LWT - Food Science and Technology 50 (2013) 496–502

Contents lists available at SciVerse ScienceDirect

LWT - Food Science and Technology

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

Discrimination of arabica coffee cultivars by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and chemometrics

Rafael Garrett [a,*], Eduardo M. Schmidt [b], Luiz Filipe P. Pereira [c,d], Cíntia S.G. Kitzberger [c], Maria Brígida S. Scholz [c], Marcos N. Eberlin [b], Claudia M. Rezende [a,*]

[a] Aroma Analysis Laboratory, Institute of Chemistry, Federal University of Rio de Janeiro, 21945-970 Rio de Janeiro, RJ, Brazil
[b] ThoMSon Mass Spectrometry Laboratory, Institute of Chemistry, State University of Campinas, 13083-970 Campinas, SP, Brazil
[c] Instituto Agronômico do Paraná, IAPAR, 86001-970 Paraná, PR, Brazil
[d] Embrapa Café, Brasília, DF, Brazil

ARTICLE INFO

Article history:
Received 5 May 2012
Received in revised form
21 August 2012
Accepted 21 August 2012

Keywords:
Coffea arabica
Genetic background
FT-ICR MS
PCA
PLS-DA

ABSTRACT

Green beans of *Coffea arabica* (arabica coffee) cultivars with Sarchimor and Catuaí genetic background were grown under the same edaphoclimatic conditions, in two different regions, and analysed by direct-infusion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS) followed by a metabolomic approach. A total of 20 coffee metabolites including phenolic compounds, fatty acids, sucrose, and diterpene glycosides were identified by negative ESI FT-ICR MS with a mass error <2 ppm. Furthermore, the multivariate data analysis techniques principal components analysis (PCA) and partial least squares-discriminant analysis (PLS-DA) successfully discriminated the arabica coffee cultivars with Sarchimor from those with Catuaí genetic background, and also revealed the correlations between the coffee metabolites with the arabica cultivars and their growing region. This methodology could be used to identify coffee cultivars according to their genetic background as well as coffees from different growing regions, being a valuable tool for traceability and certification processes.

1. Introduction

Coffea arabica L., commonly known as arabica coffee, is the coffee species most appreciated by consumers. It represents around 70% of the total world coffee production and provides a high-quality brew compared to *Coffea canephora* var. *robusta* due to its intense aroma, low bitterness, and low caffeine content (Lashermes & Anthony, 2007). There are numerous arabica coffee cultivars available and many of them are generated by artificial crosses and mutations to obtain more productive cultures, adapted to the various climates, soil conditions, and resistant to pests and diseases (Sera, 2001). Sarchimor coffee cultivar is derived from the crossing of Villa Sarchi with Timor Hybrid coffee (Silveira et al., 2003). The Timor Hybrid is a natural hybrid of *C. arabica* × *C. canephora* from the island of Timor (Setotaw et al., 2010). The traits inherited from *C. canephora* make Sarchimor derived cultivars resistant to biotic stresses, such as coffee leaf rust (*Hemileia vastatrix* Berk. and Br.), and therefore, more cost efficient than traditional cultivars (Bertrand, Guyot, Anthony, & Lashermes, 2003). Catuaí is a traditional-Brazilian arabica cultivar derived from Caturra x Mundo Novo crossing. It is very productive and account for approximately 40% of all cultivated *C. arabica* in Brazil (Mariuzzo, 2009).

Coffee quality is a demanding analytical task because the genetic background of the plants, associated with agronomic practices and environmental conditions – such as climate, soil type, and altitude – can affect the final chemical composition of the grain (Amorim et al., 2009; Leroy et al., 2006; Vaast, Bertrand, Perriot, Guyot, & Génard, 2006). Furthermore, the genetic variability has been reported to contribute to coffee acidity, sugars, chlorogenic acids, lipids, and caffeine (Guerrero, Suárez, & Moreno, 2001; Kitzberger et al., 2010; Ky et al., 2001).

A powerful way to investigate differences in the profile of secondary metabolites in coffees is the metabolomic approach. Metabolomics is a nontargeted study that involves the characterization of small molecule metabolites in high-dimensional data. By the use of statistical multivariate pattern recognition methods, such as principal components analysis (PCA) and partial least squares-discriminant analysis (PLS-DA), samples can be classified or discriminated, and the biomarkers responsible for samples classification or discrimination can be discovered (Bijlsma et al., 2006; Krastanov, 2010). Ultra-high resolution and mass accuracy mass spectrometers, such as those using Fourier transform ion cyclotron

* Corresponding authors. Tel.: +55 21 2562 7121; fax: +55 21 2562 7106.
E-mail addresses: rafael_garrett@iq.ufrj.br (R. Garrett), crezende@iq.ufrj.br (C.M. Rezende).

0023-6438/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.lwt.2012.08.016

resonance (FT-ICR) analysers, have been successfully applied in metabolomics investigations for many years (Aharoni et al., 2002; Cuadros-Inostroza et al., 2010; Takahashi et al., 2008). FT-ICR MS provides high mass resolution and accuracy with mass errors of 1 ppm or less, which combined with a soft ionization, as electrospray ionization (ESI), enables the determination of molecular formulae from mass measurements only (Corilo et al., 2010; Garrett, Vaz, Hovell, Eberlin, & Rezende, 2012; Marshall & Rodgers, 2004).

Considering the importance of both genetic variability and region of cultivation in coffee quality, and that only a few studies regarding arabica coffee cultivar differentiation based on chemical profile have been described in the literature, we performed the analysis of green beans of arabica coffee cultivars with Sarchimor and Catuaí genetic background grown under the same edaphoclimatic conditions, in two different regions, in Brazil, by direct-infusion ESI(−) FT-ICR MS followed by chemometrics.

2. Material and methods

2.1. Coffee samples and extraction

Five arabica coffee cultivars with Sarchimor genetic background (IAPAR 59, IPR 97, IPR 98, IPR 99 and IPR 104) and three arabica coffee cultivars with Catuaí genetic background (IPR 100, IPR 101 and IPR 105) were grown under the same edaphoclimatic conditions at the Agricultural Technology Park of Cooperative COCARI, in Mandaguari, Paraná state, Brazil (23°32′52″ S, 51°40′15″ W, 655 m above sea level, and average annual temperatures of 20–21 °C). Cherry fruits were selected from the harvest of May to July 2010, washed and sun-dried in patio. Samples were standardized to sieve size of 6.5 mm, had their defective beans removed, and were frozen at −18 °C until analysis. These arabica coffee cultivars described above were also grown under the same edaphoclimatic conditions at the Experimental Field of IAPAR, in Londrina, Paraná state, Brazil (23°18′36″ S, 51°09′56″ W, 585 m above sea level, and average annual temperatures of 21–22 °C). They were harvested and processed similarly as described for the Mandaguari samples.

Green coffee samples (0.5 g) were grounded to sieve size of 0.1 mm in a mill (IKA A11 Basic, Wilmington, USA) and extracted in triplicate with 10 mL of methanol (HPLC grade) for 20 min using an ultrasonic bath (40 kHz, USC-1400, Unique, São Paulo, Brazil). The methanolic extracts (1.0 mL) were centrifuged at 12,100× *g* for 5 min using a microcentrifuge MiniSpin (Eppendorf, Hamburg, Germany). Then, 100 μL of the extracts were diluted in methanol/deionized water (1:1) and used for MS analysis.

2.2. Mass spectrometry and ion identification

Mass spectra fingerprinting and MS/MS data were acquired 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 15–40 eV. Data acquisition was performed along the 100–1000 *m*/*z* range by the Xcalibur 2.0 software.

Identification of the ions was done comparing the *m*/*z* values and MS/MS data obtained by ESI FT-ICR MS in the negative-ion mode with a homemade library of coffee compounds based on literature search (Alonso-Salces, Guillou, & Berrueta, 2009; Amorim et al., 2009; Clarke & Vitzthum, 2001; Clifford, Johnston, Knight, & Kuhnert, 2003; Jaiswal & Kuhnert, 2010) and standards. We considered a match between the experimental *m*/*z* value and the theoretical *m*/*z* value from our library when the mass error was <3 ppm. The isotope distribution pattern of the ions identified was also considered with the proposed chemical formula.

2.3. Multivariate analysis of data

Tables of *m*/*z* values and relative intensities containing the fifty more abundant ions from each sample (which represented, approximately, a relative intensity higher than 1%) were exported from the Xcalibur software, saved as *.csv* files and uploaded into the MetaboAnalyst web server (http://www.metaboanalyst.ca; Xia, Psychogios, Young, & Wishart, 2009) for multivariate analyses. The data were aligned using a mass tolerance of 0.005 *m*/*z* and treated by Pareto scaling to reduce the differences between large and small relative intensities of *m*/*z* values (van den Berg, Hoefsloot, Westerhuis, Smilde, & van der Werf, 2006). Four matrices, where each line represented a sample and each column a variable, were generated and submitted to principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA). Leave-one-out cross-validation (LOOCV) was employed to prevent overfitting and estimate the quality of the PLS-DA analysis. The matrices were as follow: (a) the arabica coffee cultivars grown in Londrina; (b) the arabica coffee cultivars grown in Mandaguari; (c) The arabica coffee cultivars with Sarchimor genetic background grown in Londrina and Mandaguari; (d) the arabica coffee cultivars with Catuaí genetic background grown in Londrina and Mandaguari.

3. Results and discussion

3.1. ESI(−) FT-ICR MS analysis

Fig. 1 shows the ESI(−) FT-ICR mass spectra for the methanolic extract of sample IAPAR 59 (arabica coffee cultivar with Sarchimor genetic background) and sample IPR 105 (arabica coffee cultivar with Catuaí genetic background). Although basically the same set of ions was detected in all coffee samples, the distinction between the studied cultivars of arabica coffee was achieved due to significant and reproducible differences in relative intensities of ions.

Analysis of green coffee beans was chosen rather than analysis of roasted beans to avoid changes in coffee composition due to the roasting process.

The *m*/*z* values of the data matrices of coffees samples grown in Mandaguari and Londrina generated during multivariate analyses were used for compounds identification. A total of 20 compounds, including phenolics, lipids, and diterpenes were identified with a mass error <2 ppm (Table 1) via comparison of the experimental *m*/*z* values with the theoretical *m*/*z* values from our coffee library. Additional confirmation of the structural assignments was done via comparison of the experimental isotopic patterns of the deprotonated molecules with the theoretical ones generated by the Xcalibur software. A mass accuracy of 3 ppm and 2% of isotopic pattern accuracy can usually remove more than 95% of false candidates and significantly reduce the number of possible molecular formulae (Kind & Fiehn, 2006). In addition, dissociation of the deprotonated molecules via MS/MS experiments was also employed to gain more confidence for the structural identification. A feature that points to the reliability of our approach is that all compounds identified have already been described in the literature for coffee samples.

Six diterpene glycosides, known as atractyloside (ATR) and carboxyatractyloside (CATR) analogues were found in all coffee cultivars. Basically, this type of diterpenoid glycosides consists of an aglycone with a perhydrophenanthrene structure and a glycoside moiety made up of glucose with sulphate and/or isovalerate (Obatomi & Bach, 1998). ATRs were first isolated from the roots of *Atractylis gummifera* L. (Asteraceae) and can also be found in other