



## Chemical profiling of guarana seeds (*Paullinia cupana*) from different geographical origins using UPLC-QTOF-MS combined with chemometrics



Givaldo Souza da Silva<sup>a</sup>, Kirley Marques Canuto<sup>b</sup>, Paulo Riceli Vasconcelos Ribeiro<sup>b</sup>,  
Edy Sousa de Brito<sup>b</sup>, Madson Moreira Nascimento<sup>a</sup>, Guilherme Julião Zocolo<sup>b</sup>,  
Janlei Pereira Coutinho<sup>a</sup>, Raildo Mota de Jesus<sup>a,\*</sup>

<sup>a</sup> State University of Santa Cruz, Rod. Jorge Amado, km 16, CEP: 45662-900 Ilhéus, BA, Brazil

<sup>b</sup> Embrapa Agroindústria Tropical, Rua Dra Sara Mesquita, 2270, Fortaleza, CE 60511 110, Brazil

### ARTICLE INFO

#### Keywords:

*Paullinia cupana*  
Geographical origin  
QTOF-MS  
PCA  
OPLS-DA

### ABSTRACT

*Paullinia cupana*, commonly known as guarana, is an Amazonian fruit whose seeds are used to produce the powdered guarana, which is rich in caffeine and consumed for its stimulating activity. The metabolic profile of guarana from the two largest producing regions was investigated using UPLC-MS combined with multivariate statistical analysis. The principal component analysis (PCA) showed significant differences between samples produced in the states of Bahia and Amazonas. The metabolites responsible for the differentiation were identified by orthogonal partial least squares discriminant analysis (OPLS-DA). Fourteen phenolic compounds were characterized in guarana powder samples, and catechin, epicatechin, B-type procyanidin dimer, A-type procyanidin trimer and A-type procyanidin dimer were the main compounds responsible for the geographical variation of the samples.

### 1. Introduction

Guarana (*Paullinia cupana*, Sapindaceae), a natural product from the Amazon, is revered for its stimulant and gastronomic qualities (Schimpl, Da Silva, Gonçalves, & Mazzafera, 2013). The interest in guarana is mainly due to the pharmacological action of caffeine, which is associated with increased alertness and improved cognitive and physical abilities. Guarana seeds are rich in caffeine (containing up to 6% dry weight), theophylline and theobromine, terpenes, flavonoids, and starch (Pagliarussi, Freitas, & Bastos, 2002). The caffeine content in guarana seeds is 4–6 times greater than that found in coffee beans, tea leaves, and cola nuts (Claudio, Ferreira, Freire, & Coutinho, 2013; Edwards et al., 2005).

Guarana is marketed in four different forms: raw guarana (roasted grain), guarana stick (raw guarana crushed and pounded, shaped into rod-shaped, dried and smoked), guarana powder (roasted and grounded), and syrups (Brasil, 2003). The majority (70%) of guarana production is directed to the soft drink industry in the form of syrup, and the remainder is consumed in powder or bar form.

Most of the guarana produced on a commercial scale is Brazilian in origin, corresponding to 3554 tons in 2014. The states of Bahia and Amazonas accounted for over 95% of the planted guarana area in

Brazil, with the production of 2691 and 624 tons, respectively (IBGE, 2016). In southern Bahia, soil and climatic conditions similar to the Amazon are found that favor development. Both regions have Tropical Rainforest vegetation. The south of Bahia has an eastern Northeast Tropical climate (super humid, no drought) with an average annual rainfall of 1946 mm and mean temperature of 24.5 °C. The state of Amazonas has an equatorial (super humid, no drought) climate with an average annual precipitation and temperature of 2145 mm and 27.4 °C, respectively. The phenolic composition of guarana grown in these two areas has not been comparatively studied.

Chemical fingerprinting is an analytical approach for sample classification based on metabolic profiles, which can change in response to plant exposure to different types of stimuli such as pathogens, edaphoclimatic factors, and genetic disorders (Dettmer, Aronov, & Hammock, 2007). Chemical fingerprinting techniques enable the verification of possible adulteration in food and reveal the geographical indications of the samples. In this context, fingerprinting has been developed for geographical indication of food products using chromatographic and spectrometric techniques (Gonzalez, Llorens, Cervera, Armenta, & De La Guardia, 2009; Costas-Rodríguez, Lavilla, & Bendicho, 2010; Furia, Naccarato, Sindona, Stabile, & Tagarelli, 2011; Cheajesadagul, Arnaudguilhem, Shiwatana,

\* Corresponding author.

E-mail address: [rmota@uesc.br](mailto:rmota@uesc.br) (R.M. de Jesus).

Siripinyanond, & Szpunar, 2013).

Ultra-performance liquid chromatography (UPLC) coupled with modern mass analyzers such as quadrupole and time-of-flight (Q-ToF) is one of the most commonly used techniques for chemical fingerprinting of plant metabolites, biological samples, and food products. This technique uses chromatographic columns with small particles ( $< 2.0 \mu\text{m}$ ) to improve resolution and peak reproducibility, increase sensitivity, and reduce analysis time (Liu, Zhang, Gao, Jiang, & Chen, 2015; Wan et al., 2013). UPLC-Q-ToF combines the chromatographic separation efficiency with the structural information provided by mass spectrometry (molecular weight and fragmentation pattern) (Nguyen, Guillaume, Rudeaz, & Veuthey, 2006). Mass spectrometry (MS) detection is significantly enhanced by UPLC because increased peak concentrations with reduced chromatographic dispersion at lower flow rates (no flow splitting) promote increased source ionization efficiency (Swartz, 2005). However, as UPLC-MS metabolic profiling generates a large amount of data, chemometric tools have become indispensable for interpretation. Principal component analysis (PCA) and the Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) have been successfully used for food analysis.

PCA is a powerful method for differentiating species, while OPLS-DA may be useful in identifying components (chemical markers) responsible for this difference. OPLS-DA is a PCA-based method of multivariate calibration that uses PCA to reduce the dimensionality of the data set to obtain a correlation. Chromatogram or spectrum (matrix X) regions with the properties of interest (matrix Y) can be analyzed by a multivariate model (Bylesjö et al., 2007; Trygg, Holmes, & Lundstedt, 2007; Westerhuis, Van Velzen, Hoefsloot, & Smilde, 2010; Wiklund et al., 2008).

OPLS-DA and PCA have been combined with LC-MS/MS to determine fruit constituents (Lin, Yang, & Kuo, 2012) for fruit juice authentication (Jandrić et al., 2014), metabolomic analysis and geographic origin of oranges (Díaz, Pozo, Sancho, & Hernández, 2014), and the detection of plant contaminants (Cho et al., 2012; Macherius et al., 2014).

We compared the phenolic profile of guarana seeds (*P. cupana*) originating from two Brazilian producing regions (Bahia and Amazonas) using UPLC-MS and chemometric tools to identify possible chemical markers. We previously used nuclear magnetic resonance to determine the composition of sugars, amino acids, lipids, phenolic compounds, and caffeine content of guarana seeds from these same regions (Silva et al., 2016).

## 2. Material and methods

### 2.1. Materials

Ultrapure water was obtained from Milli-Q water purification system (Millipore, Bedford, MA, USA). Formic acid, acetonitrile, and methanol solvents were LC-MS grade (Merck, Darmstadt, Germany). (+)-Catechin was purchased from Sigma-Aldrich (St. Louis, USA), while (–)-epicatechin, procyanidin A2 [epicatechin-(2 $\beta$  → 7; 4 $\beta$  → 8)-epicatechin], and procyanidin B2 [epicatechin-(4 $\beta$  → 8)-epicatechin] were supplied by Extrasynthese (Lyon, France). PTFE syringe filters (0.22  $\mu\text{m}$ ) were purchased from Millipore. Thirty-five guarana samples were analyzed, and fourteen samples were commercial. Nine samples were purchased in the state of Amazonas (Maués County-AM) and five in the state of Bahia (Ilhéus County-BA). Among the twenty-one samples obtained from farmers, fifteen were harvested in the Low South Bahia region (Valença, Taperoá, and Ituberá Counties), while six samples were harvested in Una County, South Bahia as shown in Supplementary Table 1. The samples were stored at 2 °C until the time of analysis.

### 2.2. Extract preparation

First, 20 mg of ground guarana was suspended in 2.0 mL of 70% methanol. The mixture was sonicated in an ultrasound bath for 2 min, centrifuged for 2 min at 600 rpm, and the supernatant was collected for analysis.

### 2.3. Chromatographic analysis by UPLC-QTOF-MS<sup>E</sup>

The analyses were accomplished on an Acquity UPLC (Waters, USA) coupled to a Xevo Quadrupole and Time-of-Flight mass spectrometer (Q-TOF, Waters). The chromatographic runs were performed on a Waters Acquity BEH UPLC column (150 mm × 2.1 mm I.D., 1.7  $\mu\text{m}$ ) at 40 °C. The mobile phase consisted of water with 0.1% formic acid (A) and acetonitrile with 0.1% of formic acid (B), elution at 2–95% B (0–15 min), 100% B (15.1–17 min), and equilibration with 2% B (17.1–19.1 min) at a flow rate of 0.4 mL min<sup>-1</sup> and an injection volume of 5.0  $\mu\text{L}$ . Ionization was performed with an electrospray ionization source (negative mode ESI), acquired in the range of 110–1180 Da with a source temperature of 120 °C, desolvation temperature of 350 °C, desolvation gas flow rate of 500 L h<sup>-1</sup>, extraction cone of 0.5 V, and a capillary voltage of 2.6 kV. In the low scan, the cone voltage was 35 V with collision energy of 5 eV (trap). In the high scan, the cone voltage was 35 V, and the collision energy ramp was 20–40 eV (trap). Leucine enkephalin was used as the lock mass. The mode of acquisition was MS<sup>E</sup>. The instrument was controlled by MassLynx 4.1 software (Waters Corporation). Three biological replicates were injected in triplicate.

### 2.4. Quantification of catechin and epicatechin

The compounds were quantified using external standard calibration curves of catechin and epicatechin. Stock solutions were prepared with 5 mg of each compound dissolved in 1:1 acetonitrile/water and were transferred to 5 mL volumetric flasks. The preparation of the calibration curves followed dilutions of the stock solution ranging from 2 to 200  $\mu\text{g L}^{-1}$  ( $n = 7$ ). The precursor ion at  $m/z$  289.07 was used as a parameter to determine the areas of the standard solutions and samples.

### 2.5. Statistical analysis

#### 2.5.1. Chemometric analysis

The UPLC-MS data analyzed by chemometrics were processed through MassLynx software (version 4.1, Waters Corporation) under the following conditions: retention time 2.50–6.00 min, mass range 120–1180 Da, and noise elimination level 5. A list of the identities of peaks was created using the retention time ( $t_r$ ) and mass/charge ratio ( $m/z$ ). The intensity of each peak was normalized to the intensities of the peaks in the sample. The alignment method used to generate the PCA was centered on the average. The results of the  $t_r$ - $m/z$  pairs (2389 pairs), sample name, and ion intensity were analyzed by PCA and OPLS-DA.

## 3. Results and discussion

### 3.1. Chemical characterization of guarana samples by UPLC-QTOF-MS<sup>E</sup>

UPLC-QTOF-MS<sup>E</sup> analysis of the guarana seed extracts identified 14 compounds, as shown in Fig. 1I–II and Table 1, as catechin, epicatechin, and procyanidin oligomers (Fig. 2). Although the presence of proanthocyanidin multicharged ions have been reported in LC-MS analysis, this type was not observed in our MS spectra. Ions with  $m/z$  575, 863, and 1151 exhibited a mass difference of 288 Da and were assigned to A-type procyanidin consisting of dimers, trimers, and tetramers, respectively. B-type procyanidins presented ions at  $m/z$  577, 865, and 1153 and were differentiated from the A-type by two mass units (Gu et al., 2003; Li et al., 2012; Qiang et al., 2015; Tala et al., 2013; Zhang & Zhu,

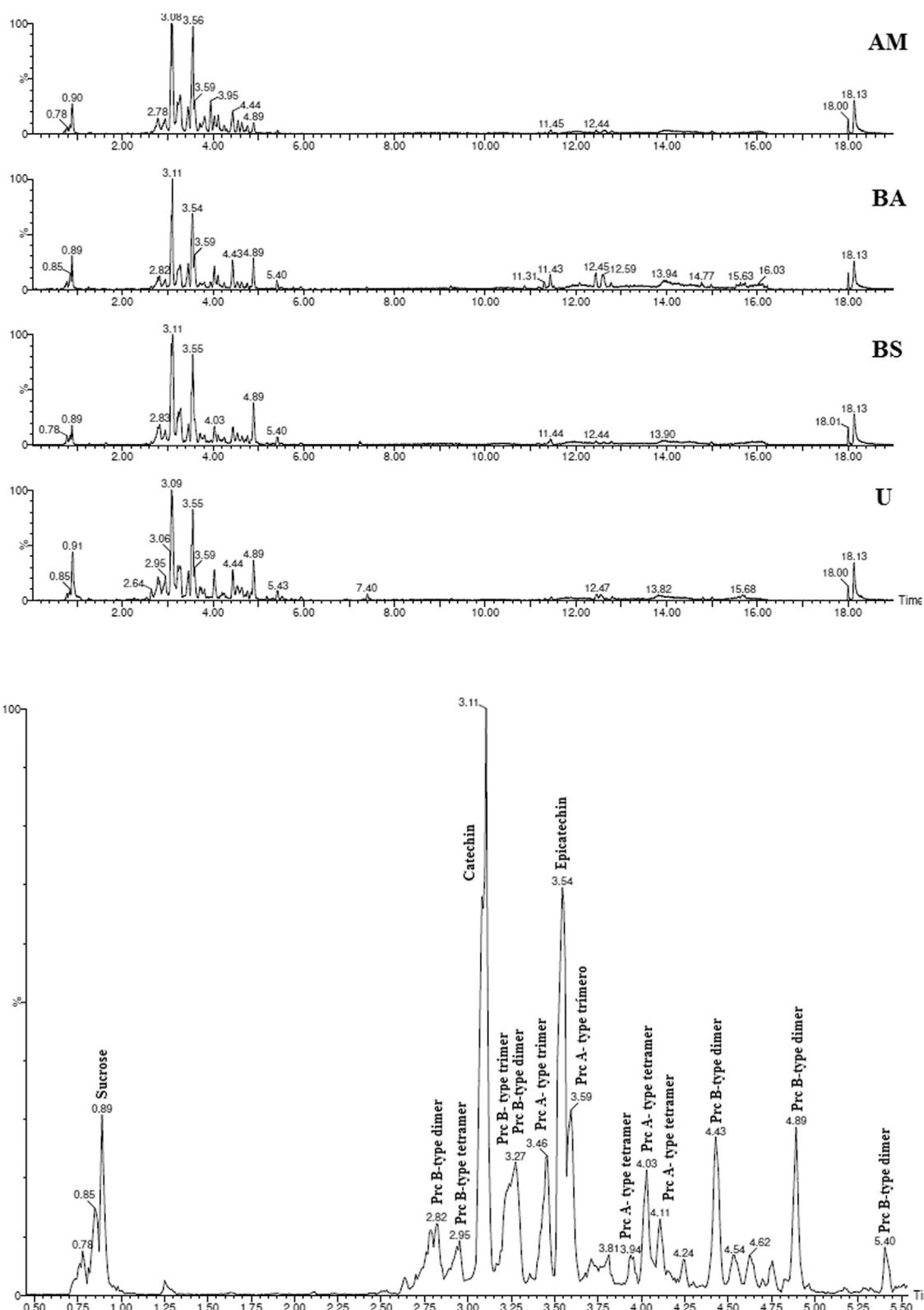


Fig. 1. Chromatogram of guarana seed samples (*Paullinia cupana*) AM (commercial samples from Amazonas State), BA (commercial samples from Bahia State), BS (samples from Low South Bahia State), U (samples from Una County Bahia State) I. Compounds identified in negative mode of the commercial sample from Bahia State (BA) II.

2015). Procyanidins were detected in guarana extracts with a degree of polymerization ranging from 2 to 4 units of condensed (*epi*)catechin. In litchi (*Litchi chinensis* Sapindaceae), the proanthocyanidins were A-type procyanidins, and the degree of polymerization varied according to the fruit part. Lv et al. (2015) used UPLC-QTOF-MS to characterize the

proanthocyanidins existing in 32 cultivar pulps and found procyanidins with up to 6 units of (*epi*)catechin. The proanthocyanidins from pericarp and seeds may contain up to 20 units (Zhou et al., 2011). For guarana, the low degree of polymerization of proanthocyanidins could result from decomposition caused by seed fermentation, which is a

**Table 1**  
Constituents identified or tentatively identified in guarana seeds (*Paullinia cupana*).

Peak no.	Tr min	[M-H] <sup>-</sup> calculated	[M-H] <sup>-</sup> found	Product ions (EM/EM)	Elemental composition	Error ppm	Tentatively identified
1	0.89	341.1080	341.1084	179.0548	C <sub>12</sub> H <sub>21</sub> O <sub>11</sub>	1.2	Sucrose
2	2.78	577.1353	577.1346	451.1090; 425.0877; 407.0781; 289.0736	C <sub>30</sub> H <sub>25</sub> O <sub>12</sub>	1.2	B-type procyanidin dimer
3	2.92	1153.2610	1153.2614	865.2078; 577.1353; 425.0789; 407.0789; 289.0721	C <sub>60</sub> H <sub>49</sub> O <sub>24</sub>	0.8	B-type procyanidin tetramer
4	3.10	289.0713	289.0712	245.0658; 205.0489	C <sub>15</sub> H <sub>13</sub> O <sub>6</sub>	0.3	Catechin <sup>a</sup>
5	3.22	865.1963	865.1980	739.1872; 577.1364; 425.0864; 407.0787; 289.0709	C <sub>45</sub> H <sub>37</sub> O <sub>18</sub>	0.3	B-type procyanidin trimer
6	3.28	577.1347	577.1346	451.1115; 425.0893; 407.0794; 289.0710	C <sub>30</sub> H <sub>25</sub> O <sub>12</sub>	0.2	Procyanidin B2 <sup>b</sup>
7	3.45	863.1849	863.1823	711.1403; 573.1063; 451.1047; 411.0740; 289.0702	C <sub>45</sub> H <sub>35</sub> O <sub>18</sub>	3.0	A-type procyanidin trimer
8	3.55	289.0707	289.0712	245.0778; 205.0495	C <sub>15</sub> H <sub>13</sub> O <sub>6</sub>	1.7	Epicatechin <sup>a</sup>
9	3.59	863.1790	863.1823	711.1334; 573.1038; 451.0998; 411.0697; 289.0720	C <sub>45</sub> H <sub>35</sub> O <sub>18</sub>	3.8	A-type procyanidin trimer
10	3.95	1151.2499	1151.2457	863.1879; 575.1196; 407.0688; 289.0729	C <sub>60</sub> H <sub>47</sub> O <sub>24</sub>	3.6	A-type procyanidin tetramer
11	4.03	1151.2496	1151.2457	863.1858; 575.1148; 407.0743; 289.0789	C <sub>60</sub> H <sub>47</sub> O <sub>24</sub>	3.4	A-type procyanidin tetramer
12	4.42	575.1178	575.1190	449.0874; 423.0763; 407.0761; 289.0748	C <sub>30</sub> H <sub>23</sub> O <sub>12</sub>	2.7	A-type procyanidin dimer
13	4.90	575.1201	575.1190	449.1015; 423.0775; 407.0698; 289.0671	C <sub>30</sub> H <sub>23</sub> O <sub>12</sub>	1.9	Procyanidin A2 <sup>b</sup>
14	5.41	861.1700	861.1667	575.1207; 449.1167; 423.0686; 289.0751	C <sub>45</sub> H <sub>33</sub> O <sub>18</sub>	3.8	A-type procyanidin trimer

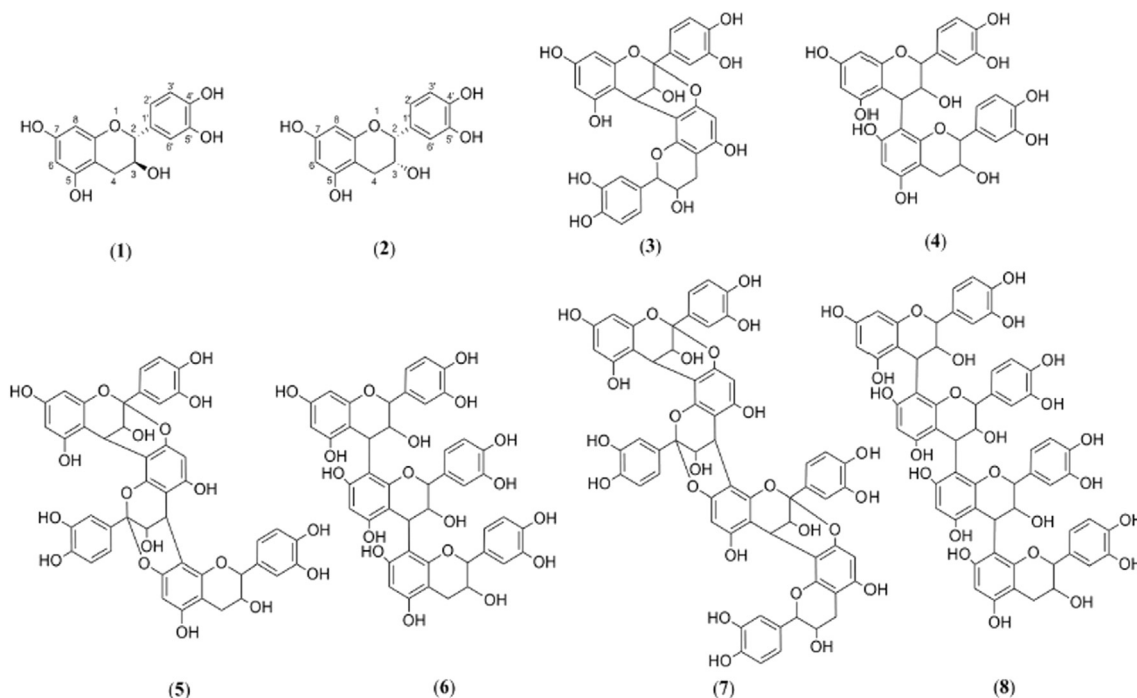
<sup>a</sup> Analytical standards analyzed under the same conditions of the samples.

process commonly used for aril removal. This phenomenon was also observed in cacao beans (D'Souza et al., 2017).

Procyanidins are a class of condensed tannins (proanthocyanidins) in the flavan-3-ol class that have been reported as components in foods and beverages such as grapes, wine, green tea, and guarana. Procyanidins can have a type A or B linkage between the flavonoid units. The B-type is characterized by individually connected flavonoid units between carbon C-4 and C-6 from the chain-extending unit or C-4 and C-8 from the chain termination unit. A-Type procyanidins differ from B-type by having one additional ether bond at C-2 (Ferreira, Marais, Coleman, & Slade, 2010). Guarana proanthocyanidin characterization was enabled by fragmentation mechanisms (Supplementary Figs. 1–3). The most commonly found mechanisms were the fission of heterocyclic ring B with a loss of 126 Da (heterocyclic ring forming fission, HRF), retro Diels-Alder with loss of 152 Da, water elimination, benzofuran formation (benzofuran-forming fission, BFF), and quinone

methide fission (QM).

Compounds **2** and **6** gave a precursor ion at  $m/z$  577. A fragmentation proposal is shown in Supplementary Fig. S1 for MS/MS fragments at  $m/z$  451 [M-H-126]<sup>-</sup>, derived from the HRF and exit of 1,3,5-trihydroxybenzene (phloroglucinol). The ions at  $m/z$  425 [M-H-152]<sup>-</sup> and 407 [M-H-152-18]<sup>-</sup> corresponded to retro Diels-Alder of one upper (epi)catechin moiety which resulted in the loss of a galloyl unit and subsequent elimination of water (likely at the C-ring 3-OH position). The precursor ion underwent QM fission with loss of the (epi)catechin unit to yield an ion at  $m/z$  289 [M-H-288]<sup>-</sup>. The structures of **2** and **6** were consistent with B-type procyanidin dimers (Qiang et al., 2015; Zhang & Zhu, 2015). Unfortunately, HR-MS data was not sufficient to determine whether procyanidins were formed by catechin or epicatechin units, arranged via (4 → 6) or (4 → 8)-type linkages, or the absolute configuration at C-4. However, the UPLC-MS data recorded for compound **6** matched the analytical standard of procyanidin B2 and



**Fig. 2.** Structures of the class substances identified in the guarana extract (*Paullinia cupana*). **1** (+)-catechin, **2** (-)-epicatechin, **3** (epi)catechin A-type dimer, **4** (epi)catechin B-type dimer, **5** (epi)catechin A-type trimer, **6** (epi)catechin B-type trimer, **7** (epi)catechin A-type tetramer, and **8** (epi)catechin B-type tetramer.



was identified as a dimer of epicatechin-(4 $\beta$  → 8)-epicatechin.

Compound **3** showed a precursor ion at  $m/z$  1153.2614 and was characterized as a B-type procyanidin tetramer. The MS/MS fragments at  $m/z$  865 and 577 [M-H-288; M-H-576]<sup>−</sup> were attributed to the QM loss of one and two units (epi)catechin, respectively. The  $m/z$  425 [M-H-576-152]<sup>−</sup> fragment was attributed to inter-flavan bond breaking followed by a retro Diels-Alder rearrangement. The fragment at  $m/z$  289 [M-H-864]<sup>−</sup> was typical of (epi)catechin (Qiang et al., 2015; Zhang & Zhu, 2015).

Compounds **4** and **8** were identified as catechin and epicatechin, respectively, using the precursor ion at  $m/z$  289 and the retention time was similar to authentic standards.

Compound **5** exhibited a precursor ion at  $m/z$  865.1980 with fragments at  $m/z$  739 [M-H-126] generated by HRF of the C-ring at the upper (epi)catechin monomer, leading to an A ring output as phloroglucinol. The ion at  $m/z$  577 [M-H-288]<sup>−</sup> was formed from the loss of one upper (epi)catechin unit by a QM mechanism. The formation of ions with  $m/z$  425 [M-H-289-152]<sup>−</sup>, 407 [M-H-289-152-18]<sup>−</sup>, and 289 [M-H-577]<sup>−</sup> were explained by the mechanisms described above. Compound **5** was characterized by these data as a B-type procyanidin trimer (Qiang et al., 2015; Tala et al., 2013).

Procyanidin trimers with only one A-type connection were evident by fragments at  $m/z$  573 and 289 generated from QM cleavage between the middle and the terminal units of (epi)catechin, or at  $m/z$  575 and 289 between the upper and the middle units (Galaverna, Sampaio, Barata, Eberlin, & Fidelis, 2015). Compounds **7** and **9** showed a precursor ion at  $m/z$  863 and fragments at  $m/z$  711, 573, and 289. The fragment at  $m/z$  711 was formed by retro Diels Alder rearrangement from the (epi)catechin unit with loss of 152 Da. Peaks **7** and **9** were identified as an A-type procyanidin trimer with (2 → 7; 4 → 8) or (2 → 7; 4 → 6) linkages between the middle and terminal units of (epi)catechin.

Compounds **10** and **11** showed a precursor ion at  $m/z$  1151 with MS/MS fragments at  $m/z$  863 [M-H-288]<sup>−</sup>, 575 [M-H-576]<sup>−</sup>, and  $m/z$  289 [M-H-862] that were attributed to successive QM cleavages with loss of (epi)catechin units. The ion fragment at  $m/z$  407.0688 [M-H-575-150-18]<sup>−</sup> was formed by RDA followed by a loss of water. Compounds **10** and **11** were characterized as A-type procyanidin tetramers (Qiang et al., 2015; Zhang & Zhu, 2015).

Compounds **12** and **13** exhibited a precursor ion at  $m/z$  575. The fragmentation proposal for this ion is shown in Supplementary Fig. S3 with MS/MS fragments at  $m/z$  449 [M-H-126]<sup>−</sup> after HRF,  $m/z$  423 [M-H-152]<sup>−</sup> from retro Diels-Alder of one (epi)catechin unit, and  $m/z$  289 [M-H-288]<sup>−</sup> from QM cleavage of one (epi)catechin unit. Both compounds were compatible A-type procyanidin dimers (C<sub>30</sub>H<sub>23</sub>O<sub>12</sub>) (Qiang et al., 2015). Compound **13** was identified as epicatechin-(2 $\beta$  → 7; 4 $\beta$  → 8)-epicatechin after comparison with UPLC-MS data of the analytical standard of procyanidin A2.

Finally, compound **14** showed a precursor ion at  $m/z$  861.1667 consistent with a procyanidin trimeric bound through two A-type linkages (Zhang & Zhu, 2015). Table 1 summarizes the compounds identified in guarana seeds.

### 3.2. Principal component analysis of guarana samples from different producing regions

PCA (Fig. 3.I) was performed to discriminate guarana seed samples from different producing regions according to their metabolic profiles as represented by retention time and the mass-to-charge ratio ( $t_r$ - $m/z$ ) from the UPLC-QTOF-MS<sup>E</sup> analysis. PCA showed 79% of the total cumulative variance in diaxial axes PC1 and PC2. A tendency to form two major groups distributed in the first principal component (PC1) was identified. A group of commercial samples from Bahia State (BA) was located to the left of PC1 (negative scores), while the majority of samples from Low South Bahia State (BS) were on the right side of PC1 (positive scores). Samples from Una County BA (U) and Amazonas (AM)

did not show any clustering tendency due to samples from these groups that were close to BA and BS samples, while others presented a randomized distribution.

Fig. 3.II represents weighted (loadings) plots with the  $t_r$  and  $m/z$  of guarana samples. The analysis of these graphs revealed which pairs of  $t_r$ - $m/z$  were responsible for the distribution and grouping of the samples in the score plot. The Low South (BS) samples were differentiated by catechin (**1**) and epicatechin (**8**) with positive scores values at PC1 and PC2, along with the procyanidin B2 (**6**), A-type procyanidin trimer (**9**), A-type procyanidin dimer (**12**), and procyanidin A2 (**13**) with negative scores values at PC1 and PC2. Procyanidin A2 (**13**), A-type procyanidin dimer (**12**) and B-type procyanidin trimer (**9**) with negative values at PC1 and positive at PC2 were responsible for the separation observed in Bahia (BA) commercial samples. In Fig. 3.III, the Hotelling's T<sup>2</sup> plots showed that the metabolic composition of samples was similar. The AM8 sample exhibited outliers in the 95% confidence interval that were considered a result of average variability among replicates in the model.

In a previous study using PCA with <sup>1</sup>H NMR data, we reported that guarana seeds from the Bahia States contained higher contents of caffeine and phenolic compounds than samples from the Amazonas States. In contrast, these latter ones were richer in fatty acids (Silva et al., 2016).

### 3.3. OPLS-DA analysis and origin discrimination

OPLS-DA chemometric analysis was applied to the data UPLC-QTOF-MS<sup>E</sup> to compare the samples according to the origin. OPLS-DA is a supervised analysis method to study ions that contribute to the experimental sample classification (Liu, Peng, Jia, Cai, et al., 2014). In the OPLS-DA model, the classification between two groups can be visualized in the form of scores chart and scatter plot (S-plot). The scores graph (Fig. 4.I) summarizes observations regarding trends and patterns in the data set and separates the noise. The S-plot used the  $t_r$ - $m/z$  to improve visualization of the classification variables by combining the covariance profiles and correlation of weights (loading) from the prediction model component of OPLS-DA. The correlation for the class or group was calculated with a maximum of 1. Ions with a high correlation value and covariance were distant from the origin and contributed to the separation of classes (Liu, Peng, Jia, Zhang, et al., 2014).

In OPLS-DA data, besides the visual separation of sample profiles through the scatter plots, we should observe other parameters such as the adjustment of the generated model. There is risk related to the adjustment of data when using this type of supervised analysis. Validation of the accuracy and reliability of the model was performed with the parameters of R<sup>2</sup>Y and Q<sup>2</sup>. Q<sup>2</sup> was used to check the predictive ability of the model, determine the fraction of Y variation that could be predicted, and evaluate the statistical quality of the model (Wu, Li, Wang, Li, & Feng, 2015). Statistical quality was expressed as a result of the cross validation to provide a qualitative measure of consistency between the predicted and actual data (Worley & Powers, 2013). R<sup>2</sup>Y was used to check the degree of model adjustment from the Y variance explained by the model. When R<sup>2</sup>Y and Q<sup>2</sup> values approach 1, the model should have a better forecast. However, when the R<sup>2</sup>Y and Q<sup>2</sup> are > 0.5, the model has a good prediction (Qi et al., 2014; Shang et al., 2014).

All OPLS-DA models showed clearly separated samples as shown in Fig. 2, and high predictive values as indicated by the corresponding R<sup>2</sup>Y (97% in BA vs AM, 97% in BA vs. U, 98% in BA vs. BS, 89% in BS vs. U) and Q<sup>2</sup> (81% in BA vs. AM, 85% in BA vs. U, 97% in BA vs. BS, 68% in BS vs U). The values of R<sup>2</sup>Y and Q<sup>2</sup> suggested that the models were well established and had good predictive ability.

#### 3.3.1. Commercial samples from Bahia State (BA) × Amazonas State (AM)

Although both groups of samples were from commercial origin, a

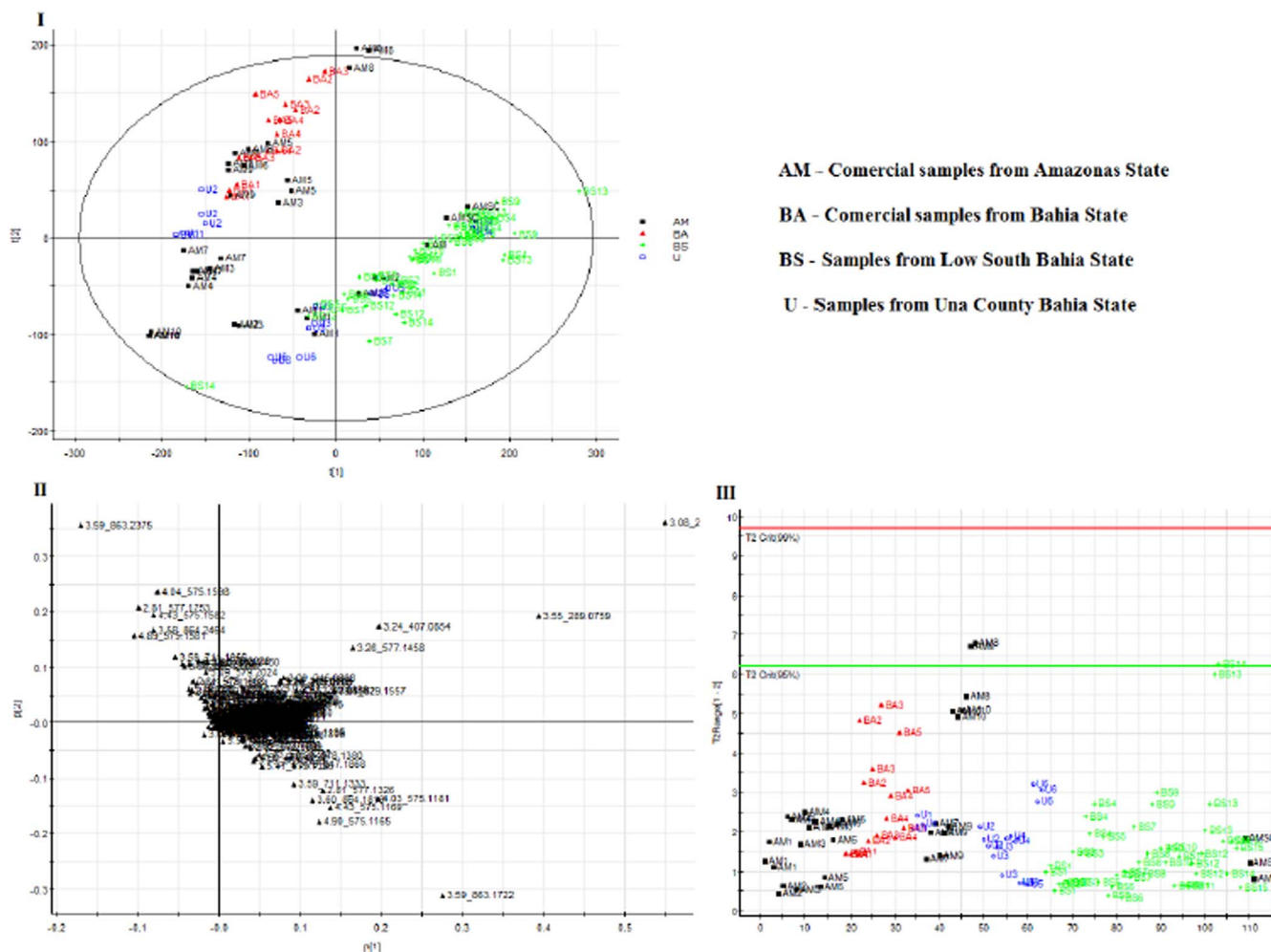


Fig. 3. Data analyzed by PCA score center scaled, negative mode, from guarana seeds (*Paullinia cupana*). In the PCA model, two PCs were used. PC1 and PC2 contributed to 56% and 23% of the total explained variance, respectively (I). Loading graph PC1  $\times$  PC2 of guarana samples by retention time ( $t_R$ ) and mass-to-charge ratio ( $m/z$ ) (II). Hotelling's  $T^2$  graphic showing 95% confidence area (III).

larger intraclass variation occurred in the Amazonas group compared to Bahia commercial samples. This intraclass variation may be related to sampling since the cultivation, origin, and processing for commercial samples may vary.

Fig. 5.I depicts the S-plot with ions responsible for the BA classification found at the bottom of the plot, whereas ions responsible for AM samples were located at the top. Ions that contributed to the separation observed in the score plot were highlighted in red on the S-plot graph as potential chemical markers with each point represented by a  $t_R$ - $m/z$  pair. Six chemical markers were identified.

A-type procyanidin dimers (12 and 13), the B-type procyanidin dimer (2), and the A-type procyanidin trimer (9) were the most distinctive compounds of the Amazonian samples. Epicatechin (8) and catechin (4) were the primary chemical markers of Bahia commercial samples.

### 3.3.2. Commercial samples from Bahia State (BA) $\times$ Una County (U)

On the scatter plot (OPLS-DA) the variance was projected against the correlation variables and the metabolites, which was essential in the model (Fig. 5.II). From the S-plot graph, six ions were highlighted as possible markers and were responsible for classification and separation of BA and Una samples.

Among six major ions, only three have been identified. The compounds procyanidin A2 (13) and A-type procyanidin trimer (9) were responsible for the differentiation of Una samples, while B-type procyanidin dimer (2) was responsible for differentiating the Bahia

commercial samples.

### 3.3.3. Commercial samples from Bahia State (BA) $\times$ Low South Bahia State (BS)

From the S-plot (Fig. 5.III), ten potential pairs of  $t_R$ - $m/z$  were selected to delineate the classes studied (BA  $\times$  BS), but only three compounds were identified. A-type procyanidin dimer (12), procyanidin A2 (13) and A-type procyanidin trimer (9) contributed to the classification of samples from Low South Bahia. Bahia commercial samples exhibited other discriminating ions that could not be identified.

### 3.3.4. Samples from Low South Bahia (BS) $\times$ Una County (U)

In Fig. 5.IV, the S-plot shows the ions ( $t_R$ - $m/z$ ) responsible for the classification and separation of BS and U samples. However, only three compounds were identified: epicatechin (8), catechin (4), and B-type procyanidin dimer (2).

The South Bahia region is divided into “sub-regions” known as Low South, South, and Far South. Una County is located in South. Despite the chemical similarity between samples due to the geographical proximity, epicatechin and catechin showed much higher relative intensities from Low South samples than in Una samples. Una samples showed a greater contribution of B-type procyanidin dimer in their compositions.

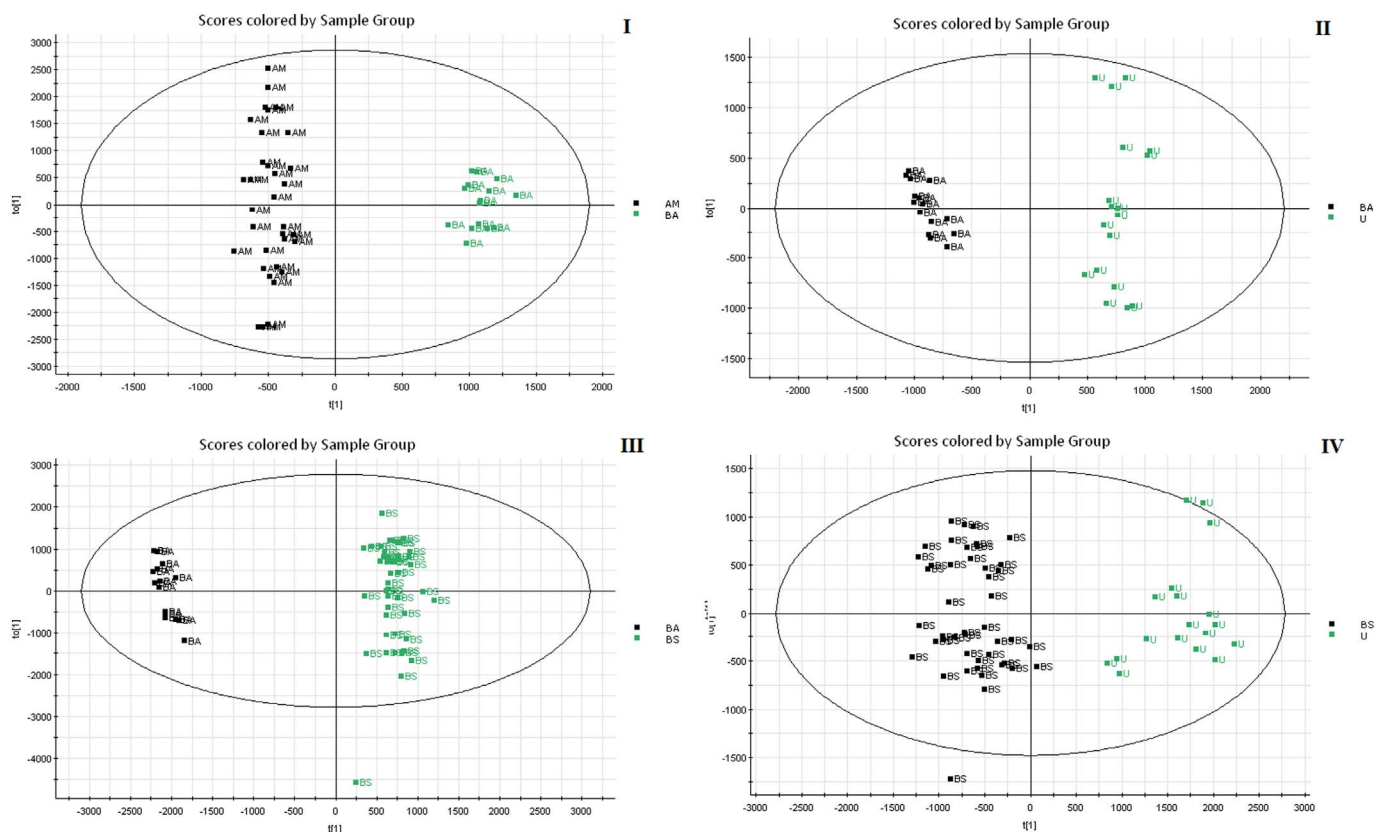


Fig. 4. Score chart in models of two classes according to OPLS-DA: (I) commercial samples of the state of Amazonas (AM) × commercial samples of the state of Bahia (BA); (II) commercial samples of Bahia (BA) × samples of the municipality of Una-Ba (U); (III) commercial samples of Bahia (BA) × samples of the Low South Bahia region (BS); (IV) samples of the Low South Bahia region (BS) × samples of the municipality of Una-Ba (U).

### 3.4. Quantification of catechin and epicatechin

Due to the differences in catechin and epicatechin levels between guarana samples from the three producing regions observed by chemometric analyses, quantitative determinations of these compounds were performed by UPLC-Q-ToF-MS<sup>E</sup>. The analytical curves were linear in the concentration ranges used (Table 2). The  $R^2$  values were above 0.99 for all calibration curves. We used an average of seven concentration levels for each buffer solution for these analyses. The variation coefficient was < 15.0% (ICH, 1995). The results presented were sufficient for the main analytical validation guides cited by ICH, as well as the concentrations of catechin and epicatechin found.

Catechin belongs to a flavonoid group known as flavan-3-ol and has been reported beneficial to human health with chemoprotective activity (Boyer & Liu, 2004; Yang, Wang, Lu, & Picinich, 2009), antioxidant activity (Majhenič, Škerget, & Knez, 2007; Jadhav & Singhal, 2014; Rashidinejad, Birch, & Everett, 2016), anti-inflammatory vascular action (Boyer & Liu, 2004; Velayutham, Babu, & Liu, 2008), and prevention of osteoporosis and neurodegenerative disease (Scalbert, Manach, Morand, Rémésy, & Jiménez, 2005).

Fig. 6 exhibits the variation of catechin and epicatechin in guarana samples. For catechin, the samples from the Low South Bahia had the highest content ( $11.6 \pm 1.3 \text{ mg g}^{-1}$  DW) followed by samples from Una ( $7.2 \pm 1.0 \text{ mg g}^{-1}$  DW), Bahia commercial samples ( $7.1 \pm 1.0 \text{ mg g}^{-1}$  DW) and Amazonas ( $5.2 \pm 0.6 \text{ mg g}^{-1}$  DW). Likewise, samples from Low South Bahia had the greatest content of epicatechin ( $9.2 \pm 0.8 \text{ mg g}^{-1}$  DW), followed by samples from Una ( $4.9 \pm 1.0 \text{ mg g}^{-1}$  DW), Amazonas ( $4.0 \pm 1.0 \text{ mg g}^{-1}$  DW), and Bahia commercial samples ( $3.7 \pm 0.3 \text{ mg g}^{-1}$  DW). In Brazilian commercial samples, Machado et al. (2018) found contents of nearly  $5.0 \text{ mg g}^{-1}$  for catechin (~0.46%) and epicatechin (~0.55%) using an HPLC-UV method. These results were compatible with the amounts

determined for samples from Amazonas State. Procyanidins A2 and B2 were quantified at levels of approximately 0.06% and 0.10%, respectively.

There are other sources of catechin and epicatechin such as cocoa, whose contents can vary respectively from 0.14 to 0.63  $\text{mg g}^{-1}$  and from 0.59 to 5.77  $\text{mg g}^{-1}$  (Natsume et al., 2000), as well as grapes containing from 0.04 to 0.61  $\text{mg g}^{-1}$  (Abe, Mota, Lajolo, & Genovesi, 2007). A low concentration (0.1  $\text{mg g}^{-1}$ ) of epicatechin was found in litchi pulp. The major source of these compounds is green tea (*Camellia sinensis*) at levels from 1.2 to 9.9  $\text{mg g}^{-1}$  catechin and from 6.4 to 21.8  $\text{mg g}^{-1}$  epicatechin (Sabhapondit, Karak, Bhuyan, Goswami, & Hazarika, 2012).

Proanthocyanidins are susceptible to environmental stimuli (e.g., sunlight, temperature, rainfall, climate, and soil) that may influence the activity of enzymes involved in the polyphenol biosynthesis (Makris, Kallithraka, & Mamalos, 2006). Due to these variations, proanthocyanidins are employed as chemical markers to determine origin, varieties, and harvest season of food, beverages, and medicinal plants, and are useful for quality control, geographic indication, and authentication of products as well as chemotaxonomy (D'Souza et al., 2017; Luengas-Caicedo, Braga, Brandão, & Oliveira, 2007; Makris et al., 2006; Naldi et al., 2014; Sivakumaran et al., 2006). For instance, procyanidins B1 and B2 were responsible for the cultivar and origin-based discrimination of young red wine from many Greek viticulture areas (Makris et al., 2006). Phenolic compounds, including proanthocyanidins, were considered chemical markers of origin for Brazilian grape juices (Margraf, Santos, Andrade, van Ruth, & Granato, 2016). Incidentally, it was previously demonstrated that Brazilian grape juices contained higher proanthocyanidin and anthocyanin levels than those produced in Europe (Granato, Koot, Schnitzler, & van Ruth, 2015). This chemical difference may have been due to geography or the type of grape variety.

Proanthocyanidins were the main discriminatory compounds for

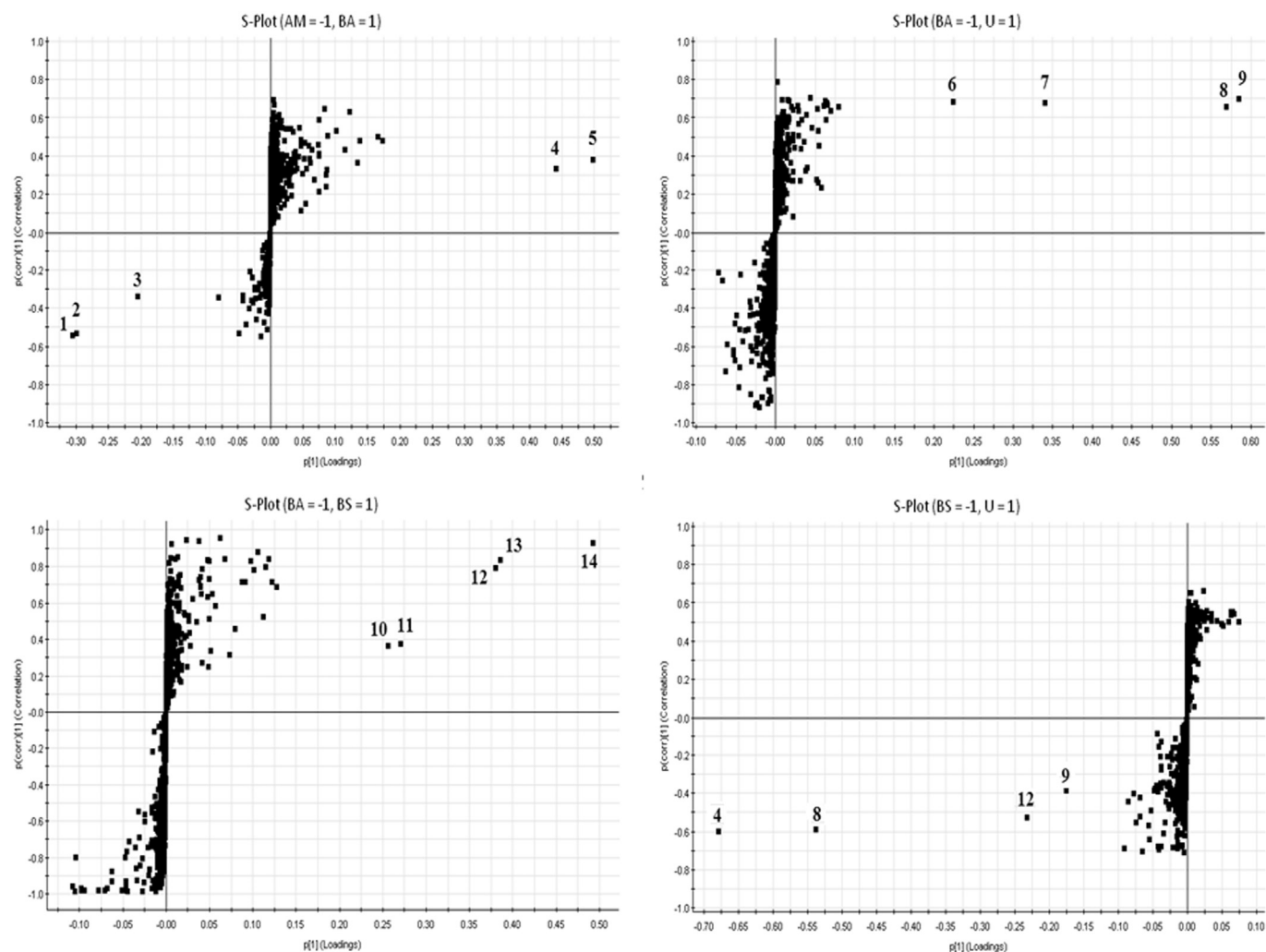


Fig. 5. OPLS-DA/S of products derived from Amazonas commercial (AM), Bahia commercial (BA), Bahia Low South (BS), municipally of Una-Ba (U). 1 (*t*, 4.42 min, *m/z* 575.1178), 2 (*t*, 4.90 min, *m/z* 575.1201), 3 (*t*, 3.59 min, *m/z* 863.1790), 4 (*t*, 3.10 min, *m/z* 289.0713), 5 (*t*, 3.55 min, *m/z* 289.0707), 6 (*t*, 2.78 min, *m/z* 577.1353), 7 (*t*, 4.42 min, *m/z* 575.1178), 8 (*t*, 4.90 min, *m/z* 575.1201), 9 (*t*, 3.59 min, *m/z* 863.1790), 10 (*t*, 3.55 min, *m/z* 289.0707), 11 (*t*, 3.10 min, *m/z* 289.0713), 12 (*t*, 4.42 min, *m/z* 575.1178), 13 (*t*, 4.90 min, *m/z* 575.1201), 14 (*t*, 3.59 min, *m/z* 863.1790).

cocoa grown in six different countries (D'Souza et al., 2017). In the medicinal plant *Cecropia glaziovii*, proanthocyanidin contents were higher in the dry season rather than in the rainy period (Luengas-Caicedo et al., 2007). *Lotus* species were distinguished from each other according to type, the degree of polymerization, and content of proanthocyanidins (Sivakumaran et al., 2006). Meanwhile, Suvanto et al. (2017) have studied Rosaceae and Ericaceae plants to understand why the species from the former predominantly biosynthesize hydrolyzable tannins, whereas the latter ones produce condensed tannins.

The influence of the weather and latitude on the composition of proanthocyanidin from sea buckthorn berries (*Hippophaë rhamnoides* L.) was investigated by Yang, Laaksonen, Kallio, and Yang (2017). These authors found no difference between wild and cultivated berries, even

though berries grown in northern Finland had greater levels of proanthocyanidins than those from southern Finland, where the temperature and solar radiation were higher. Kaundun et al. (1998) found no geographical variability of proanthocyanidin composition (prodelphinidins and procyanidins) among *Pinus halepensis* Millwild populations grown in six countries. However, there was a significant difference between cultivated and native pines. According to these authors, cultivated trees produce more proanthocyanidins in response to a stress condition. Wild trees grow under lower environmental pressure and accumulate fewer secondary metabolites.

The major factors responsible for the variation in the proanthocyanidin profile between guarana seeds from the two largest Brazilian producing regions were difficult to identify, as the samples analyzed did

Table 2

Parameters evaluated in the method validation.

Compound	RT (min)	Linearity ( $\mu\text{g mL}^{-1}$ )	$R^2$	CV % ( $n = 3$ ) ( $\mu\text{g mL}^{-1}$ )			
				2.0	10.0	LOD ( $\mu\text{g mL}^{-1}$ )	LOQ ( $\mu\text{g mL}^{-1}$ )
C	3.10	0.5–200	0.9931	5.50	1.62	0.25	0.5
EC	3.55	2.0–200	0.9959	3.18	5.72	0.25	2.0

RT: retention time.  $R^2$ : coefficient of determination. CV %: coefficient of variation. LOD: limit of detection. LOQ: quantification limit.



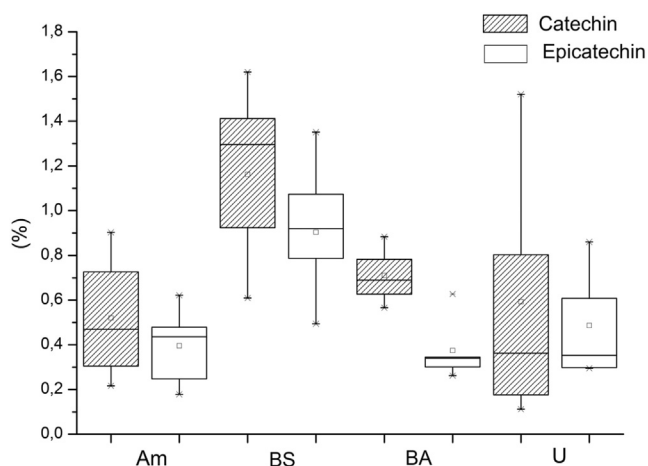


Fig. 6. Levels of catechin and epicatechin in guarana dry extract samples from different regions analyzed by UPLC-MS/MS. Commercial samples from Amazonas State (AM), Low South Bahia (BS), commercial samples from Bahia State (BA), and samples from Una County BA (U). \* Average and standard deviation of samples analyzed in triplicate.

not come from experimental fields. Our guarana samples were not produced under controlled conditions, not even those provided by farmers. Although Bahia and Amazonas states are located in different latitudes, their climates (temperature, rainfall, vegetation, and sunlight) are considered similar as was described earlier. The higher contents of proanthocyanidins found in samples from Bahia appear to be related to the guarana production system. In Bahia, guarana is cultivated in farms, while it is produced in Amazonas by extractivism from wild populations. Accordingly, the chemical variability observed between guarana seeds from these two geographic origins could be at least partially explained by the same hypothesis proposed for pines.

#### 4. Conclusion

LC-MS analyses of the guarana samples identified fourteen compounds characterized as catechin and epicatechin in addition to flavonoid oligomer derivatives known as procyanidins.

PCA applied to the chromatographic and spectroscopic data confirmed the hypothesis of a correlation between the chemical data and the geographical origin of guarana samples. Catechin, epicatechin, A-type procyanidin dimer, B-Type procyanidin dimer, and A-type procyanidin trimer were responsible for the differentiation of samples from Low South Bahia. A-type procyanidin dimer, B-type procyanidin dimer, and B-type procyanidin trimer were the differential compounds from Bahia commercial samples. Additionally, the OPLS-DA-based models showed great potential to predict the origin of guarana samples.

Our study indicated the close relationship of phenolic content to the geographical origin, which may assist in the botanical authentication of these materials and claim geographical indication in the future.

#### Acknowledgements

The authors acknowledge the financial support given by CAPES, CNPq (465335/2014-4), FINEP, and FAPESB (0011/2014). The authors also thank Prof. Dr. Rogete Mendonça from the Federal Institute of Amazonas, who kindly provided the commercial guarana samples from Amazonas State.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodres.2017.09.055>.

#### References

- Abe, L. T., Mota, R. V., Lajolo, F. M., & Genovese, M. I. (2007). Compostos fenólicos e capacidade antioxidante de cultivares de uvas *Vitis labrusca* L. e *Vitis vinifera* L. *Ciências e Tecnologia de Alimentos*, 27(2), 394–400.
- Boyer, J., & Liu, R. H. (2004). Apple phytochemicals and their health benefits. *Nutrition Journal*, 3(5), 1–15.
- Brasil - Ministério do Desenvolvimento, Indústria e Comércio Exterior. Superintendência da Zona Franca de Manaus (SUFRAMA). Projeto Potencialidades regionais Estudo de Viabilidade Econômica do Guaraná. (2003). [http://www.suframa.gov.br/publicacoes/proj\\_pot\\_regionais/guarana.pdf](http://www.suframa.gov.br/publicacoes/proj_pot_regionais/guarana.pdf) (Accessed 09.04.2017).
- Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. (2007). OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics*, 20(8–10), 398–405.
- Cheajesadagul, P., Arnaudguilhem, C., Shiowatana, J., Siripinyanond, A., & Szpunar, J. (2013). Discrimination of geographical origin of rice based on multi-element fingerprinting by high resolution inductively coupled plasma mass spectrometry. *Food Chemistry*, 141, 3504–3509.
- Cho, K., Kim, Y., Wi, S. J., Seo, J. B., Kwon, J., Chung, J. H., ... Nam, M. H. (2012). Nontargeted metabolite profiling in compatible pathogen-inoculated tobacco (*Nicotiana tabacum* L. cv. Wisconsin 38) using UHPLC-Q-TOF/MS. *Journal of Agricultural and Food Chemistry*, 60(44), 11015–11028.
- Claudio, A. F. M., Ferreira, A. M., Freire, M. G., & Coutinho, J. A. P. (2013). Enhanced extraction of caffeine from guaraná seeds using aqueous solutions of ionic liquids. *Green Chemistry*, 15, 2002–2010.
- Costas-Rodríguez, M., Lavilla, L., & Bendicho, C. (2010). Classification of cultivated mussels from Galicia (Northwest Spain) with European protected designation of origin using trace element fingerprint and chemometric analysis. *Analytica Chimica Acta*, 664(2), 121–128.
- Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, 26(1), 51–78.
- Díaz, R., Pozo, O. J., Sancho, J. V., & Hernández, F. (2014). Metabolomic approaches for orange origin discrimination by ultra-high performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry. *Food Chemistry*, 157, 84–93.
- D'Souza, R. N., Grimbs, S., Behrends, B., Bernaert, H., Ullrich, M. S., & Kuhnert, N. (2017). Origin-based polyphenolic fingerprinting of *Theobroma cacao* in unfermented and fermented beans. *Food Research International*, 99, 550–559.
- Edwards, H. G. M., Farwell, D. W., de Oliveira, L. F. C., Alia, J. M., Le Hyaric, M., & de Almeida, M. V. (2005). FT-Raman spectroscopic studies of guarana and some extracts. *Analytica Chimica Acta*, 532, 177–186.
- Ferreira, D., Marais, J. P. J., Coleman, C. M., & Slade, D. (2010). Proanthocyanidins: Chemistry and biology. In L. Mander, & H. B. Liu (Eds.). *Comprehensive natural products II* (pp. 605–661). Elsevier Science.
- Furia, E., Naccarato, A., Sindona, G., Stabile, G., & Tagarelli, A. (2011). Multielement fingerprinting as a tool in origin authentication of PGI Food Products: Tropea red onion. *Journal of Agricultural and Food Chemistry*, 59, 8450–8457.
- Galaverna, R. S., Sampaio, P. T. B., Barata, L. E. S., Eberlin, M. N., & Fidelis, C. H. V. (2015). Differentiation of two morphologically similar Amazonian *Aniba* species by mass spectrometry leaf fingerprinting. *Analytical Methods*, 7, 1984–1990.
- Gonzálvez, A., Llorens, A., Cervera, M. L., Armenta, S., & De La Guardia, M. (2009). Elemental fingerprint of wines from the protected designation of origin Valencia. *Food Chemistry*, 112, 26–34.
- Granato, D., Koot, A., Schnitzler, E., & van Ruth, S. M. (2015). Authentication of geographical origin and crop system of grape juices by phenolic compounds and antioxidant activity using chemometrics. *Journal of Food Science*, 80(3), C584–C593.
- Gu, L., Kelm, M. A., Hammerstone, J. F., Zhang, Z., Beecher, G., Holden, J., ... Prior, R. (2003). Liquid chromatographic/electrospray ionization mass spectrometric studies of proanthocyanidins in foods. *Journal of Mass Spectrometry*, 38(12), 1272–1280.
- Instituto Brasileiro de Geografia e Estatística - IBGE (2016). Levantamento Sistemático da produção Agrícola: pesquisa mensal de previsão e acompanhamento das safras agrícolas no ano civil. 29, 1–78.
- International Conference on Harmonisation (ICH). Validation of analytical procedures: Methodology. (1995). Q2B (CPMP/ICH/281/95) [https://www.uam.es/personal\\_pas/trxf/MU5.pdf](https://www.uam.es/personal_pas/trxf/MU5.pdf) (Accessed 09.04.2017).
- Jadhav, S. B., & Singhal, R. S. (2014). Laccase–gum Arabic conjugate for preparation of water-soluble oligomer of catechin with enhanced antioxidant activity. *Food Chemistry*, 150, 9–16.
- Jandrić, Z., Roberts, D., Rathor, M. N., Abraham, A., Islam, M., & Cannavan, A. (2014). Assessment of fruit juice authenticity using UHPLC-QTOF MS: A metabolomics approach. *Food Chemistry*, 148, 7–17.
- Kaundun, S. S., Lebreton, P., & Fady, B. (1998). Geographical variability of *Pinus halepensis* Mill. as revealed by foliar flavonoids. *Biochemical Systematics and Ecology*, 26, 83–96.
- Li, S., Xiao, J., Chen, L., Hu, C., Chen, P., Xie, B., & Sun, Z. (2012). Identification of A-series oligomeric procyanidins from pericarp of *Litchi chinensis* by FT-ICR-MS and LC-MS. *Food Chemistry*, 133, 31–39.
- Lin, W., Yang, C., & Kuo, B. (2012). Classifying cultivars of rice (*Oryza sativa* L.) based on corrected canopy reflectance spectra data using the orthogonal projections to latent structures (O-PLS) method. *Chemometrics and Intelligent Laboratory Systems*, 115, 25–36.
- Liu, X. W., Zhang, F., Gao, S. H., Jiang, B. B. O., & Chen, W. S. (2015). Metabolite profiling of Zi-Shen pill in rat biological specimens by UHPLC-Q-TOF/MS. *Chinese Journal of Natural Medicines*, 13(2), 145–160.
- Liu, Y. T., Peng, J. B., Jia, H. M., Cai, D. Y., Zhang, H. W., Yu, C. Y., & Zou, Z. M. (2014). UHPLC-Q/TOF MS standardized Chinese formula Xin-Ke-Shu for the treatment of

- atherosclerosis in a rabbit model. *Phytomedicine*, 21(11), 1364–1372.
- Liu, Y. T., Peng, J. B., Jia, H. M., Zhang, H. W., Ding, G., & Zou, Z. M. (2014). Urinary metabonomic evaluation of the therapeutic effect of traditional Chinese medicine Xin-Ke-Shu against atherosclerosis rabbits using UHPLC-Q/TOF MS. *Chemometrics and Intelligent Laboratory Systems*, 136, 104–114.
- Luengas-Caicedo, P. E., Braga, F. C., Brandão, G. C., & Oliveira, A. B. (2007). Seasonal and intraspecific variation of flavonoids and proanthocyanidins in *Cecropia glaziovii* Sneath leaves from native and cultivated specimens. *Zeitschrift für Naturforschung*, 62c, 701–709.
- Lv, Q., Luo, F., Zhao, X., Liu, Y., Hu, G., Sun, C., ... Chen, K. (2015). Identification of proanthocyanidins from Litchi (*Litchi chinensis* Sonn.) pulp by LC-ESI-QTOF-MS and their antioxidant activity. *PLoS ONE*, 10(3), e0120480.
- Machado, K. N., Freitas, A. A., Cunha, L. H., Faraco, A. A. G., Pádua, R. M., Braga, F. C., ... Castilho, R. A. (2018). A rapid simultaneous determination of methylxanthines and proanthocyanidins in Brazilian guaraná (*Paullinia cupana* Kunth.). *Food Chemistry*, 239, 180–188.
- Macherius, A., Seiwert, B., Schröder, P., Huber, C., Lorenz, W., & Reemtsma, T. (2014). Identification of plant metabolites of environmental contaminants by UHPLC-QTOF-MS: The *in vitro* metabolism of triclosan in horseradish. *Journal of Agricultural and Food Chemistry*, 62(5), 1001–1009.
- Majhenič, L., Škerget, M., & Knez, Ž. (2007). Antioxidant and antimicrobial activity of guarana seed extracts. *Food Chemistry*, 104, 1258–1268.
- Makris, D. P., Kallithraka, S., & Mamos, A. (2006). Differentiation of young red wines based on cultivar and geographical origin with application of chemometrics of principal polyphenolic constituents. *Talanta*, 70, 1143–1152.
- Margraf, T., Santos, E. N. T., Andrade, E. F., van Ruth, S. M., & Granato, D. (2016). Effects of geographical origin, variety and farming system on the chemical markers and *in vitro* antioxidant capacity of Brazilian purple grape juices. *Food Research International*, 82, 145–155.
- Naldi, M., Fiori, J., Gotti, R., Périat, A., Veuthey, J. L., Guilleme, D., & Andrisano, V. (2014). UHPLC determination of catechins for the quality control of green tea. *Journal of Pharmaceutical and Biomedical Analysis*, 88, 307–314.
- Natsume, M., Osakabe, N., Yamagishi, M., Takizawa, T., Nakamura, T., Miyatake, H., ... Yoshida, T. (2000). Analyses of polyphenols in cacao liquor, cocoa, and chocolate by normal-phase and reversed-phase HPLC. *Bioscience, Biotechnology, and Biochemistry*, 64(12), 2581–2587.
- Nguyen, D. T. T., Guilleme, D., Rudeaz, S., & Veuthey, J. (2006). Fast analysis in liquid chromatography using small particle size and high pressure. *Journal of Separation Science*, 29(12), 1836–1848.
- Pagliarussi, R. S., Freitas, L. A. P., & Bastos, J. K. (2002). A quantitative method for the analysis of xanthine alkaloids in *Paullinia cupana* (guarana) by capillary column gas chromatography. *Journal of Separation Science*, 25(5–6), 371–374.
- Qi, Y., Pi, Z., Liu, S., Song, F., Lin, N., & Liu, Z. (2014). A metabonomic study of adjuvant-induced arthritis in rats using ultra-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry. *Molecular BioSystems*, 10(10), 2617–2625.
- Qiang, L. V., Luo, F., Zhao, X., Liu, Y., Hu, G., Sun, C., & Li, X. (2015). Identification of proanthocyanidins from Litchi (*Litchi chinensis* Sonn.) Pulp by LC-ESI-QTOF-MS and their antioxidant activity. *PLoS One*, 10(3), 1–17.
- Rashidinejad, A., Birch, E. J., & Everett, D. W. (2016). Antioxidant activity and recovery of green tea catechins in full-fat cheese following gastrointestinal simulated digestion. *Journal of Food Composition and Analysis*, 48, 13–24.
- Sabhapondit, S., Karak, T., Bhuyan, L. P., Goswami, B. C., & Hazarika, M. (2012). Diversity of catechin in northeast Indian tea cultivars. *The Scientific World Journal*, 485193, 1–8.
- Scalbert, A., Manach, C., Morand, C., Rémésy, C., & Jiménez, L. (2005). Dietary polyphenols and the prevention of diseases. *Critical Reviews in Food Science and Nutrition*, 45, 287–306.
- Schimpl, F. C., Da Silva, J. F., Gonçalves, J. F. C., & Mazzafera, P. (2013). Guarana: revisiting a highly caffeinated plant from the Amazon. *Journal of Ethnopharmacology*, 150(01), 14–31.
- Shang, J., Liu, J., He, M., Shang, E., Zhang, L., Shan, M., ... Ding, A. (2014). UHPLC/QTOF MS-based plasma metabolic profiling analysis of the bleeding mechanism in a rat model of yeast and ethanol-induced blood heat and hemorrhage syndrome. *Journal of Pharmaceutical and Biomedical Analysis*, 92, 26–34.
- Silva, G. S., Silva, L. M. A., Alves-Filho, E. G., Canuto, K. M., Brito, E. S., & Jesus, R. M. (2016). <sup>1</sup>H quantitative nuclear magnetic resonance and principal component analysis as tool for discrimination of guarana seeds from different geographic regions of Brazil. *Magnetic Resonance in Food Science 2016 Proceedings*, 21–24.
- Sivakumaran, S., Rumball, W., Lane, G. A., Fraser, K., Foo, L. Y., Yu, M., & Meagher, L. P. (2006). Variation of proanthocyanidins in *Lotus* species. *Journal of Chemical Ecology*, 32, 1797–1816.
- Suvanto, J., Nohynek, L., Seppänen-Laakso, T., Rischer, H., Salminen, J.-P., & Puupponen-Pimia, R. (2017). Variability in the production of tannins and other polyphenols in cell cultures of 12 Nordic plant species. *Planta*, 246, 227–241.
- Swartz, M. E. (2005). UHPLC: An introduction and review. *Journal of Liquid Chromatography & Related Technologies*, 28(7–8), 37–41.
- Tala, V. R. S., Silva, C. V., Rodrigues, C. M., Nkengfack, A. E., Santos, C. L., & Vilegas, W. (2013). Characterization of proanthocyanidins from *Parkia globosa* (Jacq.) G. Don. (Fabaceae) by flow injection analysis - electrospray ionization ion trap tandem mass spectrometry and liquid chromatography/electrospray ionization mass spectrometry. *Journal of Molecular Liquids*, 18(3), 2803–2820.
- Trygg, J., Holmes, E., & Lundstedt, T. (2007). Chemometrics in metabonomics. *Journal of Proteome Research*, 6(2), 469–479.
- Velayutham, P., Babu, A., & Liu, D. (2008). Green Tea catechins and cardiovascular health: an update. *Current Medicinal Chemistry*, 15(18), 1840–1850.
- Wan, J. B., Bai, X., Cai, X. J., Rao, Y., Wang, Y. S., & Wang, Y. T. (2013). Chemical differentiation of Da-Cheng-Qi-Tang, a Chinese medicine formula, prepared by traditional and modern decoction methods using UHPLC/QTOFMS-based metabonomics approach. *Journal of Pharmaceutical and Biomedical Analysis*, 83, 34–42.
- Westerhuis, J. A., Van Velzen, E. J. J., Hoefsloot, H. C. J., & Smilde, A. K. (2010). Multivariate paired data analysis: Multilevel PLS-DA versus OPLS-DA. *Metabolomics*, 6(1), 119–128.
- Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., ... Trygg, J. (2008). Visualization of GCITOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry*, 80(1), 115–122.
- Worley, B., & Powers, R. (2013). Multivariate analysis in metabolomics. *Current Metabolomics*, 1(1), 92–107.
- Wu, X., Li, Y., Wang, Q., Li, W., & Feng, Y. (2015). Effects of berberine and pomegranate seed oil on plasma phospholipid metabolites associated with risks of type 2 diabetes mellitus by U-HPLC Q-TOF-MS. *Journal of Chromatography B*, 1007, 110–120.
- Yang, S. C., Wang, X., Lu, G., & Picinich, S. C. (2009). Cancer prevention by tea: animal studies, molecular mechanisms and human relevance. *Nature Reviews Cancer*, 9(6), 429–439.
- Yang, W., Laaksonen, O., Kallio, H., & Yang, B. (2017). Effects of latitude and weather conditions on proanthocyanidins in berries of Finnish wild and cultivated sea buckthorn (*Hippophaë rhamnoides* L. ssp. *rhamnoides*). *Food Chemistry*, 216, 87–96.
- Zhang, S., & Zhu, M. J. (2015). Characterization of polyphenolics in grape pomace extracts using ESI Q-TOF MS/MS. *Journal of Food Science and Nutrition*, 1(01), 1–10.
- Zhou, H. C., Lin, Y. M., Li, Y. Y., Li, M., Wei, S. D., Chai, W. M., & Tam, N. F.-Y. (2011). Antioxidant properties of polymeric proanthocyanidins from fruit stones and pericarps of *Litchi chinensis* Sonn. *Food Research International*, 44, 613–620.