2012

Multivariate analyses of UV-Vis absorption spectral data from cachaca wood extracts: a model to classify aged Brazilian cachacas according to the wood species used

ANALYTICAL METHODS, CAMBRIDGE, v. 4, n. 3, supl. 2, Part 3, pp. 642-646, MAR, 2012
http://www.producao.usp.br/handle/BDPI/42600

Analytical Methods

Cite this: *Anal. Methods*, 2012, **4**, 642

www.rsc.org/methods

PAPER

Multivariate analyses of UV-Vis absorption spectral data from cachaça wood extracts: a model to classify aged Brazilian cachaças according to the wood species used†

Alexandre Ataide da Silva,[a] Denis De Keukeleire,[b] Daniel Rodrigues Cardoso[a] and Douglas Wagner Franco*[a]

Received 12th October 2011, Accepted 16th January 2012
DOI: 10.1039/c2ay05670d

Multivariate analyses of UV-Vis spectral data from cachaça wood extracts provide a simple and robust model to classify aged Brazilian cachaças according to the wood species used in the maturation barrels. The model is based on inspection of 93 extracts of oak and different Brazilian wood species by a non-aged cachaça used as an extraction solvent. Application of PCA (Principal Components Analysis) and HCA (Hierarchical Cluster Analysis) leads to identification of 6 clusters of cachaça wood extracts (amburana, amendoim, bálsamo, castanheira, jatobá, and oak). LDA (Linear Discriminant Analysis) affords classification of 10 different wood species used in the cachaça extracts (amburana, amendoim, bálsamo, cabreúva-parda, canela-sassáfras, castanheira, jatobá, jequitibá-rosa, louro-canela, and oak) with an accuracy ranging from 80% (amendoim and castanheira) to 100% (bálsamo and jequitibá-rosa). The methodology provides a low-cost alternative to methods based on liquid chromatography and mass spectrometry to classify cachaças aged in barrels that are composed of different wood species.

Introduction

The flavor characteristics of alcoholic beverages such as rum, whisky, brandy, and cachaça are dependent to a great extent on maturation, which is commonly achieved in wooden barrels to improve the product's sensorial attributes.[1–4] Typically, a fresh distillate is colorless and its taste, aroma, and mouth-feel characteristics will benefit from improvements associated with aging in wooden barrels in order to improve the consumers' acceptance. During aging, various constituents are extracted from the wood into the beverage and are further altered by chemical reactions including oxidations and esterification.[5–7] The qualitative and quantitative composition of the extracted fraction is determined by a number of variables such as the specific nature of the wood species, the geographic origin, the previous treatments of the wooden barrel, the duration of the aging process, and the frequency of use of the barrel.[8–15] Barrels are chosen depending on particular preferences by the cachaça producers or by particular flavors that are derived from the wood species. Although new barrels can be used, a general trend is to involve barrels that had previously been used for maturation of alcoholic beverages. Brazilian cachaças are mainly aged in oak barrels originating from the production of whiskies.[16,17]

However, there exists a great number of Brazilian wood species that could be considered as alternatives to oak. From this perspective, knowledge of the chemical nature of the wood constituents extracted into aged cachaças and development of a suitable model to certify these cachaças and the woods used during maturation is of great interest.[18–21] Spectral fingerprint methodologies that have been reported rely on sophisticated instrumentation (mass and nuclear magnetic resonance spectrometers) for collection and interpretation of data.[19,22–25] There is undoubtedly a need for a simple, robust, and cost-effective methodology based on multivariate analyses of readily accessible spectral data such as delivered by UV-Vis absorption spectroscopy. The purpose of the present study was to develop such a model.

Materials and methods

Samples

Certified samples of non-aged cachaças (without added sugars) kindly provided by Industrias Müller de bebidas LTDA (Pirassununga, Sao Paulo, Brazil) were used as a solvent for the preparation of wood extracts. Commercial cachaças, certified by the producer and aged for 1 year, served to evaluate the model developed in this work. The Laboratório de Madeiras e Estrutura da Universidade de São Paulo (São Carlos, São Paulo,

[a]*Departamento de Química e Física Molecular, Instituto de Química de São Carlos, Universidade de São Paulo (USP), Avenida Trabalhador São-carlense 400, CP 780, CEP 13560-970 São Carlos, SP, Brazil. E-mail: douglas@iqsc.usp.br*
[b]*Ghent University, Faculty of Pharmaceutical Sciences, Harelbekestraat 72, 9000 Ghent, Belgium*

† Electronic supplementary information (ESI) available: HCA dendrogram plot of cachaça wood extracts using normalized first derivatives of spectral absorbance as variables (Fig. S1). Similarity measurement: Euclidean distance; data processing: mean-centered; and clustering technique: incremental. See DOI: 10.1039/c2ay05670d

Brazil) and Tanoaria Barro, Tonéis e Cia, and Tanoaria Paulista provided certified air-dried samples of the following Brazilian wood species: amburana (*Amburana cearensis*), amendoim (*Pterogyne nitens*), bálsamo (*Myroxylon balsamum*), cabreúva-parda (*Myrocarpus frondosus*), canela-sassafrás (*Ocotea odorifera*), castanheira (*Bertholletia excelsa*), jatobá (*Hymenaea courbaril*), jequitibá (jequitibá-rosa: *Cariniana legalis*; jequitibá-branco: *Cariniana estrellensis*), and louro-canela (*Aniba parviflora*). The certified air-dried or kiln-dried oak samples were kindly provided by Prof. Dr John Piggot (Department of Bioscience at Strathclyde University, Glasgow, Scotland). Extracts were prepared following the general procedure described previously:[19,26] dry wood sawdust (100 ± 20 mesh) of the certified wood species was extracted with non-aged cachaça (47% v/v) for 26 days on an automatic shaking table at room temperature (25 ± 2) °C. Samples were protected from light during the extraction procedure and the sample/solvent ratio was 0.010 g of wood sawdust per millilitre of spirit. The wood sawdust residues were then filtered-off using a Whatman quantitative paper filter and the remaining extracts were stored in amber flasks at room temperature (25 ± 3) °C.

UV-Vis spectroscopy

The absorption spectra were collected on a Hewlett Packard 8152A spectrophotometer (Pickering, Ontario, Canada) using 1 cm × 1 cm spectrophotometer quartz cells from Wilmad (Vineland, New Jersey, USA). Samples were diluted in ethanol/water (47% v/v) to adjust the absorbance at 220 nm to the unit value. The spectra were recorded between 220 nm and 400 nm in consecutive steps of 2 nm each. The reproducibility (for absorption values and wavelength shifts) was checked at least twice for each sample.

Multivariate analyses

PCA (Principal Components Analysis) correlation matrices were built up using the complete dataset for which rows represent samples (75) and columns represent the absorbances at a given wavelength (90). The variables (columns) were mean centered and the first derivative was applied prior to the statistical treatment. The mean centered values were calculated using the mean of each variable according to $X_j = (1/n)\sum_{i=j}^{n} X_i$, where the average X_j is the average of data contained in the column (variable). Then, this value was subtracted from the original data and the first derivative applied. This procedure was carried out on statistical software. The analytical data of each wood group were analyzed by PCA (Principal Components Analysis), HCA (Hierarchical Cluster Analysis), and LDA (Linear Discriminant Analysis) methods. The following parameters were used for HCA analysis: metric Euclidean distance, similarity plot, and complete linkage. LDA was performed using the leave-one-out cross-validation approach and the classification was carried out by confronting the unknown sample against two groups, one containing a known wood species and the other comprising all other wood species. Multivariate analyses were carried out using the Minitab R14 software (Minitab, Coventry, UK) for analysis of HCA and LDA and the Unscrambler 9.2 software (Camo Software, Oslo, Norway) for analysis of PCA.

Results and discussion

The UV-Vis absorption spectra collected from 220 nm to 400 nm for 93 samples of the cachaça wood extracts and 51 commercial cachaça samples were analyzed by PCA, HCA, and LDA in order to assess the potential for differentiation and classification of the wood species used for aging purposes. Typically, the UV-Vis absorption spectra of the cachaça wood extracts are characterized by continuous absorption throughout the UV region with some structure-less broad absorption bands between 250 nm and 300 nm, while absorption in the visible (near or above 400 nm) is negligible. Fig. 1A shows some typical spectra of representative cachaça wood extracts and Fig. 1B displays the first derivatives of the respective spectra.

The absorption features are characteristic of substituted aromatic compounds of the phenylpropanoid type that are well-known constituents of the wood lignin superstructure. In line with the typical aromatic absorption bands, intense absorptions are observed in the far-UV region (around 220 nm and lower).

Principal components analysis (PCA) of the UV-Vis dataset shows that principal components PC1, PC2, and PC3 account for 92% of the total variance (PC1: 63%, PC2: 19%, and PC3: 10%). The score plots PC1 *vs.* PC2 (Fig. 2) and PC1 *vs.* PC3 (Fig. 3)

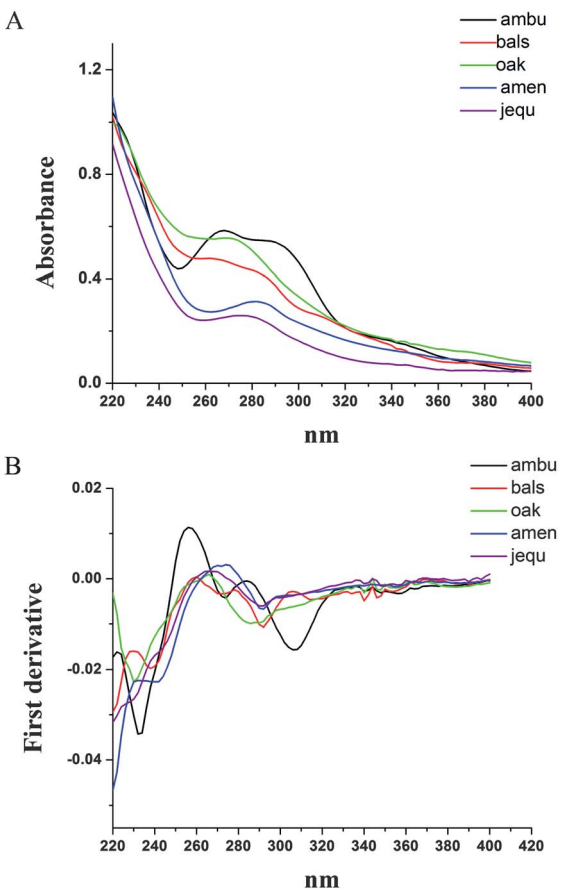

Fig. 1 (A) Typical UV-Vis absorption spectra of some representative cachaça wood extracts. (B) First derivatives of the respective UV-Vis absorption spectra. Ambu = amburana; bals = bálsamo; oak = oak; amen = amendoim; and jequ = jequitibá.

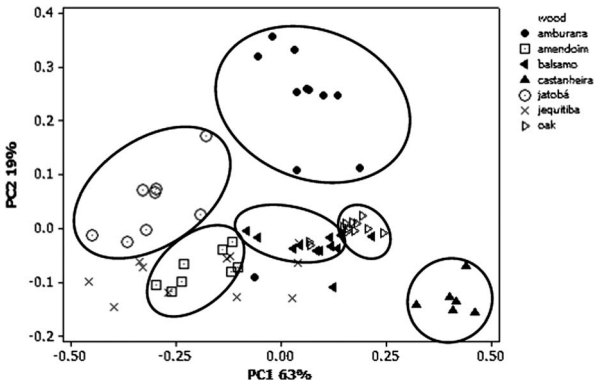

Fig. 2 Score plot for PC1 (63%) *vs.* PC2 (19%) of cachaça wood extracts.

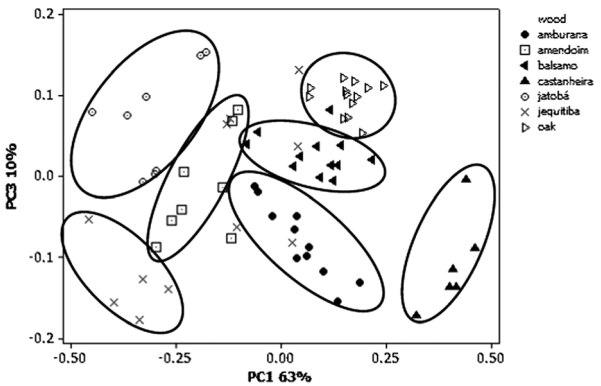

Fig. 3 Score plot for PC1 (63%) *vs.* PC3 (10%) of cachaça wood extracts.

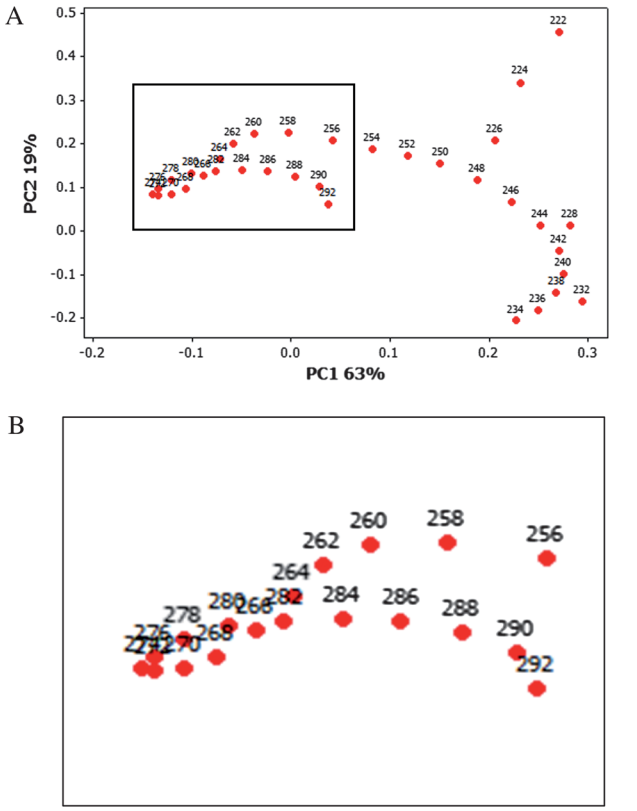

Fig. 4 (A) Loading plot of PC1 *vs.* PC2 based on the UV-Vis absorption spectral data (numbers refer to the wavelengths used in the multivariate analysis). (B) Zoom of the region marked with a square in (A).

allowed us to distinguish all cachaça wood species except jequitibá.

The 3-D scatter plot provides interesting aspects of the clustering of samples. Three well-defined large clusters were identified corresponding to amburana, castanheira, and (amendoim + bálsamo + jatobá + jequitibá + oak), respectively. For this last cluster, the first two PCs account for 87% of the total variance (PC1 = 75% and PC2 = 12%). Samples of jequitibá formed a scattered group spread along PC1, although all samples were clustered and isolated from the other samples. This may be due to the fact that the jequitibá samples belong to the same genus (*Cariniana*), but to different species, *i.e.* 7 samples are of jequitibá-branco (*Cariniana estrellensis*) and 4 samples are of jequitibá-rosa (*Cariniana legalis*).

Examination of the PC loadings (Fig. 4) indicates that PC1 (63% of the original information) is strongly correlated to positive values of amburana, bálsamo, castanheira, and oak samples and to negative values of amendoim and jatobá. A further analysis of the score and loading plots shows that a wavelength around 230 nm is most relevant (PC1, positive values) for the separation of amburana, bálsamo, castanheira, and oak extracts from the other groups, and that the region around 280 nm is most relevant (PC1, negative values) to define the amendoim, jequitibá, and jatobá clusters.

Exploration of the dataset by HCA (Hierarchical Cluster Analysis) gave similar results as observed for PCA plots. Bálsamo and oak samples showed a tendency to cluster closely to each other. At the similarity level of 20%, two clusters appear, each one containing a mixture of bálsamo and oak samples. The HCA plots, as observed by PCA, show a tendency to group the cachaça extracts according to the respective wood species. Best results were obtained for jatobá, castanheira, and amburana. Another important observation concerns the HCA analysis of the jequitibá extracts. In the PCA scatter plot, the distribution is spread out, whereas, in the HCA plot, a tendency to cluster at the similarity level of 66% is clearly evident (Fig. S1, ESI†).

Furthermore, LDA (Linear Discriminant Analysis) was applied to the data matrix in two steps. First, considering a matrix composed of cachaça wood extracts only (93 lines × 90 columns), 67 samples were used to define the calibration model and the remaining 26 samples were used for the model validation. The second step was performed using all cachaça wood extracts in the calibration model (93 samples) and the commercial cachaças (35 samples) for validation purposes. The wavelengths (descriptors) considered in LDA were those that provided the highest score values in the PCA loading plot (Fig. 4). It is interesting to point out that a previous comparison among the output results from PCA and LDA was carried out and converging results concerning the relevant variables to be used in the discriminant analysis were observed.

Application of the LDA model to 21 unknown samples (3 oak, 2 amendoim, 2 amburana, 3 bálsamo, 2 cabreúva-parda, 1 canela-sassafrás, 2 castanheira, 2 jatobá, 2 jequitibá, and 2 louro-canela) achieved a correct classification for ≥74% of the wood species studied (Table 1). The following aged cachaças were correctly assigned according to the wood species used (100%):

Table 1 Classification and cross-validation of the model derived from LDA (Linear Discriminant Analysis) for oak and Brazilian wood extracted by cachaça. Wavelengths from 220 nm to 310 nm were selected in consecutive steps of 6 nm each to minimize correlation errors between descriptors

Wood	Total number of extracts	Calibration model—number of samples	Calibration model—correct classification	Cross-validation model—number of samples	Cross-validation model—correct classification
Oak	13	10	95%	3	84%
Amburana	10	8	98%	2	98%
Amendoim	8	6	82%	2	80%
Bálsamo	12	9	91%	3	87%
Jatobá	9	7	89%	2	89%
Cabreúva-parda	8	6	78%	2	79%
Canela-sassafrás	5	3	82%	1	91%
Castanheira	12	10	100%	2	99%
Jequitibá	11	9	84%	2	81%
Louro-canela	5	3	73%	1	74%

Table 2 Cross-validation for classification of Brazilian cachaças aged in different types of wood using the model derived from LDA (Linear Discriminant Analysis)

Aged cachaça (type of wood)	Number of cachaças	Correct classification
Oak	9	88%
Amburana	6	83%
Amendoim	6	80%
Bálsamo	10	100%
Cabreúva-parda	2	100%
Canela-sassafrás	3	100%
Castanheira	5	80%
Jequitibá	7	100%
Louro-canela	1	100%
Jatobá	1	100%

básamo, cabreúva-parda, canela-sassafrás, jatobá, jequitibá, and louro-canela, followed by oak, amburana, amendoim and castanheira (Table 2). Fine tuning with respective to regional wood characteristics is currently under investigation in our laboratory.

The PCA and HCA analyses have been carried out using 93 samples from 7 different types of cachaça wood extracts. The inclusion of data on cabreúva-parda, canela sassafras, and louro-canela into the PCA and HCA plots scrambled the clustering and, therefore, has been omitted. On the other hand, the LDA model included all 10 different cachaça wood extracts proving its superior power for our purpose. The results obtained from LDA application to spectral data for 10 different types of cachaça wood extracts indicate that UV-Vis absorption spectra profiles are potentially quite useful as a routine technique to classify the type of wood used in the beverage maturation process.

Conclusions

All together the results of the analytical data obtained from the UV-Vis absorption spectra, treated by multivariate statistical analyses, proved to be a simple, robust, and cost-effective tool for routine classification of the different wood species used on the Brazilian cachaças aging. Furthermore, it holds some potential as a possible forensic tool for wood identification which could be applied to control the wood marked of endangered species. Our findings could be extended to other spirits and to a wider variety of wood species.

Acknowledgements

The authors thank Prof. Dr John Piggot from the Department of Bioscience at Strathclyde University (Glasgow, Scotland) for kindly providing the certified oak samples and Industrias Müller de Bebidas LTDA for offering the non-aged spirits. Prof. Dr Francisco Antonio Rocco Lahr (Laboratório de Madeiras e Estruturas de Madeiras, USP-São Carlos, São Paulo, Brazil) is thanked for help in preparing the wood extracts, and Prof. Dr Edenir Rodrigues Pereira Filho, Universidade Federal De São Carlos, for helpful discussions regarding the statistical software application. Fellowships and financial support from FAPESP, CAPES, and CNPq are also gratefully acknowledged.

Notes and references

1 J. J. Quesada, M. Villalon, G. H. Lopez and M. C. Lopez, *J. Agric. Food Chem.*, 1996, **44**, 1378–1381.
2 J. R. Mosedale and J.-L. Puech, *Trends Food Sci. Technol.*, 1998, **9**, 95–101.
3 J. R. Mosedale and A. Ford, *J. Sci. Food Agric.*, 1996, **70**, 273–287.
4 S. Canas, M. C. Leandro, M. I. Spranger and A. P. Belchior, *J. Agric. Food Chem.*, 1999, **47**, 5023–5030.
5 D. Matejicek, O. Mikes, B. Klejdus, D. Sterbova and V. Kuban, *Food Chem.*, 2005, **90**, 791–800.
6 J. M. Conner, *J. Inst. Brew.*, 1999, **105**, 287–291.
7 J. P. Dos Anjos, M. G. Cardoso, A. A. Saczk, H. S. Dórea, W. D. Santiago, A. M. R. Machado, L. M. Zacaroni and D. L. Nelson, *J. Braz. Chem. Soc.*, 2011, **22**, 7, 1307–1314.
8 J. O. S. Campos, F. W. B. Aquino, R. F. Nascimento, J. G. M. Costa, D. De Keukeleire and A. R. S. Casimiro, *J. Food Compos. Anal.*, 2004, **17**, 179–185.
9 V. L. Singleton, *Am. J. Enol. Vitic.*, 1995, **46**, 1, 98–113.
10 R. R. Madrera, D. B. Gomis and J. J. M. Alonso, *J. Agric. Food Chem.*, 2003, **51**, 7969–7973.
11 A. Prida and J. L. Puech, *J. Agric. Food Chem.*, 2006, **54**, 8115–8126.
12 J. R. Mosedale, *Forestry*, 1995, **68**, 3–15.
13 L. Odello, G. P. Braceschi, F. R. F. Seixas, A. A. Silva, C. A. Galinaro and D. W. Franco, *Quim. Nova*, 2009, **32**, 1839–1844.
14 M. Sanz, F. Brígida Simón, E. Esteruelas, A. M. Muñoz, E. Cadahía, T. Hernández, I. Estrella and E. Ernani Pinto, *J. Agric. Food Chem.*, 2011, **59**, 3135–3145.
15 J. Michel, M. Jourdes, M. A. Silva, T. Giordanengo, N. Mourey and P. Teissedre, *J. Agric. Food Chem.*, 2011, **59**, 5677–5683.
16 J. B. Faria, H. M. A. B. Cardello, M. Boscolo, W. D. Isique, L. Odello and D. W. Franco, *Eur. Food Res. Technol.*, 2003, **218**, 83–87.
17 E. Koussissi, V. G. Dourtoglou, G. Ageloussis, Y. Paraskevopoulos, T. Dourtoglou, A. Paterson and A. Chatzilazarou, *Food Chem.*, 2009, **114**, 1503–1509.

18 K. Nishimura and R. Matsuyama, in *The Science and Technology of Whiskies*, ed. J. R. Piggott, R. Sharp and R. E. B. Duncan, Longman, Essex, England, 1989, pp. 235–263.
19 A. A. Da Silva, E. S. P. Nascimento, D. R. Cardoso and D. W. Franco, *J. Sep. Sci.*, 2009, **32**, 3681–3691.
20 P. P. Souza, H. G. L. Siebald, D. V. Augusti, W. B. Neto, V. M. Amorim, R. R. Catharino, M. N. Eberlin and R. Augusti, *J. Agric. Food Chem.*, 2007, **55**, 2094–2102.
21 C. D. Vicente, F. C. De Abreu, M. O. F. Goulart and J. N. De Vasconcelos, *Am. J. Food Technol.*, 2011, **6**, 631–646.
22 L. Viggiani, M. Antoneietta and C. Morelli, *J. Agric. Food Chem.*, 2008, **56**, 8273–8279.
23 R. R. Catharino, R. Haddad, L. G. Cabrini, I. B. S. Cunha, A. C. H. F. Sawaya and M. N. Eberlin, *Anal. Chem.*, 2005, **77**, 7429–7433.
24 J. S. Mattson, *Anal. Chem.*, 1971, **43**, 1872–1873.
25 D. L. Luthria, S. Mukhopadhyay, R. J. Robbins, J. W. Finley, G. S. Banulos and J. M. Harnly, *J. Agric. Food Chem.*, 2008, **56**, 5457–5462.
26 D. R. Cardoso, A. M. Frederiksen, A. A. Da Silva, D. W. Franco and L. H. Skibsted, *Eur. Food Res. Technol.*, 2008, **227**, 1109–1117.

