Food Research International 90 (2016) 133–138

Contents lists available at ScienceDirect

Food Research International

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

Differentiation of cocoa nibs from distinct origins using comprehensive two-dimensional gas chromatography and multivariate analysis

Luciana F. Oliveira [a], Soraia C.G.N. Braga [a], Fabio Augusto [a], Juliana C. Hashimoto [b], Priscilla Efraim [b], Ronei J. Poppi [a,*]

[a] Institute of Chemistry, University of Campinas, P. O. Box 6194, 13084-971 Campinas, São Paulo, Brazil
[b] Department of Food Technology, Faculty of Food Engineering, University of Campinas, P.O. Box 6121, 13083-862 Caempinas, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 11 July 2016
Received in revised form 25 October 2016
Accepted 29 October 2016
Available online 01 November 2016

Keywords:
Cocoa nibs
Solid phase microextraction
Comprehensive two-dimensional gas chromatography
Fisher ratio
Principal component analysis

ABSTRACT

Headspace solid-phase microextraction (HS-SPME) combined with comprehensive two-dimensional gas chromatography was used for identification of volatiles compounds responsible for differentiation between cocoa nibs from Brazil and Ivory Coast. The unfolded GC × GC with Flame Ionization Detector (FID) chromatograms were first aligned using the correlation optimization warping (COW) algorithm and normalized. After that, Fisher ratio was calculated for each variable, and a threshold value was chosen to select the variables that best promote a separation of nibs samples from different sources in a principal component analysis (PCA) model. To identify the relevant compounds for the separation, representative samples of each source were analyzed in the same conditions by GC × GC with detection by quadrupole mass spectrometry. Finally, the average peak volumes for each key compound obtained for both classes were compared using a Student *t*-test and it was possible to identify 15 volatile compounds responsible for differentiation between cocoa nibs from Brazil and Ivory Coast.

1. Introduction

Cocoa nibs, a precursor in making processed chocolate, are obtained after several preprocess steps, where the cocoa bean is fermented, dried, roasted and crushed (Payne, Hurst, Miller, Rank, & Stuart, 2010). The control of these steps is essential to produce cocoa nibs with reliable characteristics, where the fermentation and roasted processes are primordial in the quality of the obtained nibs (Frauendorfer & Schieberle, 2008). The volatile compounds present in the cocoa are responsible for organoleptic properties, such as aroma and flavor. It is known that compounds such as pyrazines, aldehydes (cocoa aroma, nut), esters (fruity aroma), and phenolic compounds (astringent property) are directly related to aroma and they can be used for cocoa and chocolate characterization (Cambrai et al., 2010).

Depending of the origin of the cocoa nibs, different sensorial and chemical characteristics will be transferred to the produced chocolate. Thereby, the knowledge of the volatile compounds present in cocoa nibs from distinct sources is important to determine the quality of chocolate and the veracity of labeling (Saltini, Akkerman, & Frosch, 2013).

The discrimination between cocoa (and derivatives) with different characteristics has been described in the literature (Cambrai et al., 2010; Hernandez & Hernandez & Rutledge, 1994; Nightingale, Cadwallader, & Engeseth, 2012; Yanus et al., 2014). Among the several analytical techniques used for differentiation of cocoa from different origins, data fusion of near infrared spectroscopy and electronic tongue was applied for classification of cocoa beans varieties (Teye, Huang, Takrama, & Haiyang, 2014). More recently, Hori, Kiriyama, and Tsumura (2016) described the use of liquid chromatography and partial least squares-discriminant analysis (PLS-DA) in order to distingue cocoa beans from different growing regions (Dominica, Madagascar, Ghana, Ivory Coast, Venezuela and Ecuador) and phenolic compounds were indicated as important in this PLS-DA model. Also, gas chromatography using FID or mass spectrometry detection has been used to distingue different cocoa and derivatives samples (Cambrai et al., 2010; Frauendorfer & Schieberle, 2008; Rodriguez-Campos, Escalona-Buendía, Orozco-Avila, Lugo-Cervantes, & Jaramillo-Flores, 2011).

Comprehensive two dimensional gas chromatography presents several advantages when compared with traditional gas chromatography, mainly to promote the increase in the resolution power, since co-eluted compounds in the first dimension can be separated in the second one (Marriott, Chin, Maikhunthod, Schmarr, & Bieri, 2012). Moreover GC × GC chromatograms are structured, facilitating the identification of unknown compounds (Adahchour, Beens, & Brinkman, 2008; Fitz et al., 2014; Pierce, Hoggard, Mohler, & Synovec, 2008). For all these reasons GC × GC have been used to analyze a wide range of complex matrices (Dugo et al., 2014; Maikhunthod & Marriott, 2013; Radovic et al., 2014; Welke, Manfroi, Zanus, Lazzarotto, & Alcaraz Zini, 2013; Zeng, Li, Hugel, Xu, & Marriott, 2014). However, comprehensive two dimensional gas chromatography of complex samples generates large data

* Corresponding author.
E-mail address: ronei@iqm.unicamp.br (R.J. Poppi).

http://dx.doi.org/10.1016/j.foodres.2016.10.047
0963-9969/

sets, mainly if the detection is performed by mass spectrometers, since an ion fragment dimension comes up in function of the mass spectra obtained for each retention time (Matos, Duarte, & Duarte, 2012). This intrinsic fact makes necessary the use of chemometric methods in GC × GC for data analysis and interpretation.

Several chemometric methods have been successfully applied in GC × GC data analysis for classification or multivariate calibration (Gröger & Zimmermann, 2011; Hantao et al., 2013a; Kehimkar et al., 2014; Vial et al., 2011; Wei et al., 2013; Yang, Hoggard, Lidstrom, & Synovec, 2013). In order to improve the separation between classes, variable selection can be an interesting alternative since it turns the model less complex and easy for interpretation by the fact that unimportant variables are removed (Brown, Spiegelman, & Denham, 1991). Variable selection can be performed in different ways, such as loadings interpretation in Principal Component Analysis (PCA) models or using heuristic algorithms as genetic algorithm (Araújo et al., 2001; Jarvis & Goodacre, 2005).

Fisher ratio is another method that has been applied in feature selection including bi-dimensional chromatography for various matrices (Hantao et al., 2013b; Marney et al., 2013; Parsons et al., 2015; Pinkerton, Parsons, Anderson, & Synovec, 2015). In a simple way, this methodology can be described as: with the defined classes, a variation between classes for one variable is calculated and for the same variable the variation inside the class is also calculated. If the ratio between these two calculated variations is high, it means that the variable is important to discriminate the class.

The main purpose of this work was the identification of volatile compounds responsible for differentiation between cocoa nibs from Brazil and Ivory Coast using headspace solid-phase microextraction (HS-SPME) combined with comprehensive two-dimensional gas chromatography. Fisher ratio was used for selection of the variables responsible for separation of nibs samples from the different sources in a principal component analysis model.

2. Material and methods

2.1. Samples and extraction materials

Cocoa nibs samples fermented, dried and certified from three different states in Brazil (being 7 samples from state of Pará (PA), 8 samples from state of Bahia (BA) and 7 samples from state of Rondônia (RO)) and Ivory Coast (6 samples) were provided by a local cocoa processing company. Immediately after receipt the samples, they were packed in metallized polyethylene bags in temperature of 4 °C until analysis.

The isolation of the volatile fraction of the samples was performed by HS-SPME, using a procedure previously optimized (Oliveira, Braga, Filgueiras, Augusto, & Poppi, 2014). Aliquots of (1.000 ± 0.005 g) of finely ground sample were enclosed in 15 mL vials with 2 mL of NaCl satured solution. After a 5 min period for sample/headspace equilibration at 60 °C, a 50/30 μm divinilbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fiber was exposed to the sample headspace for 50 min; the fiber was then immediately exposed for 5 min to the injector of the GC × GC systems. Before the extractions the fiber was conditioned for 2 h at 250 °C according manufacturer's instructions.

2.2. GC × GC-FID and GC × GC-QMS

GC × GC-FID analysis were carried out on a prototype based on an Agilent 6890 GC fitted with a lab-made 4-jet cryogenic modulator, already described on the literature (Pedroso, Ferreira, Hantao, Bogusz, & Augusto, 2011). The column set consisted on a 30 m × 0.25 mm × 0.25 μm HP-5 column (Agilent Technologies, Wilmington, DE) connected to a 0.80 m × 0.1 mm × 0.1 μm Solgel Wax column (SGE Analytical Science, Ringwood - Victoria, Australia). The modulation period was set to 6.0 s. The split-splitless injector was operated on splitless mode at 260 °C. The column oven temperature was programmed as follows: 40 °C to 110 °C at 3 °C min^{-1} and 110 °C to 240 °C at 10 °C min^{-1} (staying during 4 min at this temperature −240 °C). High purity (99.9999%) hydrogen at 0.6 mL min^{-1} was used as carrier gas and the acquisition rate of FID detector was 100 Hz.

The same conditions were used in GC × GC–MS analysis, using a prototype built over a Shimadzu QP2010 + fast GC-QMS gas chromatograph (Hantao et al., 2013b). In this case the interface temperature was 260 °C and photomultiplier power was programmed to 0.8 kV up to 10 min of the chromatographic run, and then increased to 0.9 kV until final. The scanned mass range was set from $m/z = 40$ D to 340 D, resulting on a data collection frequency of 25 spectra s^{-1}. Also, for the calculation of 1st dimension linear temperature programming retention indexes (LTPRI) of relevant peaks, analysis of samples spiked with *n*-alkane mixture (C_8 to C_{20}) were also performed. Processing of the chromatographic data was performed using the GCImage software (Zoex Corp., Houston – TX, USA), and identification of the peaks on the GC × GC-QMS chromatograms was performed after search on NIST 2010 mass spectra library.

2.3. Alignment of signal on GC × GC chromatograms

The unfolded chromatograms were aligned in the time axis using the Correlation Optimization Warping (COW) algorithm. First, the chromatograms were split in 10 sections order to minimize the number of points in the optimization step (best number of slack size, window and the reference chromatogram) as described for Skov, van den Berg, Tomasi, and Bro (2006)). After that, the COW algorithm was applied separately for each section of the chromatograms; the aligned sections were merged resulting complete chromatograms which were used to perform Fisher Ratio analysis.

2.4. Fisher ratio and PCA analysis

The sample set was divided in two groups corresponding to Brazilian and Ivorian samples. Fisher ratios were calculated for each variable (*i.e.*, discrete signal points on the $^1t_R \times {}^2t_R$ plane) according to Eq. (3), using the unfolded GC × GC-FID chromatograms. The variation between the two classes, inside the class, and the ratio is calculated using Eqs. (1)–(3), respectively (Pierce et al., 2005).

$$\sigma_{cl}^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 n_i}{(k-1)} \tag{1}$$

where n_i is the number of variables in the *i*th class, $\bar{x}_i$ is the mean for the *i*th class, $\bar{x}$ is the global mean, and k is the number of classes.

$$\sigma_{err}^2 = \frac{\sum \left(\sum (x_{ij} - \bar{x})^2 \right) - \left(\sum (\bar{x}_i - \bar{x})^2 n_i \right)}{N-k} \tag{2}$$

where x_{ij} is the *i*th variable of the *j*th class and N is the total number of samples.

$$f_{ratio} = \frac{\sigma_{cl}^2}{\sigma_{err}^2} \tag{3}$$

These Fisher ratios consist on a vector of scalars one for each variable: wherever a value is higher than a defined threshold, it is supposed that there is a peak relevant to the differentiation between the classes (Brazilian and Ivorian samples) on the corresponding position on the chromatographic $^1t_R \times {}^2t_R$ plane – which was then located on the GC × GC-QMS chromatograms and identified.

In order to visualize the class separation after the variable selection with Fisher ratios a Principal Component Analysis with normalized data was performed using Matlab routines. Assessments of scores plots were used to define the best threshold and consequently the most important variables selected by Fisher ratios. All data analysis