Food Chemistry 329 (2020) 127129

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: wwwelseviercom/locate/foodchem

Limited genotypic and geographic variability of 16-O-methylated diterpene content in *Coffea arabica* green beans

Vincent Portaluri[a], Freddy Thomas[a,*], Sophie Guyader[a], Eric Jamin[a], Benoit Bertrand[b,c], Gérald S. Remaud[d], Elisabetta Schievano[e], Stefano Mammi[e], Elena Guercia[f], Luciano Navarini[f]

[a] *Eurofins Analytics France, 9 rue Pierre Adolphe Bobierre, B.P. 42301, F-44323 NANTES Cedex 3, France*
[b] *CIRAD, UMR IPME, F-34398 Montpellier, France*
[c] *UMR IPME, Univ Montpellier, CIRAD, IRD, F-34398 Montpellier, France*
[d] *EBSI Team, Interdisciplinary Chemistry: Synthesis, Analysis, Modelling (CEISAM), University of Nantes-CNRS UMR 6230, 2 rue de la Houssinière, BP 92208, F-44322 Nantes cedex 3, France*
[e] *Department of Chemical Sciences, University of Padova, Via Marzolo 1, 35131 Padova, Italy*
[f] *illycaffè S.p.A, Via Flavia 110, 34147 Trieste, Italy*

ARTICLE INFO

Keywords:
16-O-methylcafestol
16-O-methylkahweol
Coffea arabica
Green coffee
NMR

ABSTRACT

The acknowledged marker of Robusta coffee, 16-O-methylcafestol (16-OMC), can be quantified by NMR as a mixture with 16-O-methylkahweol (16-OMK), which accounts for approximately 10% of the mixture. In the present study, we detected and quantified 16-O-methylated diterpenes (16-OMD) in 248 samples of green *Coffea arabica* beans by NMR. We did not observe any differences between genotypes introgressed by chromosomal fragments of Robusta and non-introgressed genotypes. Environmental effects suggesting a possible protective role of 16-OMD for adaptation, as well as genotypic effects that support a high heritability of this trait were observed. Altogether, our data confirmed the presence of 16-OMD in green Arabica at a level approximately 1.5% that of a typical Robusta, endorsing the validity of 16-OMD as a marker for the presence of Robusta.

1. Introduction

Coffee is the most widely consumed beverage in the world and one of the most commercialized food products. In spite of more than 120 known botanical species (Davis, Govaerts, Bridson, & Stoffelen, 2006; Davis, Tosh, Ruch, & Fay, 2011), only two are commercially important: *Coffea arabica* L. (Arabica coffee) and *Coffea canephora* Pierre ex A. Froehner (Robusta coffee). Arabica is the most valuable coffee in the trade because it produces a very fragrant, sweet, smooth, and slightly acidic beverage with a very rich and complex aroma. Robusta coffee has rougher aromatic notes, is more bitter, and more astringent. The lower price of Robusta relative to Arabica prompted the search for new rapid methodologies to distinguish them to prevent economically motivated adulterations (Everstine, Spink, & Kennedy, 2013).

Several approaches have been suggested to distinguish these coffee species, using many different compounds as reliable discriminants (Finotello et al., 2017). In recent years, the coffee lipid fraction has attracted interest because its components can be successfully used to discern the two coffee species.

Coffee diterpenes (Speer & Kolling-Speer, 2001) are the main constituents of the unsaponifiable coffee oil fraction. They are mainly esterified with various fatty acids, and the free form is present only in small amounts (De Angelis et al., 2014). The three most important diterpenes are cafestol, kahweol and 16-O-methylcafestol (16-OMC). They are produced only by plants of the *Coffea* genus. Cafestol is found in both Arabica and Robusta coffee while kahweol is present in larger amounts in Arabica (0.1–1.0%) than in Robusta (up to 200 mg/kg) (Finotello et al., 2017). For years, 16-OMC was considered to be present exclusively in Robusta (Bonnlander, Wünnecke, & Winterhalter, 2007; Speer & Kolling-Speer, 2001; Kamm et al., 2002; Kurzrock & Speer, 2001; Oellig & Radovanovic, 2017; Pacetti, Boselli, Balzano, & Frega, 2012; Speer & Kölling-Speer, 2006). This fact made 16-OMC an excellent authenticity marker for the presence of Robusta in coffee products, considering also its thermal stability. To the best of our knowledge, only two early studies reported the presence of 16-OMC in roasted Arabica. The oldest one detected trace amounts (10 mg/kg) of 16-OMC in one Arabica sample from Guatemala (Speer & Montag, 1989), whereas the other one reported a higher content but remarkably lower than that found in Robusta (Lercker, Frega, Bocci, & Rodriguez-Estrada, 1995).

* Corresponding author.
E-mail address: FreddyThomas@eurofins.com (F. Thomas).

hiips://doi.org/10.1016/j.foodchem.2020.127129
Received 12 December 2019; Received in revised form 20 May 2020; Accepted 22 May 2020
Available online 27 May 2020

Table 1
Genetic and geographical origin of the *C. arabica* samples used in the present investigation.

	Country	Number of samples	Introgressed (samples with 16-OMD > 50 ppm)	Non introgressed (samples with 16-OMD > 50 ppm)	Introgression not determined (samples with 16-OMD > 50 ppm)
CIRAD collection	Nicaragua	96	47 (2)	49 (2)	
Commercial	South & Central America	12			12 (0)
	Africa	12			12 (0)
	Asia	5			5 (1)
	Oceania	1			1 (0)
	unknown	6			6 (0)
Breedcafs project	Nicaragua	80	60 (13)	20 (0)	
	Central America	24	13 (2)	11 (0)	
Additional	Central America	7		4 (2)	3 (0)
	South America	1			1 (0)
	Africa	3			3 (0)
	India	1			1 (0)
Total		**248**			

In addition to these three major diterpenes, 16-O-methylkahweol (16-OMK) deserves mention. First identified and quantified in Robusta coffee by Kolling-Speer, Kurzrock, and Speer (2001) and Kolling-Speer and Speer (2001), its amounts are about 10 times lower than those of kahweol and even less in roasted Arabica (Pacetti et al., 2012). The interest in 16-OMK stems from the use of the diagnostic NMR peak at 3.16–3.18 ppm to quantify coffee diterpenes (Finotello et al., 2017; Schievano, Finotello, De Angelis, Mammi, & Navarini, 2014). This peak derives from the H21 methyl groups of both esterified 16-OMC and 16-OMK (cumulatively named hereafter 16-O-methylated diterpenes, 16-OMD) (D'Amelio, De Angelis, Navarini, Schievano, & Mammi, 2013; Gunning et al., 2018; Scharnhop & Winterhalter, 2009). Gunning et al. (2018) modified the sample preparation procedure used up to that point to quantify 16-OMC by NMR (Defernez et al., 2017; Monakhova et al., 2015; Schievano et al., 2014) and improved the limit of detection. In contrast to previous NMR studies, they observed 16-OMC and 16-OMK in Arabica coffees, although at very low levels. The content of 16-OMD in roasted Arabica is approximately 1.5% that of a typical Robusta (Finotello et al., 2017; Kurzrock & Speer, 2001). This amount may represent only a problem in revealing unintentional Robusta contamination in a 100% Arabica roasted coffee blend because undeclared intentional additions of Robusta lower than 2% would hardly be economically advantageous.

The presence of trace amounts of 16-OMD in Arabica roasted beans has more relevant biochemical and genetic implications. 16-OMC has been found in other tissues of the Arabica coffee plant (Speer & Kölling-Speer, 2006), and 16-OMK has been found in leaves (Speer & Kolling-Speer, 2001), suggesting that the biochemical machinery to synthesize these diterpenes is present in the Arabica genome. A legitimate doubt remains as to whether this capability has been transferred to Arabica plants through the breeding programs implemented to confer resistance to main diseases by introgression of appropriate *C. canephora* chromosomal fragments. Today, Arabica cultivars derived from *C. canephora* via the interspecific Timor Hybrid (a spontaneous cross between *C. canephora* × *C. arabica*) represent more than 40% of the Arabica trees cultivated around the world. Unfortunately, introgression via the Timor Hybrid is sometimes accompanied by a substantial drop in cup quality (Bertrand, Guyot, Anthony, & Lasherme, 2003). Consequently, coffee buyers or roasters may wish to assess whether the coffee they are purchasing comes from introgressed Arabica varieties. The current study is meant to shed light on the possible relationship between introgression of *C. canephora* traits and the presence of 16-OMD in *C. arabica*.

Despite their genetic background, the only two backcrossed samples (one 'arabica × robusta' var. and one introgressed arabica cv.) previously examined (Gunning et al., 2018), did not show a significantly more intense 16-OMD signal. This finding stimulated the current investigation on a larger number of coffee samples in the attempt to verify the presence and discern the origin of 16-OMC and 16-OMK in Arabica coffee. To avoid any possible influence of the roasting process, screening was carried out on ground green coffee. Authentic Arabica green coffee samples were characterized by 1D ^{1}H NMR spectroscopy according to the method proposed by Schievano et al. (2014) and Monakhova et al. (2015) to quantify 16-OMD. The diagnostic NMR peak was detected and quantified in many of the investigated green Arabica beans. Screening was extended to a wide range of samples, including several clones of F1 hybrids obtained by crossing wild Ethiopians accessions × introgressed cultivars of different geographical origin or different trade quality, in the case of commercial lots. NMR experiments were also carried out on green coffee oil extracted from selected samples, and a preliminary UPLC-MS/MS method was developed to validate the quantitative NMR data and to determine the amounts of 16-OMC and 16-OMK separately.

2. Materials and methods

2.1. Coffee samples

All 248 green *C. arabica* samples used in the present investigation are reported in Table 1 and described in detail in Table 1S.

To study the possible effect of introgression on the content of 16-OMD, 96 samples were analyzed from the CIRAD collection (Nicaragua) (see Table 1S, Source: CIRAD collection). We analyzed 36 additional samples from 15 different countries of Latin America (12), Africa (12), Asia (5) and Oceania (1) and 6 of unknown origin (see Table 1S, Source: commercial) for a wide geographical representation. Harvest periods of the samples were spread over several years between 2007 and 2015. The samples were stored in tubes (10–15 g) in hermetically sealed boxes with silica gel, kept in the dark. All of the Ethiopian coffee samples were from wild accessions. The samples from the CIRAD collection represent 47 introgressed accessions and 49 non-introgressed accessions. The introgression was verified with simple sequence repeat (SSR) markers at the CIRAD institute. All information about species, varieties and other parameters are registered in the CIRAD coffee database.

Additionally, 104 green coffee beans samples were harvested in 2018 in the framework of the Breedcafs project (http://www.breedcafs.eu/) and immediately analyzed: 80 were from Nicaragua and 24 from Costa Rica (see Table 1S, Source: Breedcafs). This pool of samples includes 31 non-introgressed cultivars (Caturra, Catuai, Pakal, H3) and introgressed varieties (Starmaya Centroamerica, EvaLuna, MundoMex, Mundo Maya, Totonaca and Marsellesa). Marsellesa, Caturra and Catuai are reproduced by seeds. Starmaya is a first F1 hybrid reproduced by seeds (Georget et al., 2019). Pakal, H3, MundoMex, MundoMaya,

EvaLuna and Centroamerica are clones of F1 hybrids (Bertrand, Etienne, Cilas, Charrier, & Baradat, 2005; Bertrand et al., 2006; Bertrand et al., 2011) reproduced by somatic embryogenesis (Etienne et al., 2018).

Twelve samples including wet processed Arabica commercial lots of different geographical origin (Costa Rica, Ethiopia, Honduras, India, Brazil, Mexico and Rwanda) and Arabica samples from CATIE germplasm collection (Costa Rica) were also used (see Table 1S Source: Additional).

2.2. Sample preparation

2.2.1. Sample preparation for NMR analysis of lipophilic extracts

The green coffee beans were ground to a particle size of about 1 mm using a *Retsch MM400* ball mill (grinding jar: stainless steel, 50 mL, screw top design; grinding ball: Ø 20 mm, stainless steel; frequency 20.0 Hz; time 60 s). To 200 (± 5) mg of ground coffee, 1.5 mL of deuterated chloroform ($CDCl_3$, 99.8%-d) containing 0.1% tetramethylsilane (Sigma Aldrich, Steinheim, Germany) were added. The mixture was shaken using a BioShaker (Bruker BioSpin, Rheinstetten, Germany) for 5 min at 650 rpm and then filtered through a hydrophilic membrane filter (pore size: 0.22 μm, Sartorius; Goettingen, Germany). The extract (700 μL) was transferred directly into a 7'' NMR tube (507-PP-7 Wilmad tube of Class A, Vineland, New Jersey, USA).

2.2.2. Sample preparation for NMR analysis of coffee oil

The green coffee beans were ground after contact with liquid N_2 using a batch mill (IkaM20) to pass a 630 μm sieve (Giuliani). Coffee oil was isolated by means of Soxhlet extractions. In detail, 10 g of ground coffee samples were extracted with 130 mL of *n*-pentane (Sigma Aldrich) for 8 h at the solvent boiling point, siphoning five times per hour. The extract was paper filtered and dried over anhydrous sodium sulfate. The solvent was evaporated under reduced pressure (Buchi, Rotavapor R114), and the residue was dried to constant weight to obtain the green coffee oil.

Samples were prepared dissolving 100 L of green coffee oil in 700 L of $CDCl_3$.

2.2.3. Sample preparation for UPLC – MS/MS spectrometry analysis

The green coffee beans were ground using a *Retsch MM400* ball mill (grinding jar: stainless steel, 50 mL, screw top design; grinding ball: Ø 25 mm, stainless steel; frequency 30.0 Hz; time 20 s), which was carefully cleaned after grinding each sample.

The procedure followed the indications of De Souza and Benassi (2012) with some modifications. Ground powders (8 g) were saponified with 100 mL of ethanolic potassium hydroxide under reflux at a water bath temperature of 90 °C for 6 hs. After adding 80 mL of mQ water, the unsaponifiable fraction was extracted with 100 mL of diethyl ether three times, and the organic phase was washed twice with 100 mL of mQ water. The organic solvent was removed using a rotary evaporator. The dried organic phase was resuspended in 2 mL of mobile phase (H_2O/ACN 40/60 w/w).

2.3. Spectral acquisition

2.3.1. ^1H NMR spectroscopy

The ^1H NMR spectra of coffee oil were acquired with a Bruker DMX600 spectrometer equipped with a 5 mm TXI triple gradient probe (Bruker BioSpin, Rheinstetten; Germany). The ^1H NMR spectra of the lipophilic extracts were collected using two Bruker Avance 400 MHz spectrometers running TopSpin 3.0 and 3.2 software, both equipped with a 5 mm BBI probe with Z-gradient coils and a SampleXPress autosampler (Bruker BioSpin, Rheinstetten; Germany). The spectra were acquired at 301.8 K using a zg pulse sequence. For each spectrum, 32 scans (512 for coffee oil) of 64 k points with 4 prior dummy scans were collected using a spectral width of 20.5 ppm, acquisition time of 3.98 s, recycling delay of 8 s(2 s for coffee oil). Total acquisition time for lipophilic extracts was less than 20 min per sample (including 5 min for temperature equilibration inside the magnet). Before each acquisition, tuning, matching (atma) and shimming (topshim) were performed automatically on each tube.

2.3.2. UPLC-MS/Ms

UPLC-MS/MS analyses were performed with an Agilent 1290 UPLC interfaced to a SCIEX Triple Quad™ 4500. The chromatography column was a Waters ACQUITY HSS T3 C18 (2.1 × 100 mm, particle size 1.8 μm) maintained at 30 °C with a flow rate of 500 μL/min. A two-solvent system was used: solvent A, 0.1% formic acid in water; solvent B, acetonitrile. The linear gradient was as follows: 0 min 40% A; 6.5 min 40% A; 9.5 min 20% A; 10 min 40% A. The injection volume was 10 μL.

The mass spectrometer was run in positive ion electrospray mode, acquiring in Multiple Reaction Monitoring (MRM), which guarantees the specificity of the method. The relative standard deviation of the MRM1/MRM2 signal ($RSD_{MRM1/MRM2}$) (De sà et al., 2014) and the operating conditions were optimized using 16-OMC standard solutions. The source temperature was 350 °C. Specific compound transitions (331 → 299 *m/z*; 331 → 281 *m/z*) were confirmed based on literature data and comparison with reference solutions. The operating conditions were optimized using a 16-OMC and a kahweol standard solution, respectively, for the 16-OMC and the 16-OMK.

2.4. Data analysis

2.4.1. 16-OMD quantification by NMR

Spectra were automatically transformed and phased. After baseline correction, the spectra were exported from TopSpin as ASCII files and elaborated with Matlab version R2009a (The MathWorks, Natick, Massachusetts, USA) with Statistical Toolbox and in-house routines. The spectral shifts were adjusted using the icoshift warping method (Savorani, Tomasi, & Engelsen, 2010). Non-informative and solvent zones (< 0.22 ppm, > 10.52 ppm and chloroform singlet) were removed from the spectra. Each spectrum was normalized based on the sum of all spectral intensities to eliminate spectrometer-related gain effects. The spectra were divided using a bucket width of 0.001 ppm. Cafestol, kahweol and 16-OMD were quantified by integration of corresponding peaks, respectively, at 5.82 ppm, 6.14 ppm, and 3.10 ppm, using cafestol acetate as reference compound for response factor determination. Because of the quantitative properties of NMR in the experimental conditions used, this response factor allowed us to quantify cafestol, kahweol and 16-OMD from their peak areas. The limit of quantification of 1 g/kg for kahweol and cafestol and 0.01 g/kg for 16-OMD (uncertainty was evaluated at 15%) were determined respectively.

2.4.2. 16-OMD quantification by UPLC-MS/MS

In all samples, 16-OMC was quantified using a 16-OMC standard (Phyproof Reference Substances, PhytoLab, Germany) solution calibration curve (concentration range of 2–20 μg/mL; regression coefficient > 0.99). A standard for 16-OMK is not commercially available; therefore, 16-OMK was quantified using a kahweol (Phyproof Reference Substances, PhytoLab, Germany) standard solution calibration curve (concentration range of 1 – 10 μg/mL; regression coefficient > 0.995). Data were treated with MultiQuant software. Uncertainty was evaluated at 20%.

3. Results and discussion

3.1. 16-OMD identification and quantification by UPLC-MS/MS

Unlike HPLC-UV and GC, LC-MS/MS has not been widely used to analyze coffee 16-OMD (de Carvalho Martins et al., 2018). As far as we

know, the first HPLC-MS/MS method developed to detect and quantify 16-OMC in roasted coffee via multiple reaction monitoring (MRM) against an external standard dates back to 2007 (Bonnlander, Wünnecke, & Winterhalter). According to the authors, the method was developed to prevent overestimation observed via classical HPLC-UV analysis of traces of 16-OMC in coffee blends containing low amounts of Robusta. At that time, however, the presence of 16-OMD in Arabica had been overlooked. More recently, Gunning et al. (2018) reported a UPLC-MS method to detect 16-OMC and 16-OMK in Arabica roasted coffee. Because 16-OMD in green Arabica have not been quantified via UPLC-MS/MS so far, for the first time, a proper method was preliminarily developed and validated for this purpose.

Specific compound transitions (331 → 299 *m/z* and 331 → 281 *m/z*) and (329 → 297 *m/z* and 329 → 279 *m/z*) indicate the presence of 16-OMC and 16-OMK, respectively.

Fig. 1S_A shows chromatograms of 16-OMK and 16-OMC obtained from a typical Arabica sample (retention time of 3.26 min for 16-OMK and 3.54 min for 16-OMC). Fig. 1S_B reports the UPLC-MS/MS chromatograms of both 16-OMC standard and sample together with the corresponding calibration curve.

The amounts of 16-OMC found in selected Arabica green beans are reported in Table 2. The quantification of 16-OMK was performed on three different samples (Ethiopia, T.16713 and T.16713 batch 2 of Table 2), carefully selected because they cover a wide range of 16-OMC + 16-OMK content (from 11 mg/kg to 250 mg/kg). The amount of 16-OMK found in Ethiopia coffee sample was between 1 and 2 mg/kg whereas 24 and 30 mg/kg of 16-OMK were determined in the case of the other two samples examined, respectively. Based on these findings, it may be suggested that 16-OMC accounts for approximately 90% of the total 16-OMD present in green Arabica coffee.

3.2. ^{1}H NMR spectra: focus on 16-OMD

Fig. 1A shows two expanded regions of a typical NMR spectrum obtained from Arabica green coffee oils. When present in coffee, 16-OMD (Fig. 1B) exhibit a singlet in the ^{1}H NMR spectrum at 3.16 – 3.18 ppm (protons 21), which is well-resolved and isolated from other signals even at 60 MHz (Defernez et al., 2017; Gunning et al., 2018; Schievano et al., 2014) and for this reason, it is used as a diagnostic peak that can be easily integrated. Unfortunately, the discrimination between 16-OMC and 16-OMK cannot be achieved using NMR methods. Based on UPLC-MS/MS findings, we conclude that the ^{1}H NMR singlet at 3.17 ppm originates from protons 21 of both 16-OMC and 16-OMK. Protons 17 of 16-OMD form a second-order system at 3.78 ppm in the free form while, in the fatty acid ester, two doublets are obtained at 4.28 and 4.45 ppm, respectively, with a scalar coupling constant ^{2}J = 12.8 Hz (Finotello et al., 2017). Only the doublet at 4.45 ppm is well-resolved in the spectrum, while the other doublet partially overlaps the signals of the glycerol moiety of triglycerides. In the Arabica coffee extracts, 16-OMD are present prevalently in the esterified form, as highlighted in Fig. 1A.

Table 2
Quantification of 16-OMD (NMR) and 16-OMC (UPLC-MS/MS) in green Arabica coffee oil samples. See Table 1S, Info Source: additional and *Breedcafs.

Sample (Geographical Origin)	NMR (mg/kg)	UPLC – MS/MS (mg/kg)
Coffee0137 (Nicaragua)	< 5	< 0.6
190 (Costa Rica)	< 10	2 ± 0.4
Sample 4 (Costa Rica)*	< 10	2 ± 0.4
312 (Honduras)	< 10	3 ± 0.6
218 (Rwanda)	< 10	4 ± 0.8
120 (India)	< 10	8 ± 1.6
935 (Ethiopia)	10 ± 1.5	11 ± 2.2
Sample 10 (Costa Rica)*	70 ± 10	50 ± 10
T.16713 (Costa Rica)	200 ± 30	230 ± 46
T.16713 batch2 (Costa Rica)	260 ± 39	250 ± 50

As shown in Table 2, a satisfactory match between the quantitative data obtained by UPLC-MS/MS and those by NMR on green coffee oils was found.

Isolation of coffee oil is time and solvent consuming, and this process makes both UPLC-MS/MS and NMR analyses of coffee oil not conceivable for extensive screening purposes. In this work, diterpenes were quantified in Arabica samples using NMR spectra of $CDCl_3$ extracts. The method does not require any sample manipulation; it is fast and specific and is therefore the most advantageous approach when the number of samples to be analyzed is very large. The much lower content of 16-OMD in Arabica beans than in Robusta beans may be analytically challenging when trace amounts have to be determined quantitatively. On the other hand, the diagnostic methyl protons singlet is more intense in the NMR spectra of isolated green coffee oil than on $CDCl_3$ lipophilic extracts. The accuracy of the NMR method using $CDCl_3$ extracts has been verified by comparing the 16-OMD content with that determined in coffee oil (selected samples: 131, 132, 241, 242, 77, 73, 18, see Table 1S, Supp. Info Source: Breedcafs, T.16713, T.16713 batch 2. Source: additional). As shown in Fig. 1C, a satisfactory match was found.

3.3. Impact of introgression on 16-OMD

The contents of the major diterpenes, cafestol and kahweol, are in full agreement with previously reported data on green *C. arabica* (Kitzberger et al., 2013; Speer & Kölling-Speer, 2006), as shown in Table 3. In most samples, 16-OMD was detected under the experimental limit of quantification (Table 1S).

The content of the diterpenes does not discriminate between introgressed and wild Arabica accessions (Fig. 2). Thus, introgression does not affect the variability of these compounds within the Arabica species. The introgressed varieties are created adding *C. canephora* chromosomal fragments to Arabica varieties followed by several selections to suppress as many of the *C. canephora* genes as possible for optimum cup quality. Evidently, these selections have retained the Arabica features as far as the content of diterpenes.

The average content of 16-OMD was 23 mg/kg for introgressed samples and 13 mg/kg for non-introgressed ones as outlined in Table 3. The 16-OMD content of 22 samples out of 248 (8.9%) was higher than 50 mg/kg, independent of introgression (Table 1 and Table 1S-sheet 2).

3.4. Exploring possible sources of variability in the 16-OMD content

The effect of the environment on variability in the 16-OMD content was studied by cultivating a wide range of genotypes, including a wide range of *cultivars* and F1 hybrids, in different geographical areas of Central America in the framework of the Breedcafs H2020 EU project (see Table 1S). Growers in Latin America currently have the choice between two types of homozygous lines propagated by seed: American traditional *cultivars* (Bourbon, Typica, Caturra, Catuai) and Catimor/Sarchimor *cultivars*. These recent latter *cultivars* are derived from cv. Timor Hybrid, which is a natural cross between *C. arabica* and *C. canephora* (Bertrand et al., 2011). Since 1997, coffee growers have had access to new F1 hybrids, derived from crosses between wild Sudan-Ethiopian and American traditional *cultivars* or Catimor *cultivars,* considerably expanding the narrow genetic base of coffee trees cultivated in Latin America (Bertrand et al., 2011). The wide range of genotypes of the present investigation is representative of American traditional *cultivars* and new F1 hybrids.

The content of 16-OMD was < 10 mg/kg in all Caturra and Marsellesa *cultivars* as well as Starmaya, Centroamerica (H1), Totonaca (H14), Mundo Maya (H16) and Pakal (H17) F1 hybrids cultivated in different geographical areas (different farms at different altitudes) within Nicaragua. On the other hand, the variability within the F1 hybrids H3, Mundo Mex (H15), Evaluna (H18) and Nayarita (H19) was high, sometime remarkably high, as shown in Table 4. This variability

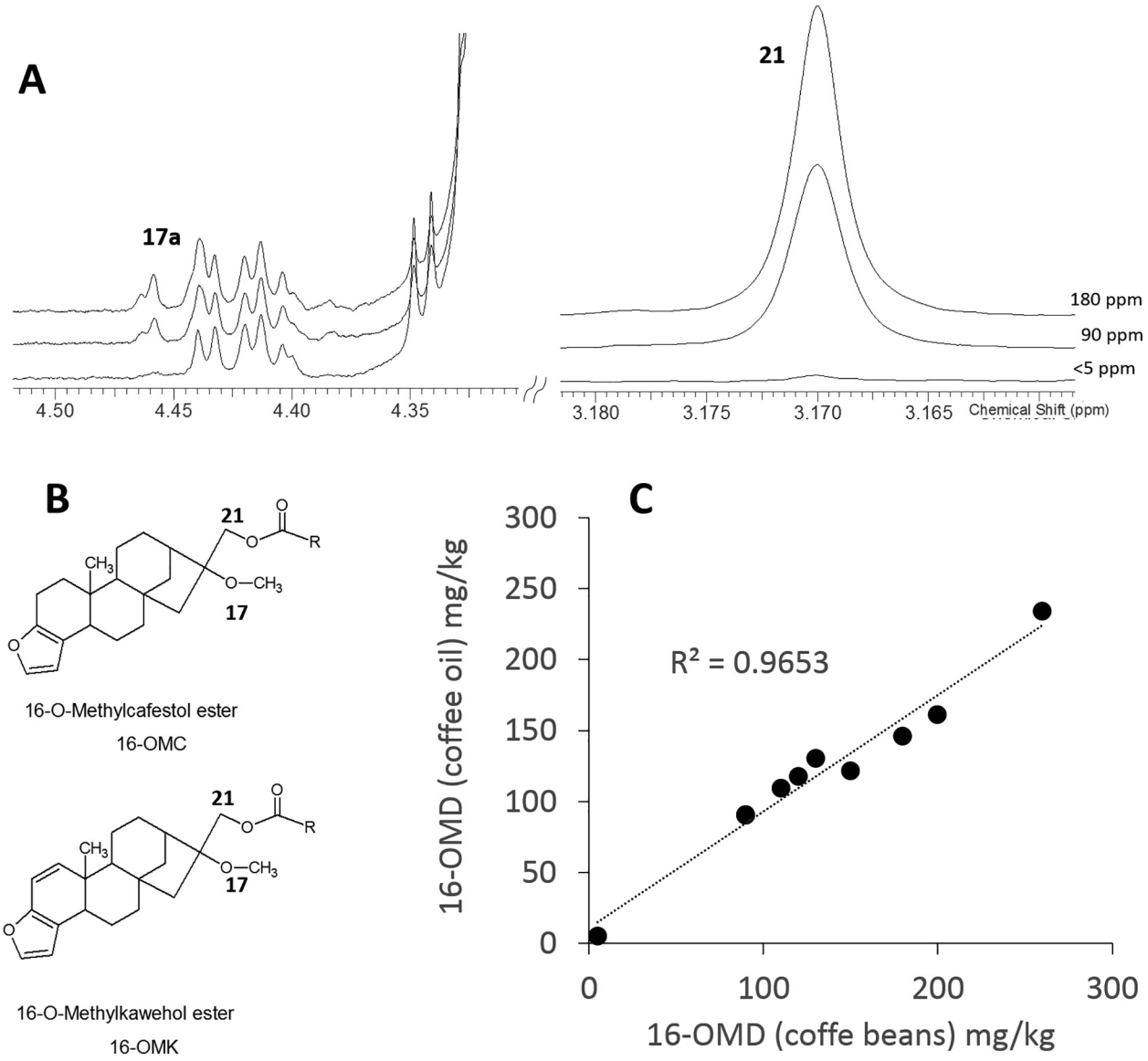

Fig. 1. A: selected regions of the 600 MHz spectra of three coffee oil samples (Breedcafs samples code 242, 18, 245, ordered from high to low 16-OMD content) showing the diagnostic peaks for 16-OMD esters. **B**: chemical structure of 16-OMC and 16-OMK. **C**: Correlation between the 16-O-methylated diterpenes content measured with a direct extraction on the coffee beans (horizontal scale) or from the coffee oil (vertical scale).

highlights a notable effect of the environment, particularly for H3 and Nayarita (H19) F1 hybrids. Two different samples of the F1 hybrid Mundo Mex (H15) from Costa Rica contained 70 and 20 mg/kg of 16-OMD, respectively, suggesting also in this case a non-negligible effect of the environment. Similarly, three different samples of the F1 hybrid H3 from Costa Rica contained < 10 mg/kg of 16-OMD, remarkably lower than the content of the same hybrid cultivated in Nicaragua. The F1 hybrid H1 contained less than 10 mg/kg when produced in Nicaragua (4 different farms at different altitudes) and a higher amount when produced in Costa Rica (up to 60 mg/kg) (Table 1S).

The entire data set shows that the great majority of green Arabica coffee samples contained 16-OMD < 10 mg/kg with the systematic exception of F1 hybrids Mundo Mex (H15), Evaluna (H18) and Nayarita (H19) and sporadic exceptions of F1 hybrids Centroamerica (H1) and H3 depending on where they were grown. It is noteworthy that these F1 hybrids are propagated in clonal form. Therefore, the effects observed are clearly due to the environment. It may be hypothesized that high levels of diterpenes could play a protection role linked with adaptation to abiotic stresses.

The experimental data on F1 hybrids Mundo Mex (H15), Evaluna (H18) and Nayarita (H19) seem to suggest a possible genetic effect on 16-OMD variability. These F1 hybrids share the same mother tree

Table 3
Concentration of 16-OMD, cafestol and kahweol for non-introgressed and introgressed Arabica samples.

	Number of samples	16-OMD (mg/kg)			kahweol (g/kg)			cafestol (g/kg)		
		min	max	Average	min	max	Average	min	max	Average
Non-introgressed	84	< 10	200	13	1.1	9.5	4.5	1.3	11.1	5.2
Introgressed	120	< 10	180	23	0.6	9.7	5.3	1.0	10.0	5.2

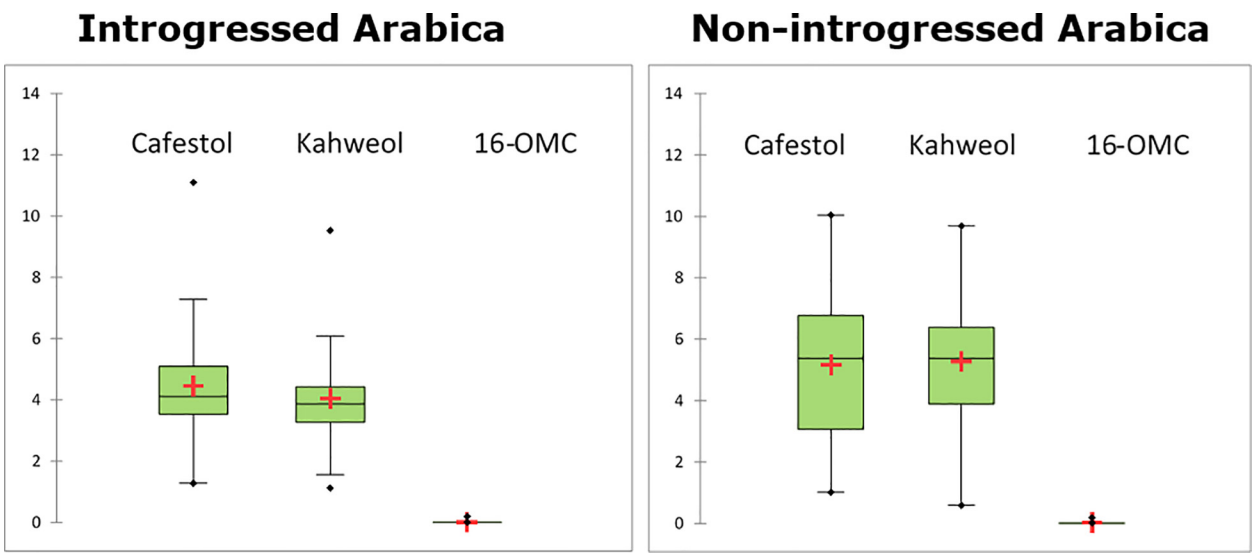

Fig. 2. Boxplot of the diterpenes content (g/kg) for introgressed (left) and non-introgressed (right) Arabica.

Table 4
16-OMD content of F1 hybrids H3, Mundo Mex (H15), Evaluna (H18) and Nayarita (H19) from different environment (crop 2018 where not indicated).

Country	F1 hybrid	Farm - Location	16-OMD (mg/kg)
Nicaragua	H3	Boaco 710 m	30
Nicaragua	H3	Las Marias 1190 m	50
Nicaragua	H3	La Aurora 1240 m	30
Nicaragua	H3	Albania 1250 m	60
Costa Rica	H3	Bella Triunfo 1400 m	< 10
Costa Rica	H3	San Ignacio 1500 m	< 10
Costa Rica	H3	Palmichal 1450 m	< 10
Costa Rica	Mundo Mex (H15)	Limonal	70
Nicaragua	Mundo Mex (H15)	Boaco 710 m	90
Nicaragua	Mundo Mex (H15)	Las Marias 1190 m	80
Nicaragua	Mundo Mex (H15)	La Aurora 1240 m	60
Nicaragua	Mundo Mex (H15)	Albania 1250 m	90
Costa Rica	Mundo Mex (H15)	Rio Grande 1200 m	20
Nicaragua	Evaluna (H18)	Boaco 710 m	90
Nicaragua	Evaluna (H18)	Las Marias 1190 m	130
Nicaragua	Evaluna (H18)	La Aurora 1240 m	110
Nicaragua	Evaluna (H18)	Albania 1250 m	90
Nicaragua	Nayarita (H19)	Boaco 10 m	120
Nicaragua	Nayarita (H19)	Las Marias 1190 m	180
Nicaragua	Nayarita (H19)	La Aurora 1240 m	150
Nicaragua	Nayarita (H19)	Albania 1250 m	100

(T17931 Catimor) and F1 hybrids Mundo Mex (H15) and Nayarita (H19) are full-sibs (same female T17931 and same male ET26). The high content in these two hybrids comes more from the Ethiopian male parent (ET26) than from T17931, which has a content < 10 mg/kg.

It has been speculated that the diversity for diterpene content both at the interspecific and intraspecific levels in *Coffea* sp. suggests the existence of genetic polymorphism of the enzyme controlling the cafestol/kahweol biosynthesis pathways (Sridevi, Giridhar, & Ravishankar, 2010). Recently, within *C. arabica* sp., a genome-wide association study identified several SNPs associated with cafestol and kahweol content and cafestol/kahweol ratio (Sant'Ana et al., 2018). In the case of 16-OMD, however, no genetic studies have been undertaken yet.

Among the wild accessions, samples of the non-introgressed variety ET26 ORSTOM (CIRAD Collection) were the only ones available from two different geographical environments: Nicaragua germplasm (coffee0154 and coffee0178) and CATIE germplasm collection in Costa Rica (T.16713, T.16713 batch 2). The latter showed a remarkably higher content of 16-OMD than those from Nicaragua. These data seem to further suggest a relevant effect of the environment as a source of variability in the content of 16-OMD. In the case of *C. canephora*, significant differences between genotypes and growing sites for the content of 16-OMD have been observed (Mori et al., 2016).

To complete the screening, we extended the study to commercial lots representative of different relevant geographical origins. Without exceptions, 16-OMD contents < 20 mg/kg were determined in full agreement with data reported by Gunning et al. (2018) on roasted Arabica cultivars.

4. Conclusions

A screening of 248 green *C. arabica* samples was conducted to ascertain the presence of 16-OMD and to quantify them by NMR spectroscopy, taking advantage of the speed of analysis and the simplicity of preparation of lipophilic extracts. A preliminary UPLC-MS/MS method was also developed "*ad hoc*", which confirmed the co-presence of 16-OMC and 16-OMK in the unsaponifiable fraction of green Arabica coffee lipids and allowed the quantification of their relative amount. Approximately 90% w/w of the 16-OMD is represented by 16-O-methylcafestol, and 16-O-methylkahweol represent the remaining 10% w/w of the mixture. The accuracy of the NMR method was checked on a representative sample pool by performing a parallel quantification of 16-OMD also on isolated coffee oil.

184 green Arabica samples out of 248 (74.2% of the analyzed samples) contained less than 10 mg/kg of 16-OMD while 42 samples (16.9%) were characterized by a content higher than 20 mg/kg. The 16-OMD content of 8 samples (3.2%) was over 100 mg/kg and below 260 mg/kg. Maximum amounts equal to 200 mg/kg and 260 mg/kg were determined on two non-commercially available samples from CATIE germplasm collection (Costa Rica), whereas amounts in the range 100 – 180 mg/kg were determined on certain F1 hybrids under study from an agronomical performance point of view. The presence of 16-OMC and 16-OMK in green *C. arabica* is fully confirmed in agreement with recently reported data on roasted Arabica (Gunning et al., 2018). The samples collected in Ethiopia at Jimma (i.e. in the center of the natural origin of the species) showed 16-OMD values ranging from 0 to 50 mg/kg.

324 samples have been investigated to verify the effect of introgression of *C. canephora* chromosomal fragments carrying desirable genes, a successfully exploited practice to induce disease resistance. The maximum content of 16-OMD was 260 mg/kg for non-introgressed accessions and 180 mg/kg for introgressed accessions. This study

therefore shows no effect of introgression.

An environmental effect on the observed 16-OMD variability was detected. Some varieties studied in five environments (such as H3) showed variations from < 10 to 70 mg/kg suggesting the capacity of some cultivars to adapt the level of 16-OMD in response to the environment. On the other hand, Mundo Mex (H15) and Nayarita (H19), two full-sibs from the same cross (T17931 × ET26), showed high and relatively constant levels of 16-OMD in any environment (20–90 mg/kg and 100–180 mg/kg, respectively) suggesting a high heritability of this trait and therefore a genetic effect. Further dedicated studies to unravel the role played by the environment and the genotypes are required.

The present investigation confirms that 16-OMC cannot be considered a diterpene exclusively present in Robusta; however, its content can pose only limited concern in coffee authenticity assessment. The presence of 16-OMC in pure Arabica at a level commensurate with 1.5% Robusta addition can be considered an unintentional contamination rather than an intentional adulteration in view of the extremely poor economic benefit ensured by this misbehavior. Finally, the quantification of other coffee terpenes should be considered for an optimal determination of Arabica and Robusta blends.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Fundacion NicaFrance for its technical support in the filed in Nicaragua. This research was partially funded by the BREEDCAFS Project (http://www.breedcafs.eu), which was supported by the European Commission under the Horizon 2020 – Research an Innovation Programme H2020-SFS-2016-2, grant agreement number: 727934.

William Solano Sánchez from Agriculture, Livestock and Agroforestry Program, CATIE, Turrialba, Costa Rica is acknowledged for kind delivery of some selected samples to Illycaffè spa.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2020.127129.

References

Bertrand, B., Guyot, B., Anthony, F., & Lasherme, P. (2003). Impact of the Coffea canephora gene introgression on beverage quality of C. arabica. *Theoretical and Applied Genetics, 107*(3), 387–394.

Bertrand, B., Etienne, H., Cilas, C., Charrier, A., & Baradat, P. (2005). Coffea arabica hybrid performance for yield, fertility and bean weight. *Euphytica, 141*, 255–262.

Bertrand, B., Vaast, P., Alpizar, E., Etienne, H., Davrieux, F., & Charmetant, P. (2006). Comparison of bean biochemical composition and beverage quality of Arabica hybrids involving Sudanese-Ethiopian origins with traditional varieties at various elevations in Central America. *Tree Physiology, 26*, 1239–1248.

Bertrand, B., Alpizar, E., Lara, L., SantaCreo, R., Hidalgo, M., Quijano, J. M., et al. (2011). Performance of *Coffea arabica* F1 hybrids in agroforestry and full-sun cropping systems in comparison with American pure line cultivars. *Euphytica, 181*, 147–158.

Bonnlander, B., Wünnecke, H., & Winterhalter, P. (2007). Rapid and sensitive detection of Robusta traces in blends of raw and roasted coffee. *Lebensmittelchemie, 61*, 73–104.

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 genus Coffea (Rubiaceae). *Botanical Journal of the Linnean Society, 152*, 465–512.

Davis, A. P., Tosh, J., Ruch, N., & Fay, M. F. (2011). Growing coffee: Psilanthus (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of Coffea. *Botanical Journal of the Linnean Society, 167*, 357–377.

De Angelis, E., Colomban, S., Lonzarich, V., Chetta, S., Navarini, L., Finotello, C., Mammi, S., & Schievano, E. (2014). Coffee Diterpenes: from Green Beans to Espresso Coffee in Proceeding of 25th International Scientific Colloquium on Coffee (42-48), Armenia, Colombia ASIC.

de Carvalho Martins, V., de Oliveira Godoy, R.L., Senna Gouvea, A.C.M., Pessanha de Araujo Santiago, M.C., Galhardo Borguini, R., de Oliveira Braga, E.C., et al. (2018). Fraud investigation in commercial coffee by chromatography. Food Quality and safety, 3, 121-133.

De sà, M., Ferreira, J. P., Queiroz, V. T., Vilas-Boas, L., Silva, M. C., Almeida, M. H., Guerra-Guimaraes L., & Bronze, M. R. (2014). A liquid chromatography/electrospray ionisation tandem mass spectrometry method for the simultaneous quantification of salicylic , jasmonic and abscisic acids in Coffea arabica leaves. Journal of the Science of Food and Agriculture, 94, 529–536.

De Souza, R. M. N., & 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.

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

Etienne, H., Breton, D., Breitler, J.-C., Bertrand, B., Déchamp, E., Awada, R., & Ducos, J.-P. (2018). Coffee somatic embryogenesis: How did research, experience gained and innovations promote the commercial propagation of elite clones from the two cultivated species? *Frontiers in Plant Science, 9*, 1630–1651.

Everstine, K., Spink, J., & Kennedy, S. (2013). Economically motivated adulteration (EMA) of food: Common characteristics of EMA incidents. *Journal of Food Protection, 76*(4), 723–735.

Finotello, C., Forzato, C., Gasparini, A., Mammi, S., Navarini, L., & Schievano, E. (2017). NMR quantification of 16-O-methylcafestol and kahweol in Coffea canephora var. robusta beans from different geographical origins. *Food Control, 75*, 62–69.

Georget, F., Marie, L., Alpizar, E., Courtel, P., Bordeaux, M., Hidalgo, J.M., Marraccini, P., Breitler, J.-C., Déchamp, E., poncon, C., Etienne, H., & Bertrand, B. (2019). Starmaya: The first Arabica F1 coffee hybrid produced using genetic male sterility. Frontiers in Plant Science, 10, 1344-1357.

Gunning, Y., Defernez, M., Watson, A. D., Beadman, N., Colquhoun, I. J., Le Gall, G., & Kemsley, E. K. (2018). 16-O-methylcafestol is present in ground roast Arabica coffees: Implications for authenticity testing. *Food Chemistry, 248*, 52–60.

Kamm, W., Dionisi, F., Fay, L., Hischenhuber, C., Schmarr, H., & Engel, K. (2002). Rapid and simultaneous analysis of 16-O-methylcafestol and sterols as markers for assessment of green coffee bean authenticity by on-line LC – GC. *Journal of the American Oil Chemists' Society, 79*(11), 1109–1113.

Kitzberger, C. S. G., Scholz, M. B. D., Pereira, L. F. P., Viera, L. G. E., Sera, T., Silva, J. B. G. D., & Benassi, M. D. T. (2013). Diterpenes in green and roasted coffee of Coffea arabica cultivars growing in the same edapho-climatic conditions. *Journal of Food Composition and Analysis, 30*, 52–57.

Kolling-Speer, I., Kurzrock, T., & Speer, K. (2001). Contents of diterpenes in green coffees assigned for the European market in Proceeding of 19th International Scientific Colloquium on Coffee. Trieste, Italy ASIC.

Kolling-Speer, I.; Speer, K. (2001) 16-O-methylcafestol in robusta coffees in Proceeding of 19th International Scientific Colloquium on Coffee. Trieste, Italy ASIC.

Kurzrock, T., & Speer, K. (2001). Diterpenes and diterpenes esters in coffee. *Food Reviews International, 17*(4), 433–450.

Lercker, G., Frega, N., Bocci, F., & Rodriguez-Estrada, M. T. (1995). High resolution gas chromatographic determination of diterpenic alcohols and sterols in coffee lipids. *Chromatographia, 41*, 29–33.

Monakhova, Y. B., Ruge, W., Kuballa, T., Ilse, M., Winkelmann, O., Diehl, B., et al. (2015). Rapid approach to identify the presence of Arabica and Robusta species in coffee using ^{1}H NMR spectroscopy. *Food Chemistry, 182*, 178–184.

Mori, A. L. B., Kalschnea, D. L., Gava Ferrão, M. A., da Fonsecab, A. F. A., Gava Ferrão, R., & de Toledo Benassi, M. (2016). Diterpenes in *Coffea canephora. Journal of Food Composition and Analysis, 52*, 52–57.

Oellig, C., & Radovanovic, J. (2017). Screening for 16-O-methylcafestol in roasted coffee by high-performance thin-layer chromatography – fluorescence detection – Determination of Coffea canephora admixtures to Coffea arabica. *Journal of Chromatography A, 1525*, 173–180.

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*, 1569–1574.

Sant'Ana, G.C., Pereira, L.F.P., Pot D., Ivamoto, S.T., Domingues, D.S., Ferreira, R.V., et al. (2018). Genome-wide association study reveals candidate genes influencing lipids and diterpenes contents in Coffea arabica L. Scientific Reports 8, 465-477.

Savorani, F., Tomasi, G., & Engelsen, S. B. (2010). icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *Journal of Magnetic Resonance, 202*(2), 190–202.

Scharnhop, H., & Winterhalter, P. (2009). Isolation of coffee diterpenes by means of high-speed countercurrent chromatography. *Journal of Food Composition and Analysis, 22*, 233–237.

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), 12309–12314.

Speer, K., & Kolling-Speer, I. (2001) Lipids. In COFFEE Recent Developments, 1st Ed.; Clarke, R. J., Vitzhum, O.G., Eds.; Blackwell Science Ltd, London (UK), 7, 33-49.

Speer, K., & Kölling-Speer, I. (2006). The lipid fraction of the coffee bean. *Brazilian Journal of Plant Physiology, 18*(1), 201–216.

Speer, K., & Montag, A. (1989). 16-O-Methylcafestol – ein neues diterpen im kaffee erste ergebnisse: Gehalte in roh- und röstkaffees. *Deutsche Lebensmittel-Rundschau, 12*, 381–384.

Sridevi, V., Giridhar, P., & Ravishankar, G. A. (2010). Free diterpenes cafestol and kahweol in beans and *in vitro* cultures of *Coffea* species. *Current Science, 99*(8), 1101–1104.