LWT - Food Science and Technology 103 (2019) 205–211

Contents lists available at ScienceDirect

LWT - Food Science and Technology

journal homepage: www.elsevier.com/locate/lwt

Direct-infusion electrospray ionization-mass spectrometry analysis reveals atractyligenin derivatives as potential markers for green coffee postharvest discrimination

Anna Tsukui[a], Pedro H. Vendramini[b], Rafael Garrett[a], Maria Brígida S. Scholz[c], Marcos N. Eberlin[b], Humberto R. Bizzo[d], Claudia M. Rezende[a,*]

[a] Universidade Federal Do Rio de Janeiro, Instituto de Química, CEP 21941-909, Rio de Janeiro, Brazil
[b] Universidade Estadual de Campinas, Instituto de Química, Laboratório ThoMSon de Espectrometria de Massas, CEP 13083-970, Campinas, São Paulo, Brazil
[c] Instituto Agronômico Do Paraná, CEP 86001-970, Londrina, Paraná, Brazil
[d] Embrapa Agroindústria de Alimentos, CEP 23020-470, Rio de Janeiro, Brazil

ARTICLE INFO

Keywords:
ESI (−) FT-ICR MS
Data analysis
Postharvest
Atractyligenin derivatives
Coffee

ABSTRACT

Coffee quality is strongly related to postharvest processes. Dry and wet are the most traditional and employed processes. However, more recently, the semi-dry process has been introduced and applied for producing coffees of good quality in a more environmental friendly manner compared to the wet process. In this work, green Arabica coffee cherries from two consecutive crops were submitted to the dry and semi-dry postharvest processes and their methanol:water extracts were characterized by ESI(−) FT-ICR MS followed by chemometric data analysis. Thirty-three compounds were identified in the extracts by their high-resolution and accurate *m/z* values and MS/MS spectra being nine of them diterpenes related to atractyligenin derivatives. The dry and semi-dry processes led to distinct extracts in which compounds such as atractyloside II, carboxyatractyloside II and carboxyatractyloside III showed the greatest contributions ($p < 0.05$) based on *t*-Student's test and PLS-DA statistical analyses. These compounds were able to discriminate the postharvest processes for coffees from two consecutive crops. These results can help to better understand the effects of the postharvest processes in the final chemical composition and coffee quality.

1. Introduction

Postharvest treatment of freshly harvested coffee cherries into green coffee beans can be performed by the dry, wet or semi-dry processing methods (Selmar, Kleinwachter, & Bytof, 2014; Teixeira, Brando, Thomaziello, & Teixeira, 2005). The wet process, which generates the "washed coffee bean", generally leads to beverages with pleasant acidity and full aroma, whereas the dry process results in beverages with the so-called full body sensation (Borém, Isquierdo, & Taveira, 2014).

This sensorial variation is undoubtedly mostly related to changes in the chemical composition caused by the different postharvest processes. For example, in the dry process, which generates the natural coffee (NC), the fruits remain intact during direct sun or machine drying, keeping the hull, pulp and mucilage that still surround the bean. In the wet process, a physical action removes both the hull and pulp, whereas the mucilage is removed by a fermentative step (Teixeira et al., 2005; Toleto, Pezza, Pezza, & Toci, 2016). Germination of the coffee bean has already been observed after the fermentative process (Selmar, Knopp, & Breitenstein, 2006). In washed coffee beans, glucose and fructose content is drastically reduced, whereas no change is observed for caffeine, sucrose, and proteins contents. Variation in polysaccharides, such as galactomannans, cellulose and arabinogalactans have also been observed (Tarzia, Scholz, & Petkowicz, 2010).

Due to the mechanical pulping step, only fully ripe coffee cherries can be treated by the wet process, a limitation that clearly has a positive impact on the quality of the beverages. Simple drying and hulling of the cherries is however employed in the natural coffee and it is common to observe ripe, unripe and overripe fruits being processed together, which negatively impacts the quality of some dry coffee (Teixeira et al., 2005; Toleto et al., 2016).

* Corresponding author.
E-mail addresses: annatsukui.at@gmail.com (A. Tsukui), ph_vendramini@yahoo.com.br (P.H. Vendramini), rafael_garrett@iq.ufrj.br (R. Garrett), mbscholz@iapar.br (M.B.S. Scholz), mneberlin@gmail.com (M.N. Eberlin), humberto.bizzo@embrapa.br (H.R. Bizzo), claudia.rezendeufrj@gmail.com (C.M. Rezende).

https://doi.org/10.1016/j.lwt.2018.12.078
Received 15 April 2018; Received in revised form 27 December 2018; Accepted 30 December 2018
Available online 31 December 2018
0023-6438/

The semi-dry process generates the pulped natural coffee (PC) and was initially developed in Brazil but is also used in other South and Central America countries, as well as in Indonesia. As in the wet process, only ripe cherries can be treated, but the polysaccharide rich mucilage that covers the beans is not removed, adding body and sweet notes after roasting, producing high quality beverages (Teixeira et al., 2005; Toleto et al., 2016).

In pulped natural coffee, it has been observed a reduction in the chlorogenic acids, trigonelline, sucrose, glucose, and fructose contents, whereas caffeine remains unchanged (Duarte, Pereira, & Farah, 2010; Selmar, Bytof, & Knopp, 2008). However, more comprehensive studies on the chemical changes of this type of coffee are urgent.

One important aspect of the semi-dry process is the substantial reduction of water consumption, ca. 10 times less than the wet process, resulting also in a reduction of the waste from the fermentation process. The semi-dry process is therefore more environmental friendly and leads to high quality brews (Borém et al., 2014).

Electrospray ionization-mass spectrometry (ESI-MS) fingerprinting followed by chemometric data analysis is being successfully employed in coffee science for years. For example, it has allowed the discrimination of coffee samples based on their geographical origin (Amorim et al., 2009; Garrett, Rezende, & Ifa, 2013a), species (Arabica *x* Robusta) (Correia et al., 2016; Garrett, Vaz, Hovell, Eberlin, & Rezende, 2012), detection of adulteration on roasted coffee (Aquino et al., 2014) and discrimination on difference cultivars of Arabica coffee (Garrett et al., 2013b). In all precedent examples, direct analysis of the extracts without previous separation was performed.

Regarding differentiation among the postharvest processes, desorption electrospray ionization–mass spectrometry (DESI-MS) has been already used and applied directly to the surface of green coffee beans. Samples from wet, dry and semi-dry treatments were clearly differentiated by applying multivariate data analysis (Garrett et al., 2014).

The aim of this work was to investigate whether coffees from two different crops (2012–2013) processed by the dry and semi-dry treatments could be differentiated by their chemical profiles using direct-infusion electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS) as well as to unveil the chemical markers responsible for the postharvest processes differentiation.

2. Material and methods

2.1. Green coffee beans

Arabica coffee samples were obtained from Instituto Agronômico do Paraná (IAPAR), Brazil. Only the ripe fruits were used for this work. For the dry process (which originates the natural coffee, NC), whole fruits were dried under the sun on cement terrace to 12% moisture. For the semi-dry (which originates the pulped natural coffee, PC), the coffee cherries were washed mechanically, peeled and dried also under the sun. After drying, the coffee samples were pulped and kept under 11% moisture at controlled atmosphere. A total of 49 coffee samples were analyzed, 30 samples from the 2012 crop (11 PC and 19 NC) and 19 samples from the 2013 crop (11 PC and 8 NC).

2.2. Sample extraction

Coffee beans were frozen in liquid nitrogen, grounded in a Perten 3600 mill and standardized to sieve size of 0.5 mm. The extraction was carried out following Garrett et al. (2013b) with adaptations. Approximately, 1 g of ground coffee was extracted with 4 mL of methanol:water (4:1) in an ultrasonic bath (Cristófoli, Paraná, Brazil) for 20 min, filtered by a 0.2 μm PVDF membrane and diluted (1:10) with 0.1% ammonium hydroxide solution before injection into the mass spectrometer. All samples were extracted in triplicate.

2.3. Electrospray ionization fourier transform-ion cyclotron resonance mass spectrometry (ESI FT-ICR MS)

Samples extracts were analyzed using a 7.2T LTQ FT Ultra mass spectrometer (Thermo Scientific, Bremen, Germany) equipped with a chip-based direct-infusion nanoelectrospray ionization source (Advion BioSciences, Ithaca, NY, USA) operating in the negative-ion mode at the following conditions: capillary voltage 1.6 kV, tube lens 160 V, temperature 270 °C, and fragmentation energy 10–40eV. Data were acquired for 3 min in the *m/z* range of 100–1000 by the Xcalibur 2.0 software.

2.4. Data analysis

Values of relative abundances and *m/z* for characteristic ions from each sample analyzed by FT ICR MS were exported as.*txt* file from Xcalibur software and uploaded into the *Matlab* software (MathWork, United States) to process the data. Identification of chemical compounds was done using the high-resolution and accurate *m/z* values and tandem MS data obtained by collision-induced dissociation (CID) experiments. These data were compared to standards (caffeoyl and dicaffeoylquinic acids) and to a homemade library of coffee compounds based on the literature (Shu et al., 2014; Garrett et al., 2013b, 2013a and; Alonso-Salces, Serra, Reniero, & Heberger, 2009; Clifford, Knight, & Kuhnert, 2005; Clifford, Johnston, Knight, & Kuhnert, 2003).

After substances identification (see Table 1), a single table containing 49 rows (samples) and 33 columns (*m/z* value of identified compounds) was used for statistical data analysis by Statistica 7.0 (Statsoft, South America) and Unscrambler X v.10.3 (CAMO, Norway) software.

Significant difference ($p < 0.05$) among postharvest processing was applied for *t*-Student's test. Besides, partial least square discriminant analysis (PLS-DA) was applied for sample discrimination. Quality assessment of the classification model was checked by the coefficient of determination of the calibration model (R^2) and cross-validation (Q^2), root mean square error of calibration (RMSEC) and cross-validation (RMSECV), slope of calibration and cross-validation (Ferreira, 2015).

3. Results and discussion

3.1. Compound identification in coffee by ESI(−) FT-ICR MS

Fig. 1 shows the typical ESI(−) FT-ICR mass spectra for the methanol:water extract of green coffee Arabica beans and Table 1 lists the chemical constituents identified by their accurate *m/z* values and main MS/MS fragments in the negative ion mode. All identified compounds have already been described in literature for coffee samples.

Thirty-three ions could be associated to deprotonated molecules $[M-H]^-$, such as sugars, fatty acids, diterpenes, chlorogenic acid, among others (Garrett et al., 2014; Garrett et al., 2013b; Garrett et al., 2013a; Rodrigues & Bragagnolo, 2013; Amorim et al., 2009; Alonso-Salces et al., 2009; Clifford et al., 2005). The mass (*m/z*) error was less than 2.0 ppm for all ions, due to the high mass accuracy of this technique (Garrett et al., 2013b; Garrett et al., 2012).

Nearly 40% of the identified compounds are related to chlorogenic acids found in coffee, such as caffeoylquinic acid (*m/z* 353.0878, $[M-H]^-$), feruloylquinic acid (*m/z* 367.1034, $[M-H]^-$) and dicaffeoylquinic acid (*m/z* 515.1190, $[M-H]^-$). This class of compounds is formed by the esterification of *trans*-cinnamic acids, such as caffeic, ferulic and *p*-coumaric with quinic acid (Clifford et al., 2003). Among these, the most intense compound was caffeoylquinic acid (*m/z* 353.0878, $[M-H]^-$), followed by quinic acid (*m/z* 191.0561, $[M-H]^-$).

Palmitic (*m/z* 255.2329, $[M-H]^-$), linoleic (*m/z* 279.2329, $[M-H]^-$), oleic (*m/z* 281.2480, $[M-H]^-$), stearic (*m/z* 283.2642, $[M-H]^-$), arachidic (*m/z* 311.2955, $[M-H]^-$) and behenic (*m/z* 339.3268, $[M-H]^-$) acids account for 18% of the total compound identified. They are present in coffee beans as free or esterified acids such as triacylglicerols, diterpenic and sterol esters, phosphatides and serotonin amides (Speer & Kölling-Speer, 2006).

Table 1
Identification of chemical compounds in green Arabica coffee beans by ESI (−) FT-ICR MS.

N°	Compound	Formula		Theoretical (*m/z*)	Crop 2012		Crop 2013		ESI (−) MS/MS	Fragmentation energy (eV)
					Experimental (*m/z*)	Error (ppm)	Experimental (*m/z*)	Error (ppm)		
1	Caffeic acid	$C_9H_8O_4$	$[M-H]^-$	179.0349	179.0350	0.55	179.0348	0.55	135.0442	13
2	Hexose	$C_6H_{12}O_6$	$[M-H]^-$	179.0561	179.0560	0.34	179.0563	0.99	–	25
3	Ferulic acid	$C_{10}H_{10}O_4$	$[M-H]^-$	193.0506	193.0505	0.51	193.0507	0.51	178.0267 149.0600 134.0365	10
4	Quinic acid	$C_7H_{12}O_6$	$[M-H]^-$	191.0561	191.0560	0.33	191.0561	0.00	127.0392 93.0335 85.0283	25
5	Palmitic acid	$C_{16}H_{32}O_2$	$[M-H]^-$	255.2329	255.2328	0.50	255.2329	0.00	273.2225	37
6	Linoleic acid	$C_{18}H_{32}O_2$	$[M-H]^-$	279.2329	279.2328	0.28	279.2328	0.35	261.9409	37
7	Oleic acid	$C_{18}H_{34}O_2$	$[M-H]^-$	281.2480	281.2482	0.71	281.2485	1.77	59.0125	39
8	Stearic acid	$C_{18}H_{36}O_2$	$[M-H]^-$	283.2642	283.2641	0.28	283.2641	0.35	265.2540	40
9	Arachidic acid	$C_{20}H_{40}O_2$	$[M-H]^-$	311.2955	311.2956	0.32	311.2958	0.80	293.2855	38
10	Behenic acid	$C_{22}H_{44}O_2$	$[M-H]^-$	339.3268	339.3265	0.88	339.3269	0.29	321.3169	30
11	Caffeoylshikimic acid	$C_{16}H_{16}O_8$	$[M-H]^-$	335.0772	335.0773	0.29	335.0775	0.86	179.0345 161.0455 135.0443	25
12	Coumaroylquinic acid	$C_{16}H_{18}O_8$	$[M-H]^-$	337.0928	337.0925	0.48	337.0929	0.29	191.0557 163.0394	13
13	Sucrose	$C_{12}H_{22}O_{11}$	$[M-H]^-$	341.1089	341.1088	0.42	341.1090	0.29	179.0555 161.0448 89.0232 71.0126 59.0126	20
14	Caffeoylquinic acid	$C_{16}H_{18}O_9$	$[M-H]^-$	353.0878	353.0876	0.49	353.0870	1.41	191.0557 179.0345 173.0450	14
15	Ferulic acid hexoside	$C_{16}H_{20}O_9$	$[M-H]^-$	355.1034	355.1031	0.84	355.1033	0.06	193.0604 173.0450	15
16	Caffeoyltryptophane	$C_{20}H_{18}N_2O_5$	$[M-H]^-$	365.1143	365.1146	0.82	365.1142	0.27	229.0620 186.0556 135.0443	13
17	Coumaroyltryptophane	$C_{20}H_{18}N_2O_4$	$[M-H]^-$	349.1193	349.1188	1.43	349.1196	0.95	229.0619 186.0556	15
18	Feruloylquinic acid	$C_{17}H_{20}O_9$	$[M-H]^-$	367.1034	367.1032	0.49	367.1032	0.54	191.0557 193.0501 173.0450	15
19	Dimethoxycinnamoylquinic acid	$C_{18}H_{22}O_9$	$[M-H]^-$	381.1191	381.1198	0.26	381.1194	0.73	207.0661 173.0449	25
20	Atractyligenin	$C_{19}H_{28}O_4$	$[M-H]^-$	319.1914	319.1912	0.77	363.1916	0.62	275.2021 257.1550 255.1757 83.0490	30
21	Carboxytractyligenin	$C_{20}H_{28}O_6$	$[M-H]^-$	363.1813	363.1809	1.10	363.1814	0.27	319.1920	14
22	Atractyloside II	$C_{25}H_{38}O_9$	$[M-H]^-$	481.2443	481.2439	0.87	481.2444	0.09	319.1920 301.1815	40
23	Coumaroyl-caffeoylquinic acid	$C_{25}H_{24}O_{11}$	$[M-H]^-$	499.1246	499.1247	0.27	499.1251	0.93	353.0886 337.0937 191.0557 173.0450 163.0394	25
24	Mascaroside I or Mozambioside [a]	$C_{26}H_{36}O_{10}$	$[M-H]^-$	507.2236	507.2235	0.19	507.2240	0.80	345.1713 327.1609	28
25	Dicaffeoylquinic acid	$C_{25}H_{24}O_{12}$	$[M-H]^-$	515.1190	515.1191	0.19	515.1193	0.58	353.0886 191.0557 173.0450	20
26	Carboxyatractyloside II	$C_{26}H_{38}O_{11}$	$[M-H]^-$	525.2341	525.2337	0.76	525.2344	0.57	481.2451	12
27	Feruloyl-caffeoylquinic acid	$C_{26}H_{26}O_{12}$	$[M-H]^-$	529.1351	529.1349	0.73	529.1356	0.94	367.1041 353.0887 193.0503 179.0345	20
28	Diferuloylquinic acid	$C_{27}H_{28}O_{12}$	$[M-H]^-$	543.1507	543.1515	1.47	543.1511	0.73	367.1042 193.0502	20
29	Atractyloside III	$C_{30}H_{46}O_{10}$	$[M-H]^-$	565.3018	565.3012	1.00	565.3022	0.70	481.2453 463.2350 319.1921 301.1815	40
30	Carboxyatractyloside III	$C_{31}H_{46}O_{12}$	$[M-H]^-$	609.2916	609.292	0.67	609.2922	0.98	565.3030	15
31	Atractyloside V	$C_{32}H_{52}O_{14}$	$[M-H]^-$	659.3284	659.3290	0.91	659.3286	0.30	335.2235	40
32	Atractyloside I	$C_{36}H_{56}O_{15}$	$[M-H]^-$	727.3546	727.3543	0.41	727.3545	0.16	643.2987 625.2880	15

(*continued on next page*)

Table 1 (*continued*)

N°	Compound	Formula		Theoretical (*m/z*)	Crop 2012		Crop 2013		ESI (−) MS/MS	Fragmentation energy (eV)
					Experimental (*m/z*)	Error (ppm)	Experimental (*m/z*)	Error (ppm)		
33	Carboxyatractyloside I	$C_{37}H_{56}O_{17}$	$[M-H]^-$	771.3444	771.3437	0.90	771.3449	0.64	727.3562	15

[a] Compounds that could not be differentiated by MS/MS experiment.

Approximately 30% of the identified compounds were glycosylated diterpenes, such as glycosylated atractyligenins (denominated as atractylosides I, II, III, V) and glycosylated carboxyatractyligenins (denominated as carboxiatractylosides I, II, III) (Lang et al., 2013, 2014), as well as mascaroside I (Shu et al., 2014) or mozambioside (Lang, Klade, Beusch, Dunkel, & Hofmann, 2015), which has a kaurane diterpene structure, similar to cafestol and kaweol, the most abundant diterpenes in coffee oil.

Caffeoyltryptophane (*m/z* 365.1143, $[M-H]^-$) and coumaroyltryptophane (*m/z* 349.1193, $[M-H]^-$) were also detected, corresponding to 6% of the total identified composition. These compounds have been identified in green coffee beans by ESI-MS (Alonso-Salces et al., 2009; Garrett et al., 2012) and in roasted robusta coffee (Rodrigues & Bragagnolo, 2013). Hexose and sucrose were also be identified (Table 1).

3.2. Postharvest process differentiation by targeted ESI-MS/MS analysis

Compounds variation between the two postharvest processes were accessed by *t*-student test considering 5% of significance level using values from the relative areas of their $[M-H]^-$ ion peaks.

Compounds with significant variations (p-value < 0.05) for the 2012 crop were palmitic, linoleic and oleic acid; coumaroylquinic acid, feruloylquinic acid diferuloylquinic acid; atractyloside I, II and III; carboxyatractyloside I and II. For the 2013 crop, however, it was observed a distinct profile, composed mainly of ferulic acid; linoleic and oleic acid; sucrose; coumaroylquinic acid, coumaroyl-caffeoylquinic acid, diferuloylquinic acid, dimethoxycinnamoylquinic acid, atractyloside I, II and III; carboxyatractyligenin; and carboxyatractyloside I, II and III (see Table 2).

Differences between the PC and NC samples for both crops were observed for linoleic and oleic acid, coumaroylquinic acid, glycosylated derivatives of atractyligenin and carboxiatractyligenin, such as: atractyloside I, II, and III, carboxyatractyloside I, II and III (Fig. 2), suggesting that these components could be potential markers for group differentiation (see Table 2).

Multivariate data analysis was applied using the partial least squares-discrimination analysis (PLS-DA). Fig. 3 shows two PCs that best

Table 2
Chemical compounds and their p-values by *t*-student between PC (pulped natural coffee) and NC (natural coffee) for 2012 and 2013 crop.

N°	Compound	p-value	
		Crop 2012	Crop 2013
1	Caffeic acid	0.410	0.584
2	Hexose	0.292	0.730
3	Ferulic acid	0.805	0.029
4	Quinic acid	0.286	0.915
5	Palmitic acid	0.041×10^{-2}	0.278
6	Linoleic acid	$\mathbf{0.063 \times 10^{-2}}$	**0.027**
7	Oleic acid	$\mathbf{0.022 \times 10^{-2}}$	**0.019**
8	Stearic acid	0.180	0.554
9	Arachidic acid	0.090	0.391
10	Behenic acid	0.556	0.275
11	Caffeoylshikimic acid or caffeoylquinide	0.880	0.338
12	Coumaroylquinic acid	$\mathbf{0.09 \times 10^{-2}}$	$\mathbf{0.05 \times 10^{-4}}$
13	Sucrose	0.490	0.012
14	Caffeoylquinic acid	–	–
15	Ferulic acid hexoside	0.795	0.658
16	Caffeoyltryptophan	0.331	0.398
17	Coumaroyltryptophan	0.062	0.065
18	Feruloylquinic acid	0.025×10^{-1}	0.598
19	Dimethoxycinnamoylquinic acid	0.228	0.019
20	Attractyligenin	0.473	0.312
21	Carboxytractyligenin	0.199	0.027×10^{-3}
22	Atractyloside II	$\mathbf{0.097 \times 10^{-2}}$	$\mathbf{0.021 \times 10^{-1}}$
23	Coumaroyl-caffeoylquinic acid	0.636	0.026
24	Mascaroside I or Mozambioside	0.192	0.085
25	Dicaffeoylquinic acid	0.155	0.550
26	Carboxyatractyloside II	$\mathbf{0.07 \times 10^{-2}}$	$\mathbf{0.043 \times 10^{-2}}$
27	Feruloyl-caffeoylquinic acid	0.502	0.880
28	Diferuloylquinic acid	**0.033**	**0.007**
29	Atractyloside III	$\mathbf{0.070 \times 10^{-4}}$	$\mathbf{0.011 \times 10^{-2}}$
30	Carboxyatractyloside III	0.0524	0.06×10^{-4}
31	Atractyloside V	0.346	0.095
32	Atractyloside I	$\mathbf{0.037 \times 10^{-3}}$	$\mathbf{0.011 \times 10^{-2}}$
33	Carboxyatractyloside I	$\mathbf{0.04 \times 10^{-4}}$	$\mathbf{0.012 \times 10^{-4}}$

(−) 100% relative intensity in all samples; p-value less than 0.05 is statistically significant. Bold numbers indicate common chemical markers for both crops.

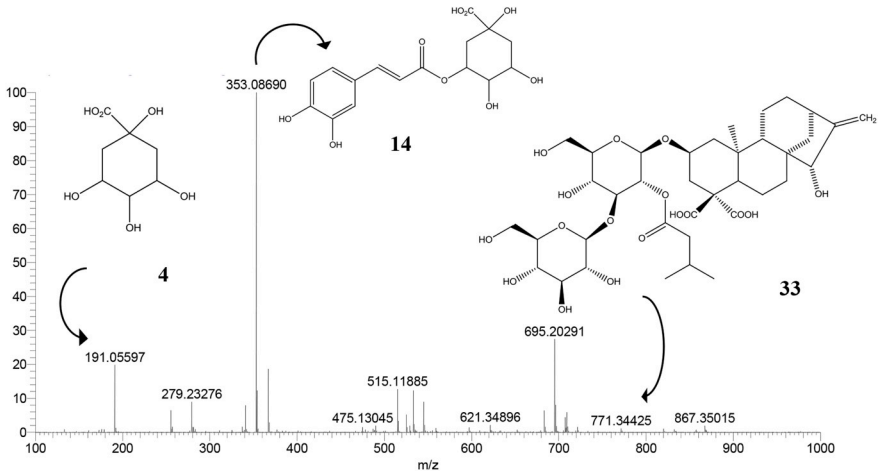

Fig. 1. Typical ESI(−) FT-ICR mass spectra for the methanol:water (4:1) extract of green Arabica coffee beans. The x-axis is the *m/z* value and the y-axis is the relative abundance of coffee substances. (**4**) quinic acid, (**14**) caffeoylquinic acid, (**33**) carboxyatractyloside I. Mass spectrum conditions: negative-ion mode, capillary voltage 1.6 kV, tube lenses 160 V, temperature 270 °C.

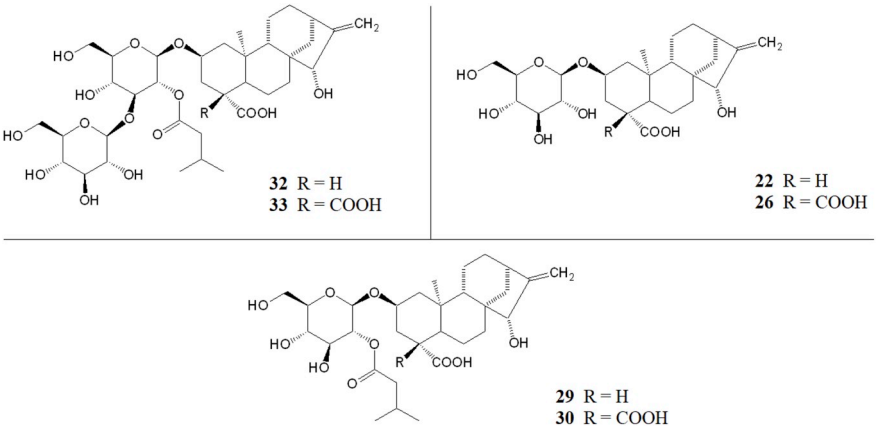

Fig. 2. Important glycosylated derivatives of atractyligenin and carboxiatractyligenin for postharvest differentiation: atractyloside I (**32**), atractyloside II (**22**), atractyloside III (**29**), carboxyatractyloside I (**33**), carboxyatractyloside II (**26**) and carboxyatractyloside III (**30**).

discriminated PC and NC for the crops 2012 and 2013. Three latent variables were chosen for the classification model. The reliability of the classification model among the coffee processes was evaluated by the root mean square error of calibration (RMSEC) and cross-validation (RMSECV), coefficient of determination for calibration (R^2) and coefficient of determination for cross-validation (Q^2) (Table 3) (Bassbasi, de Luca, Ioele, Oussama, & Ragno, 2014). The R^2 and Q^2 values for 3 latent variables showed good adjustment of the classification model (Table 3). PLS-DA score and loading plots are shown on Fig. 3 and a separation between the coffees from the two different postharvest processes is observed. Is important to reinforce that the model has not been tested by predicting a real test set. In this case, all samples are correctly classified according to the cross-validation.

In 2012 crop, the first 3 components (PC1 to 3) (Table 3) explained 77.1% for calibration and 58.7% for validation of the total variation. For the 2013 crop, these values were 86.9% and 75.6%, respectively. The score plot shows the sample grouping according to mass spectra profile similarities and the loading plot indicates how important a compound (*m/z* value) is to discriminate the groups, which were showed for PC1 *x* PC2 only (Fig. 3).

The chemical compounds that mostly contributed to the separation of the groups in the 2012 and 2013 crops were basically the same as those observed in the *t*-student test. However, glycosylated derivatives of atractyligenin (Fig. 2) were the main classes of compounds responsible for the classification and differentiation of PC and NC, suggesting that these compounds are really important for classification of postharvest processes. In general, they showed the highest loading weights (Fig. 4) and thus are the most contributory variables in class discrimination of the PLS-DA models.

A previous work has shown that the considerable difference in content observed for atractyligenin derivatives in Arabica and Robusta coffees could be used to help in the authentication of green coffee samples (Aeschbach, Kusy, & Maier, 1982). Garrett et al. (2013b) have also demonstrated that atractyloside II was important in the differentiation of two varieties of green Arabica coffees (Catuaí and Sarchimor) grown under the same edaphoclimatic conditions. In addition, Correia et al. (2016) also showed that atractyloside II, identified in Arabica and Robusta coffees subjected to dark roast, could be used to differentiate these two species. More recently, Souard et al. (2018)

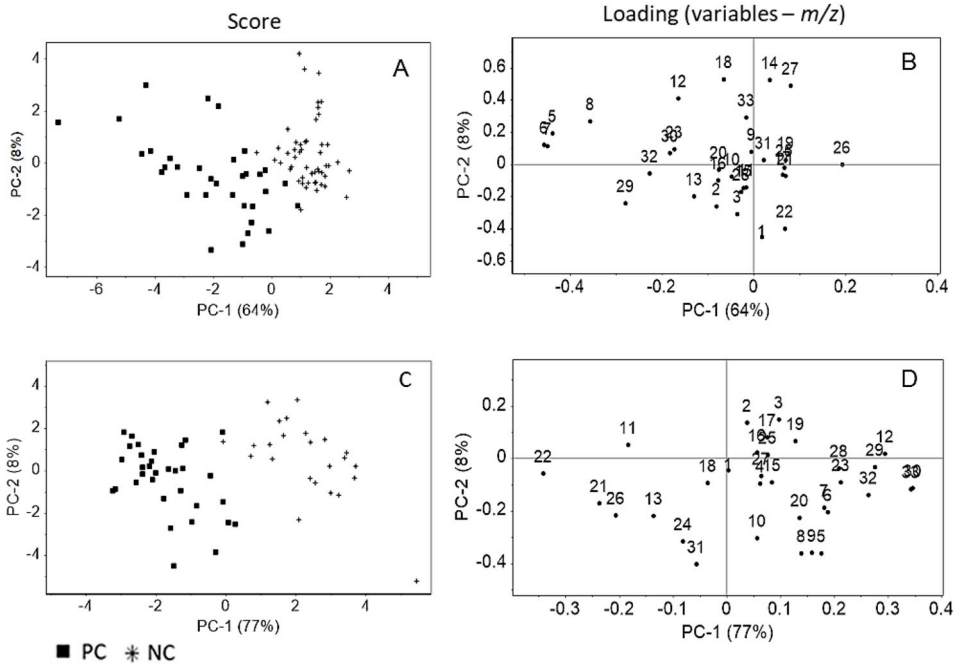

Fig. 3. PLS-DA score and loading plots of pulped natural (PC) and natural (NC) processes of coffees from crops 2012 (A and B), and 2013 (C and D). Codes in loading plots are presented in Table 1.

Table 3
Parameters of the classification model and number latent variables by PLS-DA using 33 variables (*m/z*).

Latent variable	Crop 2012					
	RMSEC	$R^2_{calibration}$	r	RMSECV	$Q^2_{cross\text{-}validation}$	r
1	0,5849	0,6414	0,8008	0,6548	0,5506	0,7420
2	0,5178	0,7189	0,8479	0,6418	0,5726	0,7567
3	**0,4671**	**0,7712**	**0,8782**	**0,6326**	**0,5874**	**0,7664**
4	0,4436	0,7937	0,8909	0,6369	0,5866	0,7659
5	0,4174	0,8173	0,9040	0,6490	0,5868	0,7660
6	0,4048	0,8281	0,9100	0,6468	0,5943	0,7709
7	0,4002	0,8320	0,9121	0,6400	0,6013	0,7754
8	0,3986	0,8334	0,9129	0,6388	0,6014	0,7755
Latent variable	Crop 2013					
	RMSEC	$R^2_{calibration}$	r	RMSECV	$Q^2_{cross\text{-}validation}$	r
1	0,4691	0,7742	0,8799	0,5221	0,7271	0,8527
2	0,3810	0,8510	0,9225	0,4727	0,7713	0,8782
3	**0,3577**	**0,8687**	**0,9320**	**0,4924**	**0,7559**	**0,8694**
4	0,3256	0,8912	0,9440	0,5564	0,7130	0,8444
5	0,3154	0,8979	0,9475	0,6015	0,6785	0,8237
6	0,3001	0,9075	0,9526	0,6436	0,6407	0,8004
7	0,2874	0,9152	0,9566	0,6749	0,6253	0,7908
8	0,2834	0,9175	0,9579	0,6901	0,6142	0,7837

RMSEC: Root mean square error of calibration; RMSECV: Root mean square error of cross-validation. R^2: Coefficient of determinaiton for calibration; Q^2: Coefficient of determinaiton for cross-validation.

using LC-HRMS showed that ent-kaurane diterpenoid derivatives in coffee leaves greatly varies according to the coffee species. In this way, it is observed that this class of compounds contributes in many differentiation events among coffee samples, including the postharvest processes proposed in this work.

Besides the importance of atractyligenin derivatives for coffee differentiation, this class of compounds possesses important biological activity. Lang et al. (2013) demonstrated that glycosylated carboxyatractyligenins present in coffee significantly inhibit mitochondrial adenosine-nucleotide-translocase (ANT) activity, leading to reduced respiration, but its content is drastically reduced during coffee roasting. Due to the increased use of extracts and oils from green coffee beans in manufactured products and the concern of the toxicity associated with the atractylosides analogues, the same group has analyzed dietary

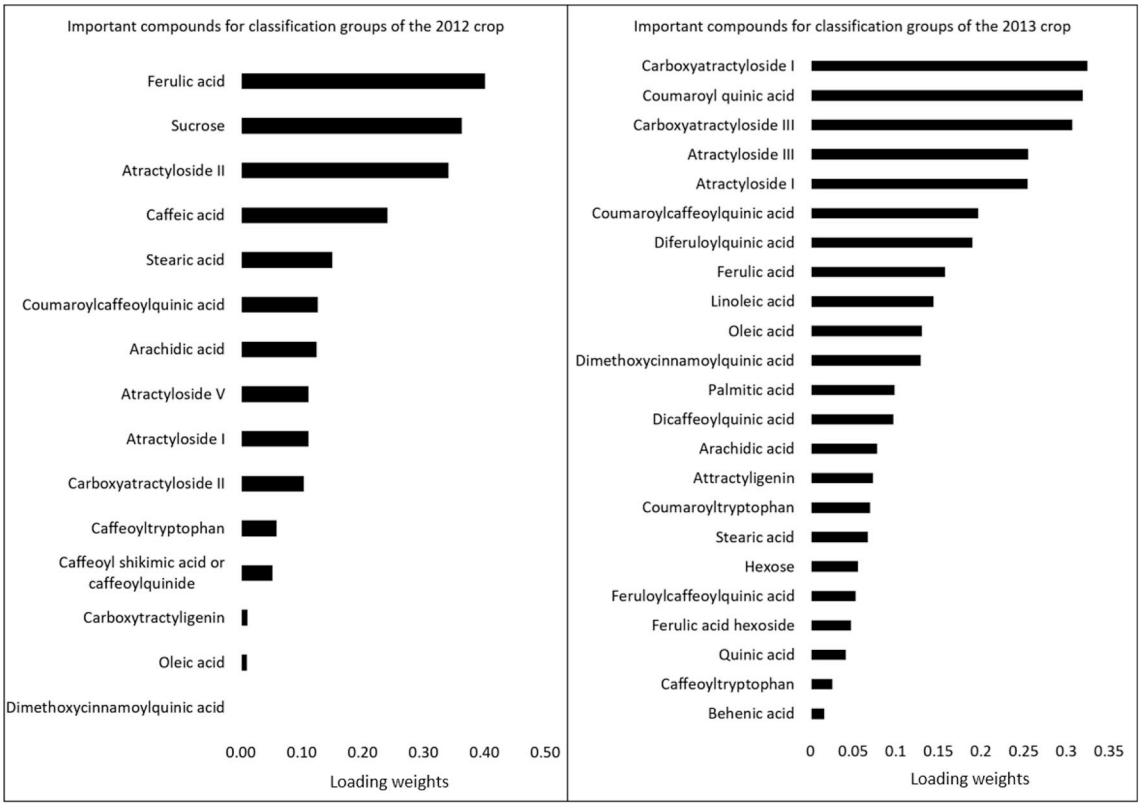

Fig. 4. Compounds with high loading weights and thus higher impact on group (**PC**: pulped natural coffee) and (**NC**: natural coffee) classification by PLS-DA analysis.

supplements based on green coffee and significant levels of atractyligenin derivatives, especially carboxyatractyloside II (**26**), were observed in commercial products based on green coffee beans. This finding implies that more studies are necessary to monitor these constituents in food supplements based on green coffee beans (Lang et al., 2014).

4. Conclusion

The direct analysis of methanol:water extracts of green coffee beans by the high-resolution and accurate ESI (−) FT-ICR MS led to the identification of 33 compounds. Among them, glycosylated atractyligenins and carboxyatractyligenins were identified by statistical methods (*t*-student and PLS-DA) as most important chemical markers to differentiate pulped natural coffees (PC) from natural coffee (NC) in two consecutive crops, showing that despite the expected variation in climate, processing and handling of coffees from one year to another, atractyligenin derivatives could still be used to evaluate coffee processing. Furthermore, samples from the two postharvest processes also showed significant differences regarding the composition of fatty acids and chlorogenic acids. Such distinct chemical composition reflects some specific chemical modifications suffered by green coffee beans when subjected to either of these postharvest treatments.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank CNPq, FAPERJ, CAPES and Embrapa Café for financial support.

References

Aeschbach, R., Kusy, A., & Maier, H. G. (1982). Diterpenoide in Kaffee. *Zeitschrift für Lebensmittel-Untersuchung und Forschung, 175*, 337–341.

Alonso-Salces, R. M., Serra, F., Reniero, F., & Heberger, K. (2009). Botanical and geographical characterization of green coffee (*Coffea arabica* and *Coffea canephora*): Chemical evaluation of phenolic and methylxanthine contents. *Journal of Agricultural and Food Chemistry, 57*, 4224–4235.

Amorim, A. C. L., Hovell, A. M. C., Pinto, A. C., Eberlin, M. N., Arruda, N. P., Pereira, E. J., et al. (2009). Green and roasted Arabica coffees differentiated by ripeness, process and cup quality via electrospray ionization mass spectrometry fingerprinting. *Journal of the Brazilian Chemical Society, 20*, 313–321.

Aquino, F. J. T., Augusti, R., Alves, J. O., Diniz, M. E. R., Morais, S. A. L., Alves, B. H. P., et al. (2014). Direct infusion electrospray ionization mass spectrometry applied to the detection of forgeries: Roasted coffees adulterated with their husks. *Microchemical Journal, 117*, 127–132.

Bassbasi, M., de Luca, M., Ioele, G., Oussama, A., & Ragno, G. (2014). Prediction of the geographical origin of butters by partial least square discriminant analysis (PLS-DA) applied to infrared spectroscopy (FTIR) data. *Journal of Food Composition and Analysis, 33*, 210–2015.

Borém, F. M., Isquierdo, E. P., & Taveira, J. H. S. (2014). In Flavio Meira Borém (Ed.). *Coffee Processing: Handbook of coffee post-harvest technology*(1th ed.). (Chapter 4).

Clifford, M. N., Johnston, K. L., Knight, S., & Kuhnert, N. (2003). Hierarchical Scheme for LC-MS identification of chlorogenic acids. *Journal of Agricultural and Food Chemistry, 51*, 2900–2911.

Clifford, M. N., Knight, S., & Kuhnert, N. (2005). Discriminating between the six isomers of dicaffeoylquinic acid by LC-MS. *Journal of Agricultural and Food Chemistry, 53*, 3821–3832.

Correia, R. M., Loureiro, L. B., Rodrigues, R. R. T., Costa, H. B., Oliveira, B. G., Filgueiras, P. R., et al. (2016). Chemical profiles of robusta and arabica coffee by ESI(-)FT-ICR MS and ATR-FTIR: A quantitative approach. *Analytical Methods, 8*, 7678–7688.

Duarte, G. S., Pereira, A. A., & Farah, A. (2010). Chlorogenic acids and other relevant compounds in Brazilian coffees processed by semi-dry and wet post-harvesting methods. *Food Chemistry, 118*, 851–855.

Ferreira, M. M. C. (2015). *Quimiometria - conceito, métodos e aplicações* (1th ed.). Campinas: Editora da Unicamp (Chapter 5).

Garrett, R., Rezende, C. M., & Ifa, D. R. (2013a). Coffee origin discrimination by paper spray mass spectrometry and direct coffee spray analysis. *Analytical Methods, 5*, 5944–5948.

Garrett, R., Schmidt, E. M., Pereira, L. F. P., Kitzberger, C. S. G., Scholz, M. B. S., Erbelin, M. N., et al. (2013b). Discrimination of Arabica coffee cultivars by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and chemometrics. *LWT – Food Science and Technology, 50*, 496–502.

Garrett, R., Shwab, N. V., Cabral, E. C., Henrique, B. V. M., Ifa, D. R., Eberlin, M. N., et al. (2014). Ambient mass spectrometry employed for direct analysis of intact Arabica coffee beans. *Journal of the Brazilian Chemical Society, 25*, 1172–1177.

Garrett, R., Vaz, B. G., Hovell, A. M. C., Eberlin, M. N., & Rezende, C. M. (2012). Arabica and robusta coffees: Identification of major polar compounds and quantification of blends by direct-infusion electrospray ionization mass spectrometry. *Journal of Agricultural and Food Chemistry, 60*, 4253–4258.

Lang, R., Fromme, T., Beusch, A., Lang, T., Klingenspor, M., & Hofmann, T. (2014). Raw coffee based dietary supplements contain carboxyatractylignin derivatives inhibiting mitochondrial adenine-nucleotide-translocase. *Food and Chemical Toxicology, 70*, 198–204.

Lang, R., Fromme, T., Beusch, A., Wahl, A., Klingenspor, M., & Hofmann, T. (2013). 2-O-β-D-glucopyranosyl-carboxyatractyligenin from *Coffea* L. inhibits adenine nucleotide translocase in isolated mitochondria but is quantitatively degraded during coffee roasting. *Phytochemistry, 93*, 124–135.

Lang, R., Klade, S., Beusch, A., Dunkel, A., & Hofmann, T. (2015). Mozambioside is an arbica-specific bitter-tasting furokaurane glucoside in coffee beans. *Journal of Agricultural and Food Chemistry, 63*, 10492–10499.

Rodrigues, N. P., & Bragagnolo, N. (2013). Identification and quantification of bioactive compounds in coffee brews by HPLC-DAD-MS. *Journal of Food Composition and Analysis, 32*, 105–115.

Selmar, D., Bytof, G., & Knopp, S. E. (2008). The storage of green coffee (*Coffea arabica*): Decrease of viability and changes of potential aroma precursors. *Annals of Botany, 101*, 31–38.

Selmar, D., Bytof, G., Knopp, S. E., & Breitenstein, B. (2006). Germination of coffee seeds and its significance for coffee quality. *Plant Biology, 8*, 260–264.

Selmar, D., Kleinwachter, M., & Bytof, G. (2014). Metabolic responses of coffee beans during processing and their impact on coffee flavor. In R. F. Schwan, & G. H. Fleet (Eds.). *Cocoa and coffee fermentations* (pp. 431–476). New York: CRC Press.

Shu, Y., Liu, J.-Q., Peng, X.-R., Wan, L.-S., Zhou, L., Zhang, T., et al. (2014). Characterization of diterpenoid glucosides in roasted Puer coffee beans. *Journal of Agricultural and Food Chemistry, 62*, 2631–2637.

Souard, F., Delporte, C., Stoffelen, P., Thévenot, E. A., Noret, N., Dauvergne, B., et al. (2018). Metabolomics fingerprint of coffee species determined by untargeted-profiling study using LC-HRMS. *Food Chemistry, 245*, 603–612.

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

Tarzia, A., Scholz, M. B. S., & Petkowicz, C. L. O. (2010). Influence of the postharvest processing method on polysaccharides and coffee beverages. *International Journal of Food Science and Technology, 45*, 2167–2175.

Teixeira, A. A., Brando, C. H. J., Thomaziello, R. A., & Teixeira, R. (2005). The raw beans. In A. Illy, & R. Viani (Eds.). *Espresso coffee – the science of quality* (pp. 87–178). Amsterdam: Elsevier Academic Press.

Toledo, P. R. A. B., Pezza, L., Pezza, H. R., & Toci, A. T. (2016). Relationship between the different aspects related to coffee quality and their volatile compounds. *Comprehensive Reviews in Food Science and Food Safety, 15*, 705–719.