyst*, *124*(7), 999–1002.
Pacetti, D., Boselli, E., Balzano, M., & Frega, N. G. (2012). Authentication of Italian Espresso coffee blends through the GC peak ratio between kahweol and 16-O-methylcafestol. *Food Chemistry*, *135*(3), 1569–74.
Pizarro, C., Esteban-Díez, I., & González-Sáiz, J. M. (2007). Mixture resolution according to the percentage of robusta variety in order to detect adulteration in roasted coffee by near infrared spectroscopy. *Analytica Chimica Acta*, *585*(2), 266–76.

Romano, R., Santini, A., Le Grottaglie, L., Manzo, N., Visconti, A., & Ritieni, A. (2014). Identification markers based on fatty acid composition to differentiate between roasted Arabica and Canephora (Robusta) coffee varieties in mixtures. *Journal of Food Composition and Analysis*, *35*(1), 1–9.

Rubayiza, A. B., & Meurens, M. (2005). Chemical discrimination of arabica and robusta coffees by Fourier transform Raman spectroscopy. *Journal of Agricultural and Food Chemistry*, *53*(12), 4654–9.

Rui Alves, M., Casal, S., Oliveira, M. B. P. P., & Ferreira, M. A. (2003). Contribution of FA profile obtained by high-resolution GC/chemometric techniques to the authenticity of green and roasted coffee varieties. *Journal of the American Oil Chemists' Society*, *80*(6), 511–517.

Schievano, E., Finotello, C., De Angelis, E., Mammi, S., & Navarini, L. (2014). Rapid authentication of coffee blends and quantification of 16-O-methylcafestol in roasted coffee beans by nuclear magnetic resonance. *Journal of Agricultural and Food Chemistry*, *62*(51), 12309–14.

Servillo, L., Giovane, A., Casale, R., Cautela, D., D'Onofrio, N., Balestrieri, M. L., & Castaldo, D. (2016). Homostachydrine (pipecolic acid betaine) as authentication marker of roasted blends of Coffea arabica and Coffea canephora (Robusta) beans. *Food Chemistry*, *205*, 52–57.

Souza, R. M. N. de, & Benassi, M. T. (2012). Discrimination of commercial roasted and ground coffees according to chemical composition. *Journal of the Brazilian Chemical Society*, *23*(7), 1347–1354.

Spaniolas, S., Tsachaki, M., Bennett, M. J., & Tucker, G. A. (2008). Evaluation of DNA extraction methods from green and roasted coffee beans. *Food Control*, *19*(3), 257–262.

Speer, K., & Kolling-Speer, I. (2001). Lipids. In R. J. Clarke & O. G. Vitzthum (Eds.), *Coffee: Recent Developments* (pp. 33–49). Oxford, UK: Blackwell Science.

Trantakis, I. A., Spaniolas, S., Kalaitzis, P., Ioannou, P. C., Tucker, G. A., & Christopoulos, T. K. (2012). Dipstick test for DNA-based food authentication. Application to coffee authenticity assessment. *Journal of Agricultural and Food Chemistry*, *60*(3), 713–7.

Waller, J.M. , Bigger, M., Hillocks, R. J. (Ed.). (2007). *Coffee Pests, Diseases and their Management*. Egham, Surrey, UK: CAB International.

Wermelinger, T., D'Ambrosio, L., Klopprogge, B., & Yeretzian, C. (2011). Quantification of the Robusta fraction in a coffee blend via Raman spectroscopy: proof of principle. *Journal of Agricultural and Food Chemistry*, *59*(17), 9074–9.

Table 1. Contents of 16-O-methylcafestol, Cafestol and Kahweol in *Coffea canephora* from different geographical origins

	Roasted					
Origin	16-OMC	SD	Cafestol	SD	Kahweol	SD
	mg/kg					
AFRICA	1793	20	1879	36		
	1964	12	2256	16		
	1204	6	1876	4	197	4
	1390	7	2145	52		
	1649	13	1875	19		
	1844	20	1817	51		
	1257	4	1791	23		
	1622	5	2058	33		
	2236	16	2816	48		
	1706	14	2445	44	34.6	0.5
	1906	4	1698	27		
	2208	4	2725	35		
	1889	13	2607	29		
INDIA	1804	2	1805	18	76	1
	1749	5	1494	6	143.4	1.6
	1774	5	1710	50		
	1807	7	1666	23		
	1849	13	2077	31		
	1774	9	1781	34		
	1867	20	1577	39		
	1844	4	2116	9	62.6	0.8
	2007	11	2557	13		
	1719	5	1702	19		
INDONESIA	2123	15	2053	16	99.2	0.8
	1731	9	1562	14	65.9	0.9
	1837	20	1297	10		
	1979	4	1939	21		
	1873	11	1616	44	25.4	0.4
	1921	8	1811	24	62.9	0.8
[illegible]	1892	8	1683	44		
	1848	7	1830	11		
	1933	12	2233	11		
	1863	4	1623	44		
	1907	23	2510	53	200.4	1.4
	1895	15	1484	31		