doi: 10.1255/mrfs.5

^{1}H quantitative nuclear magnetic resonance and principal component analysis as tool for discrimination of guarana seeds from different geographic regions of Brazil

Givaldo Souza da Silva[a], Lorena Mara Alexandre e Silva[b]*, Elenilson de Godoy Alves Filho[b], Kirley Marques Canuto[b], Edy Sousa de Brito[b] and Raildo Mota de Jesus[a]

[a]**Santa Cruz State University, Ilhéus-BA, Brazil**

[b]**Embrapa Tropical Agroindustry, Fortaleza-CE, Brazil. Corresponding Author: lorena.mara@embrapa.br**

Guarana (*Paullinia cupana*) is an Amazonian native fruit, whose seeds are rich in caffeine and phenolic compounds. Due to the economic importance of guarana, it is necessary to improve the knowledge of its chemical characteristics, which may vary depending on several factors, such as soil characteristics, crop variety, and environmental conditions. In this context, this study describes the application of quantitative nuclear magnetic resonance associated with principal components analysis to evaluate the composition of guarana seeds from the two largest Brazilian producer regions: Amazonas and Bahia states. The principal component analysis discriminated the samples according to the source, revealing that the seeds from Bahia contain higher caffeine and phenolic contents (catechin, epi-catechin), but the seeds from Amazonas present higher content of fatty acids. Additionally, caffeine was quantified using quantitative nuclear magnetic resonance and liquid chromatography coupled to photodiode array detector, which corroborate the chemometrics. The results show a slightly larger content of caffeine for seeds from Bahia. Therefore, *q*NMR combined with chemometrics was important to detect the metabolic variability of guarana seeds from different regions.

Introduction

Paullinia cupana (Sapindaceae), also known as guarana, is a native plant from Amazon (Brazil), where was traditionally used by ancient indigenous people from Sateré-Mawé tribe in the preparation of a drink with medicinal properties (e.g. diuretic and for treating headache, fever, and cramps) as well as a stimulant drug[1]. The Paullinia genus accounts for one hundred and ninety five species, besides many varieties. Until the 1970's, guarana was produced on a commercial scale only in Maués county (Amazonas state). However, in 1980's the Brazilian government encouraged the cultivation of guarana in the states of Bahia, Mato Grosso, Pará, Acre, and Rondônia. The most successful case of adaptation was in Bahia state that, currently, is the largest producer of guarana [2].

The guarana is commercialized in four different forms: as raw guarana (roasted grain), guarana stick (raw crushed guarana, dried, smoked and shaped as stick), guarana powder (roasted and ground seeds), and syrup (extract from guarana seed)[3], which is the usual commercial form. About 70 % of guarana is utilized by the soft drink industry and the 30 % remaining are exported or used for domestic consumption, mainly as dietary supplement[2]. Guarana seed is rich in caffeine (up to 7 %), presenting 4 to 6 times higher content than found in coffee beans, tealeaves, and cola nuts[4]. The seeds also contain high concentration of polyphenols, especially proanthocyanidins[5] as well as other methylxantines (theophylline and theobromine), terpenes, epi/catechin, and starch[6]. As these metabolites arise from the plant secondary metabolism, their levels can change in response to disease, genetic disorders or environmental conditions.

The "digital printing" or "fingerprint" techniques have become one of the most powerful approaches to food quality control, providing chemical information to characterize, for instance, the so-called geographical origin indication. Currently, several scientific studies have been published on the development of fingerprinting and geographical indications for food using various techniques[7-11]. Traditionally, NMR spectroscopy is used for identification and elucidation of molecules, since it allows the unequivocal characterization of substances in all physical states[12]. Nowadays, it has been used as a useful tool for analysis of complex mixture, such as plant metabolites and food for quality control[13, 14]. However, NMR generates highly complex spectroscopy matrices with elevated similarity, which makes infeasible the visual analysis of the data. Therefore, chemometrics can be used to solve chemical problems based on a large number of samples and measured variables. The aim of this work was to characterize the main compounds found in guarana seeds, and to differentiate guarana seeds produced in different geographical areas of Brazil. In addition, caffeine (major bioactive compound found in guarana) was quantified using quantitative nuclear magnetic resonance (*q*NMR) and high performance liquid chromatography coupled to photodiode array detector (HPLC-PDA) in order to attest its variation indicated in chemometric models.

Experimental

Sample

Guarana seeds (Figure 1a) were collected from the two main production states of Brazil: Amazonas and Bahia (highlighted in Figure 1b). The samples were processed according to the "Serviço Brasileiro de Respostas Técnicas – SBRT (2008)" guidelines[15] with modifications. The fruits of guarana were fermented for two days into plastic bags at room temperature. Thereafter, the seeds were

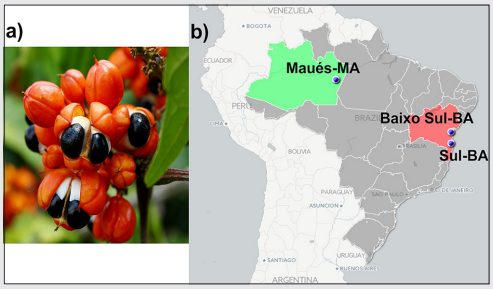

Figure 1. a) *Paullinia cupana* (guarana) fruit. Source: Embrapa archive; b) geographic regions of guarana sampling.

washed with distilled water, peeled, placed in oven at 50 °C for 12 h, and grounded using a Ika™ A11 mill.

NMR measurements and chemometrics

The 30 mg of powder of guarana seeds were suspended in 49 µL of tetradeuterated methanol (CD_3OD-d_4) and 210 µL of deuterated water (D_2O) containing 1.6 mg.mL^{-1} of sodium salt of trimethylsilyl-propanoic and 5.6 mg.mL^{-1} of EDTA (ethylene-diamine-tetra acetic acid). Then, the mixture was sonicated for 2 min, 605 × g was applied for 2 min, and the supernatant was transferred to 5 mm NMR tubes. The NMR experiments were acquired on an Agilent DD2 600-MHz spectrometer (14.1 T), equipped with 5 mm (H-F/^{15}N-^{31}P) inverse detection One Probe™ with actively shielded z-gradient. The ^{1}H NMR experiments were acquired under quantitative conditions: 90° pulse calibrated to 7.45 µs; acquisition time of 3.3 s, and relaxation delay of 23.0 s. It was acquired 32 transients in a spectral window of 16 ppm and 32 k of points at 298 K. The identification of the constituents was performed through two-dimensional NMR experiments as well as the comparison to the existing data[16-19].

For chemometrics, the region between 0.70 and 8.50 ppm (excluding water and methanol resonances) in the ^{1}H NMR spectra were bucketed (0.04 ppm) and used as input variables for the chemometric analysis using the Amix™ v. 3.9.11 (Bruker BioSpin GmbH) program. To perform the principal component analysis (PCA), the bucket tables were pre-processed by mean-centered. In addition, the full cross-validation method was built leaving one of the samples out of the model and then, the sample was predicted by the model. This process was repeated until all the samples have been left out once for each model. Also, in order to determine the significance of climate differences between the regions the forecast data were obtained at the website of the Brazilian Institute of Meteorology and compared with the data using analysis of variance (ANOVA) single factor, available in Origin 8.1 software (Microcal Software Inc., Northampton, MA, USA).

Caffeine quantification

Caffeine was quantified by using two approaches: ^{1}H *q*NMR, and HPLC-UV. For *q*NMR the resonance at 3.98 ppm was found to be a unique marker of the presence of caffeine in guarana[19] and was used for quantification. The sample preparation and the HPLC-UV method were based on the methodology described by Meinhart, et al. (2010)[20], with modifications. The extraction was performed by weighting 10.0 mg of the powdered guarana seeds and mixed with 60 mL of chloroform. To this solution was added 10.0 mL of NaOH 0.2 mol.L^{-1} and the chloroform portion was collected and dried. This dried fraction was mixed with 10.0 mL of H_2O and filtered prior the injection. The HPLC-UV instrument consisted of a Prominence LC-20A from Shimadzu (Japan) coupled with a photodiode array detector (PAD) set to 274 nm and oven set to 40 °C. A volume of 20 µL was injected in a Shim-pack CLC-ODS column (Shimadzu C18 4.6 × 250 mm, 5 µm) with a flow rate of 0.74 mL.min^{-1}. The mobile phase consisted of a combination of solvent A mixed with 0.1 % (v/v) of acetic acid, and solvent B as methanol. The elution gradient was varied linearly from 0 % to 100 % B in 9.85 min, then reduced back to 0 % of B and held there for 5 min. Quantification was performed with a external calibration curve of caffeine ranging from 10.0 mg.L^{-1} to 100.0 mg.L^{-1}. The results were expressed as percentage of caffeine per grams of sample.

Results and discussions

Characterization

The ^{1}H NMR spectrum of the guarana seeds is displayed in Figure 2. The extraction method (aqueous methanol) was used in order to favor the extraction of primary and secondary metabolites from the seeds. Therefore, signals were notice from corresponding low molecular weight primary metabolites such as sugars, organic acids, fatty acids, cyanolipids, as well as the secondary metabolites caffeine, cathechin, and epi-catechin. In the high field region it is observed the characteristics signals of metabolites as fatty acids (1.31 ppm for -CH_2-; 0.88 ppm for -CH_3; and 5.42 ppm for methylene hydrogen); cyanolipids (1.31 ppm for -CH_2-; 0.88 ppm for -CH_3; 4.85 ppm for -CH_2- of the acylglycerol moiety; and 5.42 ppm for methylene hydrogen); valin (1.02 ppm for -CH_3 and 1.07 ppm for the other -CH_3); ethanol (1.24 ppm for -CH_3); threonine (1.33 ppm for -CH_3); alanine (1.49 ppm for -CH_3); gamma-aminobutyric acid named as GABA (alpha -CH_2- at 2.3 ppm, and gamma -CH_2- at 3.02 ppm); acetic acid (1.93 ppm for -CH_3); catechin and epi-catechin with -CH_2- groups at region from 2.50 to 2.90 ppm; malic acid with the -CH_2- group overlapped with catechin and epi-catechin signals. At the sugar region was observed the characteristic signal of malic acid (4.42 ppm for -CH-); α-glucose (5.23 ppm for -CH-);

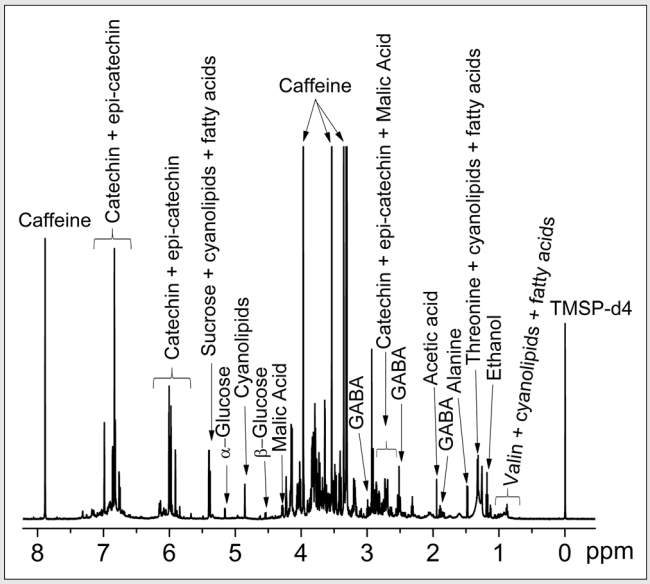

Figure 2. ^{1}H NMR spectrum of guarana seeds showing signals labelled with the compounds identified. Legend: GABA – gamma aminobutyric acid.

β-glucose (4.65 ppm for -CH-), and sucrose (5.42 ppm for -CH-), beyond the signals from the methyl groups of caffeine at 3.41, 3.59, and 4.02 ppm. In the aromatic region was observed the -CH- signals from catechin and epi-catechin from 5.0 to 6.20 ppm and 6.70 to 7.10 ppm.

Chemometrics

In order to identify potential relationships between the geographical origin and chemical composition of guarana samples, the ^{1}H NMR spectra were used as input data for PCA analysis. Figure 3a presents the biplot PC1 *vs*. PC2 showing the separation of the seeds according to the geographical origin. The samples were symbolized regarding the region: Amazonas as gray square; south of the state of Bahia as circle; Baixo sul region as black asterisks, and commercial from Bahia as triangle. At positive scores of PC1 (48.3 % of the total variance) are located the majority of the samples collected in Amazonas. On the other hand, the samples collected at Bahia (from Baixo Sul region and the south of Bahia) are located in negative scores of PC1. It was also observed that the commercial samples (dark triangles) are distributed all over the PCA graphic. This might indicate the poor quality control of the powder, with a mixture of powders from Bahia and Amazonas.

The loadings graph of PC1 (Fig. 3b) presented seeds from Bahia with highest amount of carbohydrates, phenolic compounds as cathechin and epi-catechin, besides the caffeine. In addition, the samples from Amazonas exhibited the highest amount of fatty acids and cyanolipids. These differences were clearly observed on the expansion of the ^{1}H NMR spectra (from 0.6 to 3.2 ppm) of the samples (Figure 3 c, d, and e). This finding indicated that the edaphoclimatic conditions of Bahia (tropical eastern north climate) when compared with Amazon (equatorial climate) seem to have ideal conditions for guarana cultivation (mostly higher phenolic content). Additionally, the temperature (°C) and total rainfall (mm) of the sampling period (September of 2013 to February of 2015)[21] was examined. The variance analysis showed significant differences for temperature wherein Maués-area presented higher temperatures than Bahia-area while rainfalls in both regions were quite similar. The temperature difference might influence the chemical composition of guarana seeds.

Additionally, after the evaluation of the Figure 3c and Figure 3d, it was noticed that the seeds from Baixo sul region (Fig. 3c) and from south of Bahia (Fig 3d) also presented variations on the chemical composition. Therefore, the PCA was also performed for detailed evaluation of these samples. The biplot PC1 *vs.* PC2 presented in Figure 4a shows a tendency of separation of the guarana seeds according to the geographic origin, with samples from Baixo sul region of Bahia (star symbol) at negative scores of PC1, and samples from south of Bahia (circle symbols) at positive scores of PC1. In the PC1 loadings plot (Figure 4b) it is possible to observe that the seeds from Baixo sul region of Bahia presented higher amount of the phenolic compounds, while samples from south of Bahia contained higher amount of carbohydrates. The variance analysis for climate data was also examined showing similarity between the climates for both regions, which shows that this variability can be associates to other parameter beyond the edaphoclimatic condition.

Quantification of caffeine

About 70 % of the guarana production is absorbed by the beverage industry, located in Amazon state [2]. This high demand is driven by the energetic properties of the seed that is mostly attributed to the elevated caffeine content. Therefore, in order to obtain the content of caffeine for the different regions, ^{1}H *q*NMR and HPLC-UV were employed to quantify caffeine in guarana seeds. The Figure 5 shows the box-and-whisker graphs for the quantification of caffeine.

The caffeine content obtained from HPLC-UV method ranges from 1.5 to 5 % per gram of seed while through *q*NMR, the caffeine content ranges from 1 to 3.5 % per gram of seed. Both methods were in accordance with publish data that shows that the caffeine ranges from 2.5 to 6 % [2]. However, the quantification using the HPLC-UV method presented values higher than the ones obtained through the ^{1}H *q*NMR. This occurs because the HPLC-UV experiments were performed with samples that undergo exhaustive extraction yielding up to 98 % of caffeine extraction (data not shown). In addition, according to statistical analysis by ANOVA for

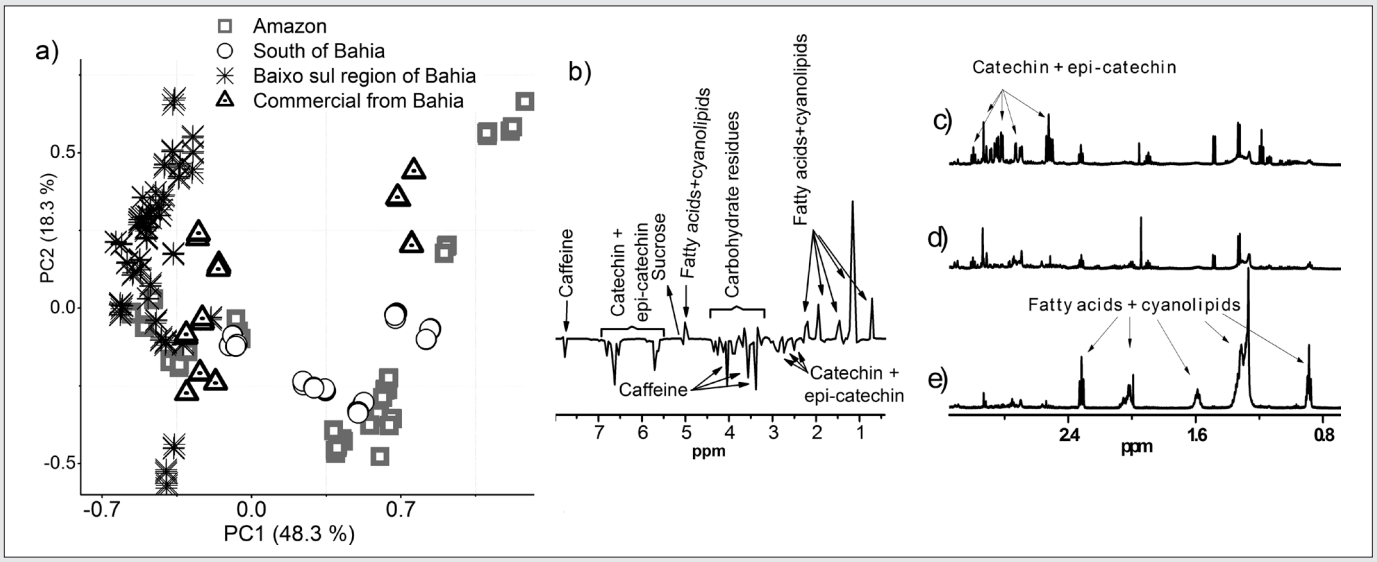

Figure 3. a) PC1xPC2 scores plot of guarana seeds extract from different geographic origin; b) PC1 loadings graph. Comparison among the ^{1}H NMR spectra from: c) seeds of Baixo sul region of Bahia state; d) seeds of south of Bahia state; e) seeds of the Amazonas state.

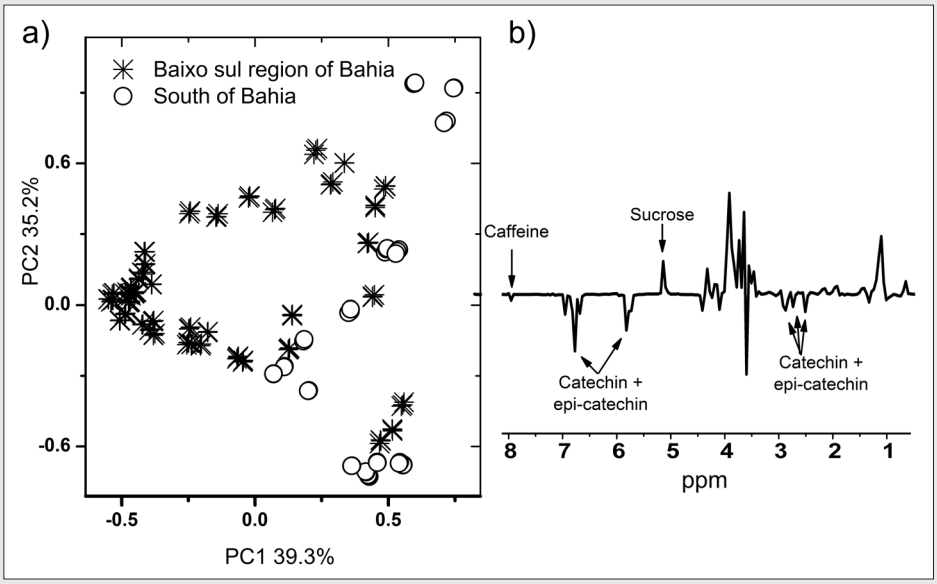

Figure 4. a) PC1 × PC2 of guarana seeds extract from the Bahia state; and b) loadings graph plotted in line from PC1 axis (top) and ^{1}H NMR spectra (bottom) to illustrate the NMR regions.

^{1}H *q*NMR, the caffeine contents found in the samples from Amazonas and Bahia are significantly different with 95 % of significance level. However, the HPLC-UV results showed that the caffeine content from both regions is statistically equal, with the same significance level. The higher dispersion of the data obtained by HPLC-UV induces a tendency of similarity among the values. On the other hand, the results obtained by ^{1}H *q*NMR has lower deviation given a small discrepancy among the samples highlighting the differences of caffeine content. Therefore, the caffeine content of samples from different geographic origin is statistically different corroborating the chemometrics.

Conclusions

Chemometric analysis discriminated the samples according to the geographic origin. Guarana seeds from Bahia state presented the highest phenolic content (catechin and epi-catechin), and guarana seeds from Amazon state presented the highest content of fatty acids and cyanolipids. Based on the *q*NMR results (less dispersed data), the content of caffeine from Bahia and Amazonas are significantly different. Therefore, as the methodologies were performed with biological replicates and *q*NMR corroborate the chemometrics, this study emphasizes the remarkable advantage of ^{1}H *q*NMR for food analysis based on the identification and quantification of compounds in complex mixtures without the use of standards.

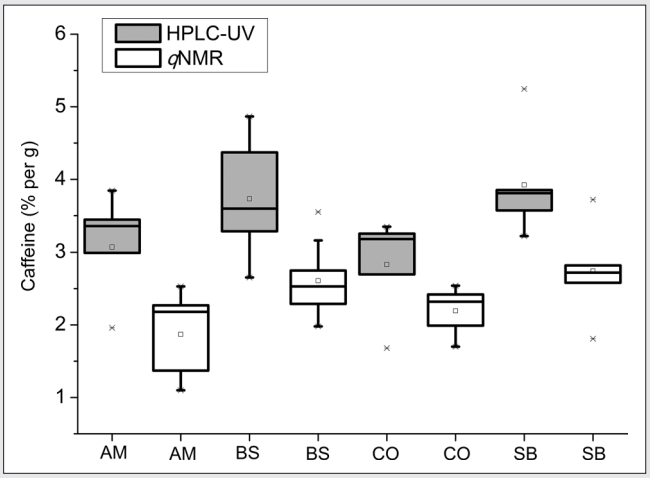

Figure 5. Box and whisker graph for quantification of caffeine through HPLC and ***q***NMR. Legend: AM – Amazon (Maués); BS – Baixo Sul region of Bahia; CO – Commercial from Bahia; SB – South of Bahia.

References

1. A. Basile, L. Ferrara, M.D. Pezzo, G. Mele, S. Sorbo, P. Bassi and D. Montesano, "Antibacterial and Antioxidant Activities of Ethanol Extract from Paullinia Cupana Mart", *J Ethnopharmacol*, **102**, 32 (2005). doi: http://dx.doi.org/10.1016/j.jep.2005.05.038
2. F.C. Schimpl, J.F. da Silva, J.F.C. Gonçalves and P. Mazzafera, "Guarana: Revisiting a Highly Caffeinated Plant from the Amazon", *J Ethnopharmacol*, **150,** 14 (2013). doi: http://dx.doi.org/10.1016/j.jep.2013.08.023
3. L.L.A. Veiga, H. Amorim, J. Moraes, M.C. Silva, R.S.L. Raices and S.L. Quiterio, "Quantification of Polycyclic Aromatic Hydrocarbons in Toasted Guaraná (Paullinia Cupana) by High-Performance Liquid Chromatography with a Fluorescence Detector", *Food Chem*, **152,** 612 (2014). doi: http://dx.doi.org/10.1016/j.foodchem.2013.11.154
4. A.F.M. Claudio, A.M. Ferreira, M.G. Freire and J.A.P. Coutinho, "Enhanced Extraction of Caffeine from Guarana Seeds Using Aqueous Solutions of Ionic Liquids", *Green Chem*, **15,** 2002 (2013). doi: http://dx.doi.org/10.1039/C3GC40437D
5. L.S. Bittencourt, D.C. Machado, M.M. Machado, G.F.F. Dos Santos, T.D. Algarve, D.R. Marinowic, E.E. Ribeiro, F.A.A. Soares, F. Barbisan, M.L. Athayde and I.B.M. Cruz, "The Protective Effects of Guaraná Extract (Paullinia Cupana) on Fibroblast Nih-3t3 Cells Exposed to Sodium Nitroprusside", *Food Chem Toxicol*, **53,** 119 (2013). doi: http://dx.doi.org/10.1016/j.fct.2012.11.041
6. R.S. Pagliarussi, L.A.P. Freitas and J.K. Bastos, "A Quantitative Method for the Analysis of Xanthine Alkaloids in Paullinia Cupana (Guarana) by

Capillary Column Gas Chromatography", *J Sep Sci*, **25,** 371 (2002). doi: http://dx.doi.org/10.1002/1615-9314(20020401)25:5/6%3C371::AID-JSSC371%3E3.0.CO;2-9

7. J. Sperková and M. Suchánek, "Multivariate Classification of Wines from Different Bohemian Regions (Czech Republic)", *Food Chem*, **93,** 659 (2005). doi: http://dx.doi.org/10.1016/j.foodchem.2004.10.044
8. P. Cheajesadagul, C. Arnaudguilhem, J. Shiowatana, A. Siripinyanond and J. Szpunar, "Discrimination of Geographical Origin of Rice Based on Multi-Element Fingerprinting by High Resolution Inductively Coupled Plasma Mass Spectrometry", *Food Chem*, **141,** 3504 (2013). doi: http://dx.doi.org/10.1016/j.foodchem.2013.06.060
9. G. Kouvoutsakis, C. Mitsi, P.A. Tarantilis, M.G. Polissiou and C.S. Pappas, "Geographical Differentiation of Dried Lentil Seed (Lens Culinaris) Samples Using Diffuse Reflectance Fourier Transform Infrared Spectroscopy (Drifts) and Discriminant Analysis", *Food Chem*, **145,** 1011 (2014). doi: http://dx.doi.org/10.1016/j.foodchem.2013.09.041
10. A. Caligiani, L. Palla, D. Acquotti, A. Marseglia and G. Palla, "Application of 1h Nmr for the Characterisation of Cocoa Beans of Different Geographical Origins and Fermentation Levels", *Food Chem*, **157,** 94 (2014). doi: http://dx.doi.org/10.1016/j.foodchem.2014.01.116
11. R. Consonni, L.R. Cagliani and C. Cogliati, "Nmr Based Geographical Characterization of Roasted Coffee", *Talanta*, **88,** 420 (2012). doi: http://dx.doi.org/10.1016/j.talanta.2011.11.010
12. H. Winning, F.H. Larsen, R. Bro and S.B. Engelsen, "Quantitative Analysis of Nmr Spectra with Chemometrics", *J Magn Reson*, **190,** 26 (2008). doi: http://dx.doi.org/10.1016/j.jmr.2007.10.005
13. L.M.A. Silva, E.G. Alves Filho, R. Choze, L.M. Lião and G.B. Alcantara, "1h Hrmas Nmr Spectroscopy and Chemometrics for Evaluation of Metabolic Changes in Citrus Sinensis Caused by Xanthomonas Axonopodis Pv. Citri", *J Braz Chem Soc*, **23,** 1054 (2012). doi: http://dx.doi.org/10.1590/S0103-50532012000600009
14. E.G. Alves Filho, F.D.L. Almeida, R.S. Cavalcante, E.S. de Brito, P.J. Cullen, J.M. Frias, P. Bourke, F.A.N. Fernandes and S. Rodrigues, "1h Nmr Spectroscopy and Chemometrics Evaluation of Non-Thermal Processing of Orange Juice", *Food Chem*, **204,** 102 (2016). doi: http://dx.doi.org/http://dx.doi.org/10.1016/j.foodchem.2016.02.121
15. S.n.M.M. Oliveira, "Guaraná Em Pó, Xarope E Extrato. InformaçÕEs TéCnicas Sobre a FabricaçãO De Produtos De Guaraná, Assim Como Processos, MáQuinas E Equipamentos Empregados.", *Serviço Brasileiro de Respostas Técnicas*, (2006). doi:
16. D.S. Wishart *et al.*, "Hmdb: A Knowledgebase for the Human Metabolome", *Nucleic Acids Res*, **37,** D603 (2009). doi: http://dx.doi.org/10.1093/nar/gkn810
17. R.C.A. Lago, M.d.L.P.S.C. Simoni and Â.d.C. Pinto, "On the Occurrence of Cyanolipids in Paullinia Carpopodea Cambess and P. Cupana Kunth Seed Oils", *Acta Amazonica*, **30,** 101 (2000). doi: http://dx.doi.org/10.1590/1809-43922000301105
18. A.L. Davis, Y. Cai, A.P. Davies and J.R. Lewis, "1h and 13c Nmr Assignments of Some Green Tea Polyphenols", *Magn Reson Chem*, **34,** 887 (1996). doi: http://dx.doi.org/10.1002/(SICI)1097-458X(199611)34:11%3C887::AID-OMR995%3E3.0.CO;2-U
19. J. Sitkowski, L. Stefaniak, L. Nicol, M.L. Martin, G.J. Martin and G.A. Webb, "Complete Assignments of the 1h, 13c and 15n Nmr Spectra of Caffeine", *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **51,** 839 (1995). doi: http://dx.doi.org/10.1016/0584-8539(94)00192-E
20. A.D. Meinhart, C.S. Bizzotto, C.A. Ballus, M.A. Prado, R.E. Bruns, J.T. Filho and H.T. Godoy, "Optimisation of a Ce Method for Caffeine Analysis in Decaffeinated Coffee", *Food Chem*, **120,** 1155 (2010). doi: http://dx.doi.org/10.1016/j.foodchem.2009.11.048
21. I.-I.N.D. METEOROLOGIA, "Estações Automáticas - Gráficos," vol. 2016,