JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Chemical Profile of Rums as a Function of their Origin. The Use of Chemometric Techniques for their Identification

OLÍVIA MOREIRA SAMPAIO, RONI VICENTE RECHE, AND DOUGLAS WAGNER FRANCO*

Departamento de Química e Física Molecular, IQSC, Universidade de São Paulo, Avenida Trabalhador São Carlense 400, CP 780, CEP 13560-970, São Carlos, SP, Brazil

To identify chemical descriptors to distinguish Cuban from non-Cuban rums, analyses of 44 samples of rum from 15 different countries are described. To provide the chemical descriptors, analyses of the the mineral fraction, phenolic compounds, caramel, alcohols, acetic acid, ethyl acetate, ketones, and aldehydes were carried out. The analytical data were treated through the following chemometric methods: principal component analysis (PCA), partial least square-discriminate analysis (PLS-DA), and linear discriminate analysis (LDA). These analyses indicated 23 analytes as relevant chemical descriptors for the separation of rums into two distinct groups. The possibility of clustering the rum samples investigated through PCA analysis led to an accumulative percentage of 70.4% in the first three principal components, and isoamyl alcohol, *n*-propyl alcohol, copper, iron, 2-furfuraldehyde (furfuraldehyde), phenylmethanal (benzaldehyde), epicatechin, and vanillin were used as chemical descriptors. By applying the PLS-DA technique to the whole set of analytical data, the following analytes have been selected as descriptors: acetone, *sec*-butyl alcohol, isobutyl alcohol, ethyl acetate, methanol, isoamyl alcohol, magnesium, sodium, lead, iron, manganese, copper, zinc, 4-hydroxy-3,5-dimethoxybenzaldehyde (syringaldehyde), methaldehyde (formaldehyde), 5-hydroxymethyl-2-furfuraldehyde (5-HMF), acetaldehyde, 2-furfuraldehyde, 2-butenal (crotonaldehyde), *n*-pentanal (valeraldehyde), iso-pentanal (isovaleraldehyde), benzaldehyde, 2,3-butanodione monoxime, acetylacetone, epicatechin, and vanillin. By applying the LDA technique, a model was developed, and the following analytes were selected as descriptors: ethyl acetate, *sec*-butyl alcohol, *n*-propyl alcohol, *n*-butyl alcohol, isoamyl alcohol, isobutyl alcohol, caramel, catechin, vanillin, epicatechin, manganese, acetaldehyde, 4-hydroxy-3-methoxybenzoic acid, 2-butenal, 4-hydroxy-3,5-dimethoxybenzoic acid, cyclopentanone, acetone, lead, zinc, calcium, barium, strontium, and sodium. This model allowed the discrimination of Cuban rums from the others with 88.2% accuracy.

KEYWORDS: Rum; metals; volatile compounds; chemical profiles; multivariate analysis

INTRODUCTION

Rum is a fairly tasteless and neutral spirit derived from the fermentation of sugar molasses and sugar cane syrup (*1, 2*). Alternatively, rum can also be produced from the fermentation of preconcentrated sugar cane juice. This preconcentration is carried out by heating (*2*). Once the alcohol is obtained from the fermentation and distillation processes, it undergoes further processing, such as percolation through carbon filters, aging in oak barrels, and blending, which give the rum particular sensory characteristics (*1*).

Molasses, the subproduct of sugar cane, has a different chemical composition from sugar cane juice which, upon fermentation, generates a number of compounds that will impact the chemical composition of the rum, such as aliphatic and aromatic esters, aldehydes, alcohols, furan derivatives, nucleic acids, alcohols, amino acids, and other organic acids (*2, 3*). Nicol (*2*) reports that fresh blackstrap molasses with a low viscosity, high total sugars, nitrogen, and phosphorus, and a low ash and gum content helped in the production of rum with the desired odors and tastes.

Rum and cachaça from Brazil (its sister spirit), both made from sugar cane, account for a high percentage of the distilled alcoholic beverages consumption in the world (*2*).

Caribbeans from English-speaking areas (Barbados) and French-speaking areas (Martinique), Mexicans, Venezuelans, Cubans, and Puerto Ricans produce light rums (low congeners content) with a light body (dark), and Jamaicans and Guyaneses produce navy rums (heavy). Carta Blanca and Carta Oro are

* To whom correspondence should be addressed. Telephone/Fax: + 55 16 3373 9976. E-mail: douglas@iqsc.usp.br.

10.1021/jf0726841 CCC: $40.75
Published on Web 02/15/2008

the expressions used in some countries to label light, white, or caramel-colored rums.

The type of rum produced largely depends on the type of raw materials, their treatment (*2, 3*), the type of yeast used in the fermentation process, and the equipment used for the distillation. Rum can be distilled in pot stills (batch) and columns (continuous distillation systems).

Today, the distillation of Cuban rum is mostly done in stainless steel columns, whereas the distillation of almost all other rums is done in pot stills (*4*). Column distillation is a continuous process that involves many assembled theoretical plates, and the copper pot still (or alembic) proces is known as a batch distillation which corresponds to one theoretical plate (*5–7*). Most Cuban rums are aged, and the aging process involves keeping the spirit in white American oak barrels, purchased from Canada and previously used for the aging of whiskey (*4*).

It is a common practice in pot still distillation to collect the distillate in three different fractions: head, heart, and tail, at 70, 40, and 15% alcoholic content, respectively. Only the heart fraction is commercialized.

Rums produced in pot stills are more robust and heavier than those produced in columns. Pot still distillates undergo prolonged maturation as malt whiskies. Conversely, column still products are light rums and neutral spirits, such as gin and vodka (*2*). Pot still rums are mainly produced in English- and French-speaking areas of the West Indies and the Caribbean Islands (Antilles). Continuous distillation is largely used for Cuban and Puerto Rican rums (*3*).

The presence of volatile components, such as alcohols, ethyl acetate, acetic acid, aldehydes, ketones, polyphenols, and nonvolatile compounds, such as metals ions originating from the raw materials and the fermentation, distillation, and aging processes, is essential to define the beverages composition and, therefore, provide elements for their distinction (*1–3, 8*).

For example, the aldehydes found in rum are formed during the fermentation and distillation in copper apparatus, the higher alcohols are formed during the fermentation procedure and dragged during the distillation, and the phenolics compounds are extracted from the barrels during the maturation of the rum (*3*).

Chemometric techniques such as principal component analysis (PCA), partial least square-discriminate analysis (PLS-DA), and linear discriminate analysis (LDA) to the analytical data to certify the geographic origin of the beverages were extensively used (*8–18*). Chemometry was applied by Latorre et al. (*9*) in the study of wines from northwestern Spain, and lithium, rubidium, sodium, potassium, manganese, iron, and calcium were used as descriptors.

Today, international trade is very intense, and because of health safety and economic policies, there is a worldwide growing concern regarding product typification, authenticity, origin, and falsification. Based on the chemometric treatment of collected analytical data for metal ions and organic compounds, the typification of cachaça, together with suggestions to improve the quality control, is being carried out at our laboratory. An example of the distinction between rum and cachaça by applying PCA, HCA, and PLS-DA, among other techniques, has been proposed (*8*).

In countries other than the Latin American countries, rum is by far more consumed than cachaça. In contrast to cachaça, which is a typical product of Brazil, rum is produced in many countries; therefore, there is an interest in developing a method to certify the authenticity and the origin of the beverage. The first efforts toward this have been reported by Herranz et al. (*19*) who, by using major volatile components, esters and carboxylic acids, proposed through multivariated statistical methods the differentiation of a genuine rum of a well-known and expensive brand from the others. However, the number of samples were small, and no attempts to correlate the rums with their origins were described. Herein, we have described our efforts to discriminate rums produced in Cuba from the others on the basis of the quantitative mineral and organic profiles and the analyses of these data through chemometric methods.

MATERIALS AND METHODS

Samples. A total of 44 certified rum samples were supplied by ABRABE (Associação Brasileira de Bebidas) and ICIDCA (Instituto Cubano de Investigaciones de los Derivados de la Caña de Azúcar).

The rum samples included Havana Club Silver Dry (Cuba), Havana Club Añejo 3 años (Cuba), Havana Club Añejo 7 años (Cuba), Havana Club Añejo Reserva (Cuba), XK Solera (Mexico), Bacardi Premium Black (Brazil), Bacardi Carta de Oro (Brazil), Bacardi Carta Blanca (Brazil), Montilla Carta Oro (Brazil), Montilla Carta Blanca (Brazil), Appleton White (Jamaica), Appleton Estate (Jamaica), Mount Gay (Barbados), Negrita dry and light (France), Negrita aged 8 years (France), El Dorado aged 12 years (Guyana), Captain Morgan Dark (Canada), Casino (Hungria), Soccaron white (France), Jamaica (Brazil), Havana Club Añejo Blanco (Cuba), Caribbean Club Añejo 7 años (Cuba), Bucanero Añejo 7 años (Cuba), Silver Cacique Premium (Venezuela), Siboney Añejo 7 años (Cuba), Cruzan aged 5 years (USA), Del Barrilito (Puerto Rican), Centenario Añejo Especial (Costa Rica), Vigia Gran Añejo (Cuba), Matusalem (Cuba), Varadero Añejo 7 años (Cuba), Arecha Extra Añejo (Cuba), Legendário Carta Blanca (Cuba), Bucanero Añejo 7 años (Cuba), Bacardi Reserva (Puerto Rican), Abuelo Añejo (Panama), Mulata Añejo 7 años (Cuba), Conde de Cuba Añejo 7 años (Cuba), Cortez aged 3 years (Panama), Santiago de Cuba Añejo (Cuba), Edmundo Dante's (Cuba), Flor de Caña Black Label (Nicaragua), Montilla Carta Cristal (Brazil), and Bacardi Solera (Mexico). The Cuban rums will be called class 1 and the others class 2.

Materials. Ethanol, methanol, acetonitrile, acetaldehyde, *n*-propyl alcohol, isobutyl alcohol, *sec*-butyl alcohol, butanol, isoamyl alcohol, acetic acid, ethyl acetate, acetone, and hexanol (HPLC grade) were purchased from J.T. Baker and Tedia (Phillipsburg, NJ). Water was deionized by using a Milli-Q system (Millipore, Bedford, MA). All aldehydes and ketones standards (methaldehyde, butyraldehyde, 2-furfuraldehyde, 5-hydromethyl-2-furfuraldehyde, 2-butenal, propionaldehyde, 2-methylpropionaldehyde, phenylmethanal, *n*-pentanal, isopentanal, 2-propenal, 2,3-butanodione monoxime, cyclopentanone, acetylacetone, methyl-phenyl-ketone, methyl-isoamyl-ketone, and 2,4-dinitrophenylhydrazine (DNPH)) were purchased from Merck and Sigma-Aldrich (Milwaukee, WI) and used as purchased. Polyphenols (gallic acid, 4-hydroxy-3,5-dimethoxybenzoic acid, myricetin, ellagic acid, 4-hydroxy-3, 5-dimethoxybenzaldehyde, vanillin, 4-hydroxy-3-methoxybenzoic acid, epicatechin, catechin, scopoletin, coniferaldehyde, sinapaldehyde, *trans*-resveratrol, quercetin, eugenol, and coumarin) were purchased from Sigma-Aldrich (Milwaukee, WI). Metal standard solutions (copper, iron, zinc, magnesium, manganese, cadmium, lead, nickel, cobalt, chromium, calcium, barium, sodium, lithium, and strontium) were obtained by the dilution of 1000 $mg \cdot L^{-1}$ metal standards purchased from Carlo Erba (Milano, Italy). The acids HNO_3, H_2SO_4, and H_3PO_4 were purchased from Mallinckrodt (Xalostoc, Mexico).

Analytical Procedures. The selection of the compounds to be analyzed was based on their occurrences and quantitative profiles previously reported for other alcoholic beverages, including rum (*1–3*).

When chromatographic methods were used, the identification of the components was carried out by comparing the relative retention time of the standards, obtained from a standard 40% v/v ethanol–water solution, and through the addition of the desired standard into the sample. Quantitative analyses were performed by using both the internal standard and standard addition methods. For the spectroscopic analyses, atomic absorption spectroscopy (AAS) and inductively coupled plasma

optical emission spectroscopy (ICP OES), the external standard method was used for quantification. All the analyses were always performed in duplicate.

*Volatile Compounds (**20**).* The analyzed compounds were methanol, *n*-propyl alcohol, isobutyl alcohol, *sec*-butyl alcohol, *n*-butyl alcohol, isoamyl alcohol, acetic acid, ethyl acetate, 2-propenal, and acetone.

The rum samples were spiked with an internal standard (*n*-hexanol). Aliquots of 1.0 μL were injected into the gas chromatograph system (HP 5890) by using a flame ionization detector (FID) and an HP-FFAP column coated with esterified polyethylene glycol (50 m × 0.20 mm, 0.33 μm film thickness). The analyses were performed at a 1:20 split ratio. Hydrogen was used as the gas carrier (1.2 mL·min^{-1} flow rate). The temperatures of both the injector and the FID were set at 240 °C. The oven temperature program was set at 40 °C for 2 min, raised to 240 °C in 10 °C·min^{-1} steps, and kept isothermal for 4 min.

*Polyphenols (**21**).* Aliquots of 20.0 μL were injected in the liquid chromatograph system (Shimadzu LC-10ADVP) equipped with a SPD-M6A diode array detector (Shimadzu) and a RF-551 fluorescence detector (Shimadzu) coupled on-line. The separation was performed by using a Shimpack VP-ODS column (25.0 cm × 2.0 mm id × 2.5 μm) with the following mobile phases: solvent A, water/acetic acid (98:2 v/v); solvent B, methanol/water/acetic acid (70:28:2 v/v). The gradient profile at a 0.3 mL·min^{-1} flow rate was as follows: solvent B, 0.0% isocratic for 3 min, from 0.0 to 40% B in 22 min, from 40 to 60% B in 18 min, 60% isocratic for 12 min, from 60 to 80% B in 5 min, 80% isocratic for 5 min, and from 80 to 0% B in 2 min. Detection was performed using UV–vis and fluorescence detectors coupled on-line. Measurements with the UV–vis detector were carried out at 280 nm. For the fluorescence detector, the excitation was set at 280 nm and the emission at 313 nm.

*Aldehydes and Ketones (**22, 23**).* The analysis of the aldehydes and ketones by using liquid chromatography was performed after the derivation with DNPH. The DNPH standards were prepared by mixing solutions A and B. Solution A: 10 g of DNPH dissolved in 50 mL of H_2SO_4, 70 mL of H_2O, and 250 mL of ethanol. Solution B: 1 g or 1 mL of the desired aldehyde or ketone standard dissolved in 40 mL of ethanol. The aldehydes and ketones in rum were transformed into their DNPH derivatives by mixing 1.0 mL of a solution containing 200 mg/100 mL of DNPH with 1.0 μL of H_3PO_4 and 4 mL of rum. After 2 h, a 40.0 μL aliquot was withdrawn and analyzed with HPLC technique. HPLC experiments were carried out by using a LC-10ADVP Shimadzu chromatograph, a SPD-M10A UV–vis diode array detector, and an ODS-C18 Resolve column (25 cm × 4.6 mm × 5 μm). The injection volume was 20.0 μL, and detection was performed at 365 nm. The following methanol/acetonitrile−water gradient was used: methanol/acetonitrile (8:2 v/v)−water (60:40 v/v), isocratic for 9 min, from 60:40 to 95:5 in 16 min, from 95:5 to 60:40 in 9 min, and 60:40 isocratic for 15 min. The absorption maximum (λ_{max}) for the ketones and aldehydes, as their DNPH derivatives, changed from 310 to 380 nm. Since their molar absorptivities did not change very much (less than 7%) from their respective λ_{max} to 365 nm, this wavelength was chosen for their analytical detection and quantification, except for acetylacetone (*23*). For acetylacetone, the measurement was performed at $\lambda_{max} = 310$ nm.

*Metals (**8, 24**).* The samples (50.0 mL) were placed in an open 100.0 mL beaker and digested with 5.0 mL of HNO_3 under controlled heating until the sample volume was reduced to 5.0 mL. After cooling to room temperature, the remaining liquid was quantatively transferred to a 25.0 mL volumetric flask, and the volume was adjusted up to the mark by using a 5% nitric acid solution. The analyses were performed by ICP OES (Optima 300 dual view, Perkin-Elmer) and AAS (Hitachi Z-8100). The instrumental conditions and analytical lines for each element are given in **Table 1** for ICP OES and in **Table 2** for AAS. The calibration curves were constructed by using the external standard method.

*Caramel (**25**).* A Hitachi U-3501 UV–vis spectrophotometer was used. The measurements were carried out at $\lambda = 210$ and 282 nm, following a methodology previously described in the literature (*25*). When necessary for the absorbance measurements, the samples were diluted with an aqueous ethanol solution at the proportion of 1:20 for rum.

Table 1. Operation Conditions for ICP OES for the Determination of the Metals

parameters	conditions
nebulizer Ar flow rate	0.8 L·min^{-1}
operating power	1300 W
coolant Ar flow rate	0.5 L·min^{-1}
plasma Ar flow rate	15 L·min^{-1}
nebulizer type	cross-flow pneumatic
sample flow rate	1 mL·min^{-1}

element	λ (nm)
magnesium	279.079
calcium	317.933
zinc	213.856
barium	233.527
iron	239.562
lead	220.353
manganese	260.569
cobalt	238.892
nickel	232.003
chromium	206.158
strontium	232.235
cadmium	226.502
copper	324.754

Table 2. Operation Parameters for AAS for the Determination of Lithium and Sodium

parameters	lithium	sodium
lamp current	10.0 mA	10.0 mA
wavelength	670.8 nm	589 nm
flame gas	C_2H_2−sintetic Ar	C_2H_2−sintetic Ar
fuel flow	1.7 L·min^{-1}	2.2 L·min^{-1}
oxidant pressure	160 kPa	160 kPa
burner height	7.5 mm	7.5 mm

Chemometric Data Treatment. The data from the chemical analyses of 61 analytes in 44 samples were organized in a 61 × 44 matrix, and the chemical variables were normalized before the statistical treatment.

The statistical methods, PCA, LDA, and PLS-DA, were performed by using MINITAB Release 14 (Statistical Software, PA) software.

PCA identifies, in the hyperspace of variables, the directions where most of the information is retained, thus reducing the dimensionality of the system. By projecting the samples of the data set into the space of the first few components, it is possible to demonstrate the differences among the various samples and, at the same time, to determine the variables which are more involved (*9–12, 17, 18, 26*).

The discriminate analysis applied here is based on the PLS-DA, a multivariate regression technique widely used in the chemical sciences (*8, 17, 27*). Conversely to PCA, the exploratory technique PLS-DA is able to predict a classification. Additionally, it allows the visualization of the results through the samples score plots, which is not possible when the LDA technique is used. The PLS-DA was performed from the normalized data. In the normalization step, each original value of the data set is divided by the sum of all the absolute values. The matrix data *X* (61 compounds) are related to a set of class variables, and *Y* (44 rums) are the variables used to indicate the class samples (see Supporting Information).

Later, LDA was used for classification purposes. LDA is a classification procedure (*9, 11, 18, 19*) that maximizes the variance between categories and minimizes the variance within categories. This method renders a number of orthogonal linear discriminate functions, equal to the number of categories minus one ($n - 1$).

A proper stepwise feature-selection procedure was performed as a pre-treatment of the data aiming at identifying the variables with the highest discrimination abilities, hence decreasing the number of variables to be used. The data processing was performed with the LDA

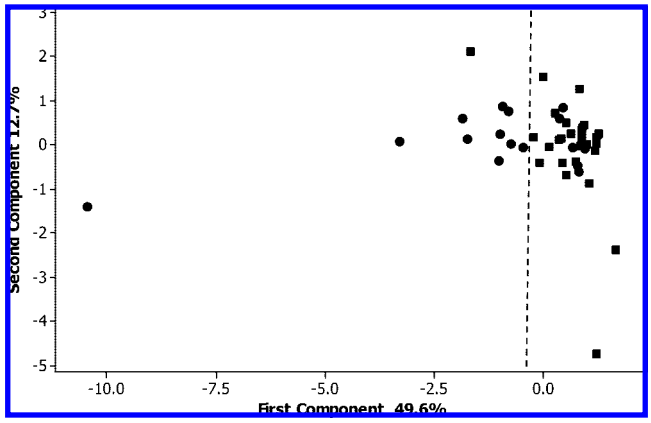

Figure 1. Score plot PCA for rums: class 1 (•), 18 Cuban rum samples; class 2 (■), 26 non-Cuban rum samples.

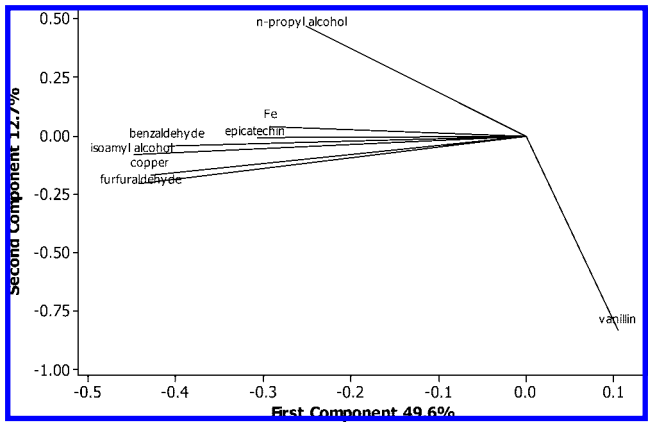

Figure 2. Loading plot of PCA (PC1 × PC2) for 44 samples of rums (Cuban and non-Cuban).

technique through calculations of linearity. This procedure eliminates those variables the information of which, although possibly significant for other objectives, does not contribute to our specific purpose.

In the LDA technique, after creating the model from the original data matrix, it was essential to evaluate its reliability, for which a division of the database was performed. After that, the same model was tested with 10 unknown samples randomly selected. This selection was performed by taking into account the distribution of the samples in **Figures 1** and **2**, avoiding sample clustering. The matrix was composed of 34 lines (samples) and 61 columns (variables), and the degree of confidence was evaluated by using the same data matrix from which the model was generated. The self-consistency of the LDA model was examined through a cross-validation procedure by using the same samples those of the model. Finally, the model was tested with the 10 samples previously selected and unknown to the model. The number of samples used to test the LDA model was in accordance with the number of rums available (*17, 18*).

During the cross-validation procedure, one sample at a time (of *n* samples) is left out, and the prediction ability is tested for the omitted sample. This procedure is repeated *n* times, resulting in *n* models, and provides an estimate of the average prediction ability for the *n* models.

RESULTS AND DISCUSSION

All the analytical data collected from the analyses of volatiles compounds, polyphenols, aldehydes, ketones, and metals (61 compounds) of the 44 rum samples are presented as Supporting Information. **Table 3** shows the median, maximum, minimum, and average values of the analytical data for the 61 analytes. The rums were separated into class 1 (Cuban rums) and class 2 (non-Cuban rums).

All the chemometric analyses were performed by using the full set of data. **Table 4** summarizes the analytical data in terms of median, maximum, minimum, and average values for metal ions, aldehydes, ketones, caramel, alcohols, and phenolic compounds, which have been selected by the chemometric methods tested herein as the most important descriptors.

In rums, for example, the higher alcohols, such as *n*-propyl alcohol and isobutyl alcohol, are formed from their correspondent keto-acid, following the route by which ethanol is converted from pyruvate (*3, 19, 28*).

The origin of the ethyl lactate and *sec*-butyl alcohol is associated with the activity of lactic acid bacteria, which also cause an increase in the concentration of ethyl hexanoate and ethyl octanoate (*3, 29, 30*). High concentrations of methyl and ethyl acetate are indications of aerobiosis in the raw material during the fermentation process or the result of an incorrect separation of the first fraction (head) during distillation (*30*). The ethyl acetate is, in general, the major ester in distilled beverages. When its perception threshold (33 $mg \cdot L^{-1}$) is exceeded, it is reported to contribute with nuances of glue and dissolvent in the Orujo spirits (567 $mg \cdot L^{-1}$) (*30*).

Strongly smelling aliphatic alcohols, such as *n*-propyl alcohol, isobutyl alcohol, isoamyl alcohol, and aromatic phenylethyl alcohol, may be formed from sugars by an anabolic process through the pathways of the amino acids synthesis (*3, 19, 28*).

Aldehydes can also be formed through the reduction of fatty acids, but they do not occur very frequently in alcoholic fermentation. Several aldehydes can be formed from the amino acids present in the sugar cane broth (*3*). The partial degradation of amino acids would account for the formation of higher alcohols which, in the presence of oxygen, can be converted into aldehydes (*3*).

5-Hydroxymethyl-2-furfuraldehyde (5-HMF) and furfuraldehyde are not formed during fermentation; however, they can appear in sugar cane juice when its harvest is preceded by the burning of the foliage that leads to partial dehydration of a small fraction of the sugars. Dehydration of pentoses and hexoses generates furfuraldehyde and 5-HMF, respectively (*28*). Aldehydes are known to reach up to 5−9% v/v in rums from Jamaica, Puerto Rico, and Martinique, acetaldehyde being the predominant compound (*3*). According to our data (see **Table 3** and Supporting Information), the acetaldehyde concentration is higher in Cuban rums than in the other samples.

As observed for cachaça (*31*), the presence of benzaldehyde in a beverage at an average level of 0.591 mg/100 mL anhydrous alcohol (A.A.) can possibly indicate that the spirit was distilled in columns, whereas the presence of methanol at an average level of 20.9 mg/100 mL A.A. would suggest that the spirit was distilled in pot stills.

High concentrations of benzaldehyde were found in the samples produced in Cuba, at an average level of 56.7 $mg \cdot L^{-1}$, whereas the samples produced in the 14 remaining countries showed higher concentrations of methanol at 34.7 $mg \cdot L^{-1}$ maximum. These data suggest that Cuban rums were distilled in columns, whereas the others were probably distilled in pot stills.

The concentrations of 5-HMF found in non-Cuban rums were higher than those found in Cuban products. The non-uniform heating or overheating of wine (fermented must) will lead to an increase in the 5-HMF concentration. Thus, 5-HMF is in general found more abundantly in homemade cachaças (pot still) than in industrial cachaças (column), and the non-uniform heating and even the overheating of the alembics contrasting with the more controlled heating of the columns would in part account for that (*28*).

Table 3. Average, Median, Maximum, and Minimum Concentrations ($mg \cdot L^{-1}$) of Constituents of the Samples Separated in Two Classes (Cuban and Non-Cuban Rums)[a]

	average		median		maximum		minimum	
chemical descriptors classification	Cuban	non-Cuban	Cuban	non-Cuban	Cuban	non-Cuban	Cuban	non-Cuban
gallic acid	NQ	NQ	NQ	NQ	NQ	NQ	NQ	NQ
4-H-3-M	0.711	0.851	0.530	0.709	2.26	3.70	0.306	ND
vanillin	0.615	1.11	0.517	0.542	1.25	6.74	ND	ND
4-hydroxy	1.45	3.13	1.03	1.08	4.86	7.37	ND	ND
4-H-3.5-D	2.22	5.27	1.59	1.49	6.96	33.6	ND	ND
scopoletin	ND	ND	ND	ND	ND	ND	ND	ND
coumarin	ND	ND	ND	ND	ND	ND	ND	ND
sinapaldehyde	ND	0.643	ND	ND	<LOQ a	0.643	ND	ND
coniferaldehyde	1.31	0.636	1.31	<LOQ b	2.27	0.671	ND	ND
trans	ND	ND	ND	ND	ND	ND	ND	ND
ellagic acid	ND	ND	ND	ND	ND	ND	ND	ND
myricetin	ND	ND	ND	ND	ND	ND	ND	ND
quercetin	ND	ND	ND	ND	ND	ND	ND	ND
eugenol	ND	ND	ND	ND	ND	ND	ND	ND
catechin	0.153	0.254	0.139	0.165	0.246	0.848	ND	0.0640
epicatechin	0.309	0.237	0.206	0.207	0.729	0.522	ND	0.0650
2-propenal	5.11	8.25	4.58	3.57	11.8	30.7	ND	ND
acetone	1.64	1.38	1.34	1.38	5.80	3.43	0.112	ND
et. acet.	43.3	106	40.5	57.0	107	564	1.32	ND
methanol	19.2	18.2	22.3	16.7	33.7	34.7	ND	ND
butOH-2	1.62	20.5	1.62	10.3	1.62	65.8	ND	ND
1-propanol	78.5	69.5	49.9	36.5	197	343	1.32	ND
isobutOH	120	50.4	96.9	33.2	553	230	1.02	ND
1-butanol	1.37	38.8	1.28	2.70	1.99	184	ND	ND
isoamyl	429	145	164	100	205×10^1	392	1.57	ND
acetic	169	186	157	105	827	661	1.75	ND
lithium	ND	ND	ND	ND	ND	ND	ND	ND
magnesium	0.780	0.683	0.786	0.380	1.64	2.06	<LOQ c	<LOQ c
sodium	12.6	10.4	12.3	10.1	34.9	29.3	1.45	0.512
calcium	0.702	1.04	0.483	0.722	2.80	3.23	0.199	<LOQ c
zinc	0.0860	0.0780	0.0350	0.0360	0.570	0.222	<LOQ d	<LOQ d
barium	<LOQ c	<LOQ c	<LOQ c	<LOQ c	<LOQ c	<LOQ a	ND	ND
iron	0.200	0.0870	0.115	0.0550	0.648	0.291	0.0210	<LOQ d
lead	0.0130	0.0150	0.0130	0.0150	0.0130	0.0160	<LOQ d	ND
manganese	0.0190	0.0450	0.0160	0.0190	0.0310	0.118	ND	ND
cobalt	ND	ND	ND	ND	ND	ND	ND	ND
nickel	ND	ND	ND	ND	ND	ND	ND	ND
chromium	ND	ND	ND	ND	ND	ND	ND	ND
cadmium	ND	ND	ND	ND	ND	ND	ND	ND
strontium	0.0440	0.0400	0.0350	0.0320	0.0800	0.0970	0.0190	ND
copper	0.897	0.424	0.836	0.373	2.76	1.08	<LOQ c	<LOQ c
acetil	4.45	1.62	ND	ND	4.45	1.62	ND	ND
formal	2.11	2.73	2.01	1.86	7.34	12.8	0.145	0.352
5-HMF	17.4	48.8	10.9	13.8	79.3	269	ND	ND
acetal	169	45.4	42.5	26.7	142×10^1	270	2.88	2.15
furfural	36.3	5.14	15.6	5.14	84.1	7.23	ND	ND
propional	1.15	0.948	1.12	0.648	2.48	3.46	ND	ND
2.3 but	8.60	9.74	1.33	2.77	46.5	72.7	ND	ND
crotonal	7.19	18.7	4.40	4.24	18.3	137	ND	ND
isobut/but	3.87	13.8	1.18	1.50	31.6	171	0.267	ND
ciclopent	0.696	1.61	0.696	1.05	1.02	4.20	ND	ND
benzal	56.7	11.5	17.4	6.86	208	46.5	ND	ND
isovaleral	18.2	12.5	2.13	1.21	246	97.1	0.332	ND
valeral	7.99	15.5	3.60	5.10	33.8	61.8	ND	ND
acetofe	3.74	5.06	4.06	3.27	5.65	23.2	ND	ND
isoamil	7.36	ND	ND	ND	7.36	ND	ND	ND
caramel	448	591	193	521	193×10^1	195×10^1	<LOQ e	<LOQ e

[a] NQ, no quatification; ND, not detected; <LOQ, smaller than the limit of quantification (a = 0.500 $mg \cdot L^{-1}$, b = 0.150 $mg \cdot L^{-1}$, c = 0.125 $mg \cdot L^{-1}$, d = 0.0100 $mg \cdot L^{-1}$, and *e* = 0.200 $g \cdot L^{-1}$); 4-hydroxy, 4-hydroxy-3,5-dimethoxybenzoic acid; 4-H-3,5-D, 4-hydroxy-3,5-dimethoxybenzaldehyde; 4-H-3-M, 4-hydroxy-3-methoxybenzoic acid; trans, trans-resveratrol; butOH-2, sec-butyl alcohol; et. acet., ethyl acetate; isobutOH, isobutyl alcohol; isoamyl, isoamyl alcohol; acetic, acetic acid; acetil, acetylacetone; formal, methaldehyde; acetal, acetaldehyde; furfural, 2-furfuraldehyde; propional, propionaldehyde; 2.3-but, 2,3-butanodione monoxime; crotonal, 2-butenal; isobut/but, isobutyraldehyde + butyraldehyde; ciclopente, cyclopentanone; benzal, phenylmethanal; isovaleral, iso-pentanal; valeral, *n*-pentanal; acetofe, methyl-phenyl-ketone; isoamil, methyl-isoamil-ketone.

Esters are key compounds in spirits because of their contribution to the aroma. In rum, ethyl acetate is typically the predominant ester. When yeast is present during distillation, the content of long-chain carboxylic acid esters derived from the yeast cells increases. The ester content of rum also depends on the yeast (*3*).

In a related paper, a model has been proposed in order to distinguish between homemade (pot still) and industrial cachaças (columns). Considering that Cuban rums are preferably distilled in columns and the others in alembics, we unsuccessfully tried to apply to rums the approach developed for the distinction of cachaças. Rums and cachaças can been distinguished on the

Table 4. Average, Median, Maximum, and Minimum Concentrations of Constituents of the Principal Chemical Descriptors Used for the Classification of All Rums[a]

chemical descriptors classification	concentrations ($mg \cdot L^{-1}$)			
	average	median	maximum	minimum
acetone	1.50	0.112	5.80	<LOQ a
4-hydroxy-3,5-dimethoxybenzoic acid	1.95	ND	7.37	ND
n-propyl alcohol	73.4	176	343	ND
acetaldehyde	96.0	35.9	142×10^1	2.15
calcium	0.883	0.640	3.23	<LOQ b
2-butenal	15.4	<LOQ c	137	ND
cyclopentanone	1.31	ND	4.20	ND
lead	0.0140	<LOQ d	0.0160	ND
ethyl acetate	79.8	46	564	ND
caramel	133	72.4	488	<LOQ e
sec-butyl alcohol	17.8	ND	65.8	ND
magnesium	0.709	<LOQ b	2.06	<LOQ b
sodium	11.3	17.2	34.9	0.512
strontium	0.0418	0.0409	0.0965	ND
n-butyl alcohol	24.8	0.834	184	0.834
vanillin	0.946	0.633	6.74	ND
zinc	0.0819	0.0974	0.570	<LOQ d
barium	0.0130	<LOQ b	0.0130	ND
catechin	0.204	<LOQ f	0.848	<LOQ f
epicatechin	0.271	<LOQ f	0.729	<LOQ f
isoamyl alcohol	270	192	205×10^1	1.57
4-hydroxy-3-methoxybenzoic acid	0.781	ND	3.7	ND
isobutyl alcohol	81.1	121	553	<LOQ g

[a] ND, not detected; LOQ, smaller than the limit of quantification (a = 0.160 $mg \cdot L^{-1}$, b = 0.125 $mg \cdot L^{-1}$, c = 0.300 $mg \cdot L^{-1}$, d = 0.0100 $mg \cdot L^{-1}$, e = 0.200 $g \cdot L^{-1}$, f= 0.0251 $mg \cdot L^{-1}$, and g = 0.500 $mg \cdot L^{-1}$).

basis of their chemical profile (*8*); therefore, it is reasonable to suppose that the most representative chemical descriptors for both beverages would not necessarily be the same. However tentatively, some general analogies between cachaça and rum composition will, by means of their respective production process, be presented in this section.

According to the PCA (see Materials and Methods), from the first 61 analytes, only the isoamyl alcohol, *n*-propyl alcohol, copper, iron, furfuraldehyde, benzaldehyde, epicatechin, and vanillin revealed superior discriminant properties. Thus, considering only these eight chemical compounds, the first three principal components [PC1 (49.6%), PC2 (12.7%), and PC3 (8.1%)] would account for 70.4% of the total variability, which is considered sufficient for exploratory analysis purposes. **Figure 1** presents the score plot obtained from PC1 × PC2. A two-dimensional plot of the objects (rums) in the space defined by the first principal component shows a tendency of the objects to separate into two groups. To illustrate this separation, a line was arbitrarily drawn on the PCA score plot. Cuban rums are on the left side of the line and non-Cuban rums on the right side. Although the same samples are located far from their group, as it can be seen on the far left-hand side of the score plot (**Figure 1**), they were correctly classified, and their withdrawal from the database did not provide a significant statistical gain, so they were kept in the analyses.

According to **Figure 2**, it is clear that benzaldehyde, isoamyl alcohol, copper, and furfuraldehyde compounds that presented high loading values in PC1 are characteristic of rums produced in Cuba, whereas compounds that presented low values are characteristic of rums produced in other countries.

In the PLS-DA, the first three components (C1, C2, and C3) explain 55.1% of the total variance of the original data and, according to the score plot, **Figure 3** exhibits a better tendency

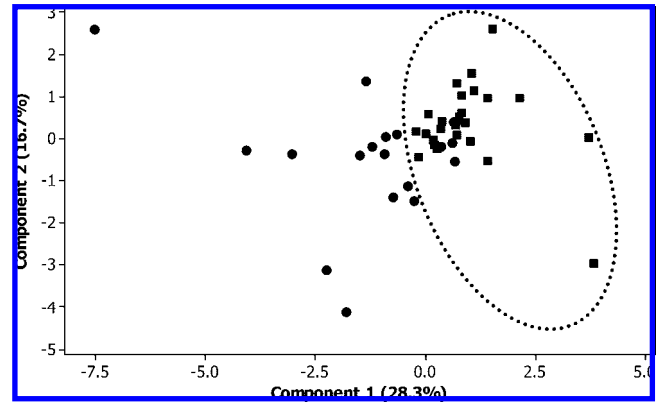

Figure 3. Score plot PLS-DA. (•) Cuban rums (class 1), 18 samples; (■) non-Cuban rums (class 2), 26 samples.

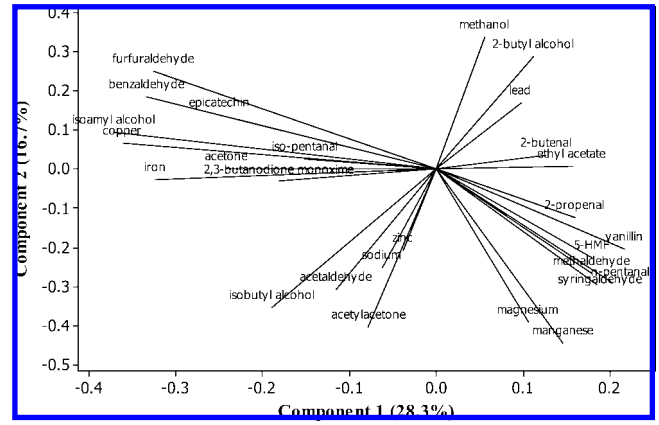

Figure 4. Loading plot PLS-DA (C1 × C2) for 44 samples of rums (Cuban and non-Cuban).

for rum separation than that exhibited in **Figure 1**. Among the 61 analytes, 26 were selected by the PLS-DA technique as the most significant descriptors. The selected discriminators are acetone, *sec*-butyl alcohol, isobutyl alcohol, ethyl acetate, methanol, isoamyl alcohol, magnesium, sodium, lead, iron, manganese, copper, zinc, syringaldehyde, formaldehyde, 5-HMF, acetaldehyde, furfuraldehyde, 2-butenal, *n*-pentanal, iso-pentanal, phenylmethanal, 2,3-butanodione monoxime, acetylacetone, epicatechin, and vanillin. During the PLS-DA, the variables were selected through an inclusion and exclusion process for latent variables by using the PC1 and PC2 loading values of each compound, and the compounds with the lowest values were discarded.

According to the loading values, **Figure 4**, furfuraldehyde, benzaldehyde, epicatechin, isoamyl alcohol, copper, acetone, iso-pentanal, iron, and 2,3-butanodione monoxime contribute more significantly to characterize the samples in class 1 (Cuban rums) in the first component. This behavior is similar to the one described for cachaça, where benzaldehyde, acetaldehyde, isoamyl alcohol, and acetone are usually more abundant in column distillates than in pot still products (*31*). To better illustrate that, an arbitrary contour resembling an ellipse was drawn in the score plot of the PLS-DA.

These four compounds are more abundant in the head and tail fractions than in the heart fraction of the pot still cachaça, and it is likely that the same phenomenon occurs for rums. However, in the column distillation, only one fraction is collected, and the distilled product is bottled, thus explaining the increased presence of these volatiles in this spirit.

The last step, following a variable reduction guided by the PCA results, was to apply LDA to the data set to generate a

Table 5. Model for the Classification of Samples Derived from LDA for the Two Groups of Rums (Cuban and non-Cuban)

	true group	
model	Cuban	non-Cuban
Cuban	13	1
non-Cuban	0	20
total number of samples	13	21
total number correctly assigned	13	20
percentage (%)	100	95.2
total of sampling	34	
total of correct classified	33	
correct assignment (%)	97.1	

Table 6. Cross-Validation for the Classification of Samples by Using the Model Derived from LDA (Cuban and non-Cuban Rums)

	true group	
cross-validation	Cuban	non-Cuban
Cuban	11	2
non-Cuban	2	19
total number of samples	13	21
total number correctly assigned	11	19
percentage (%)	84.6	90.5
total of sampling	34	
total of correct classified	30	
correct assignment (%)	88.2	

classification model rule. The reliability of the classification models achieved was studied in terms of recognition ability (percentage of the members of the training set correctly classified) and prediction ability (percentage of the members of the test set correctly classified by using the rules developed in the training step) (*13*).

The descriptors considered in the LDA model were ethyl acetate, *sec*-butyl alcohol, *n*-propyl alcohol, *n*-butyl alcohol, isoamyl alcohol, isobutyl alcohol, caramel, catechin, vanillin, epicatechin, manganese, acetaldehyde, 4-hydroxy-3-methoxybenzoic acid, 2-butenal, 2-butenal, 4-hydroxy-3,5-dimethoxybenzoic acid, cyclopentanone, acetone, lead, zinc, calcium, barium, strontium, and sodium.

The LDA model (**Table 5**) was generated by using the analytical data for 34 samples, where 13 were Cubans and 21 non-Cubans. According to **Table 5**, the LDA technique correctly identified 33 samples from a total of 34 rums, leading to a 97.1% correct assignment. It can also be observed that all 13 Cuban rum samples were correctly classified (100%), whereas only the Mexican rum was misclassified among the non-Cuban rums (class 2). Thus, a 95.2% accuracy was achieved.

To evaluate the prediction ability, the cross-validation method was applied, and the model was tested with unknown samples. The self-consistency of the LDA model was examined through a cross-validation procedure by using the same samples employed to build up the model. Thus, according to our data (**Table 6**), the LDA method described in this paper proved to be able to distinguish Cuban rum from non-Cuban rum with 88.2% accuracy.

The model was then checked against a test group formed with 10 samples unknown to the model (5 produced in Cuba and 5 in other countries). **Table 7** shows the data obtained from checking the LDA model with the unknown samples, namely, the test group. Only one sample in each group was erroneously classified, thus allowing 80% accuracy in the distinction between the two groups of rums.

It is interesting to point out that among all the compounds used in the PCA and LDA, four of them have exhibited relevant discriminant properties in both techniques, namely, isoamyl alcohol, *n*-propyl alcohol, epicatechin, and vanillin.

Table 7. Application of the Model Derived from LDA for the Classification of Samples from Unknown Origin (Cuban and non-Cuban Rums)

	true group	
group test	Cuban	non-Cuban
Cuban	4	1
non-Cuban	1	4
total number samples	5	5
total number correctly assigned	4	4
percentage (%)	80.0	80.0
total of sampling	10	
total of correct classified	8	
correct assignment (%)	80.0	

Although a large number of discriminators were analyzed, others could certainly be defined. However, despite the limited number of commercial samples, the paper presented reliable results to distinguish rums. The sample group can certainly be enlarged. Nonetheless, the experimental data analyses performed here depict the present rum chemical profiles and allow the identification of chemical discriminators able to distinguish Cuban rum from non-Cuban rums. The LDA methodology using the analytical data for 23 analytes is the most promising technique to identify rum from Cuba, and it can also be a useful tool to certify the geographic origin of the beverage. Furthermore, it assures the consumer of a genuine product. The procedure can be recommended as a routine method for forensic purpose.

ACKNOWLEDGMENT

The authors are thankful to Dr Luis O. Gálvez Taupier (ICIDCA-Cuba) and ABRABE for providing some rum samples and to Dr. Marieta Vasallo (CERALBE) for valuable information.

Supporting Information Available: Complete analytical data used in all of the statistical methods. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Pino, J.; Martí, M. P.; Mestres, M.; Pérez, J.; Busto, O.; Guasch, J. Headspace solid-phase microextraction of higher fatty acid ethyl esters in white rum aroma. *J. Chromatogr., A* **2002**, *954*, 51–57.

(2) Nicol, D. A. Rum. In *Fermented Beverage Production*, 2nd ed.; Lea, A. G. H., Piggott, J. R., Eds.; Kluwer Academic: New York, **2003**; pp 263–287.

(3) Gómez, S. M. Rum Aroma Descriptive Analysis. Tese (Master of Science), La Salle University, Mexico City, 2002. http://etd.lsu.edu/docs/available/etd-1114102–110301/unrestricted/Maza_Gomez_thesis.pdf.

(4) Dr. Marieta Vasallo,diretora do CERALBE (ICIDCA), personal communication.

(5) Nascimento, R. F.; Cardoso, D. R.; Faria, J. B.; Lima-Neto, B. S.; Franco, D. W. Influência do material do alambique na composição química das aguardentes de cana-de-açúcar. *Quim. Nova* **1998**, *21*, 735–739.

(6) Cardoso, D. R.; Nascimento, R. F.; Lima-Neto, B. S.; Franco, D. W. Influência do material do destilador na composição química das aguardentes de cana - Parte II. *Quim. Nova* **2003**, *26*, 165–169.

(7) Andrade-Sobrinho, L. G.; Boscolo, M.; Lima-Neto, B. S.; Franco, D. W. Carbamato de etila em bebidas alcoólicas (cachaça. tiquira. uísque e grapa). *Quim. Nova* **2002**, *25*, 1074–1077.

(8) Cardoso, D. R.; Andrade-Sobrinho, L. G.; Leite-Neto, A. F.; Reche, R. V.; Isique, W. D.; Ferreira, M. M. C.; Lima-Neto, B. S.; Franco, D. W. Comparison between cachaça and rum using pattern recognition methods. *J. Agric. Food Chem.* **2004**, *52*, 3429–3433.
(9) Latorre, M. J.; García-Jares, C.; Mèdina, B.; Herrero, C. Pattern recognition analysis applied to classification of wines from Galicia (northwestern Spain) with certified brand of origin. *J. Agric. Food Chem.* **1994**, *42*, 1451–1455.
(10) Sivertsen, H. K.; Holen, B.; Nicolaysen, F.; Risvik, E. Classification of French red wines according to their geographical origin by the use of multivariate analyses. *J. Sci. Food Agric.* **1999**, *79*, 107–115.
(11) Arvanitoyannis, I. S.; Katsota, M. N.; Psarra, E. P.; Soufleros, E. H.; Kallithraka, S. Application of quality control methods for assessing wine authenticity: use of multivariate analysis (chemometrics). *Trends Food Sci. Technol.* **1999**, *10*, 321–336.
(12) Kokkinofta, R. I.; Theocharis, C. R. Chemometric characterization of the cypriot spirit "Zivania". *J. Agric. Food Chem.* **2005**, *53*, 5067–5073.
(13) Alonso-Salces, R. M.; Guyot, S.; Herrero, C.; Berruela, L. A.; Drilleau, J. F.; Gallo, B.; Vicente, F. Chemometric characterisation of Basque and French ciders according to their polyphenolic profiles. *Anal. Bioanal. Chem.* **2004**, *379*, 464–475.
(14) Penza, M.; Cassano, G. Chemometric characterization of Italian wines by thin-film multisensors array and artificial neural networks. *Food Chem.* **2004**, *86*, 283–296.
(15) Martí, M. P.; Busto, O.; Guasch, J. Application of a headspace mass spectrometry system to the differentiation and classification of wines according to their origin variety and ageing. *J. Chromatogr., A* **2004**, *1057*, 211–217.
(16) Alonso-Salces, R. M.; Herrero, C.; Barranco, A.; López-Márquez, D. M.; Berruela, L. A.; Gallo, B.; Vicente, F. Polyphenolic composition of Basque natural ciders: A chemometric study. *Food Chem.* **2006**, *97*, 438–446.
(17) Togari, N.; Kobayashi, A.; Aishima, T. Pattern recognition applied to gas chromatographic profiles of volatiles components in three tea categories. *Food Res. Int.* **1995**, *28*, 495–502.
(18) Fernández-Cáceres, P. L.; Martíns, M. J.; Pablos, F.; González, A. G. Differentiation of tea (Camellia sinensis) varieties and their geographical origin according to their metal content. *J. Agric. Food Chem.* **2001**, *49*, 4775–4779.
(19) Herranz, A.; De La Serra, P.; Barro, C.; Martin-Alvarez, P. J. Multivariate statistical methods applied to the differentiation of rum brands. *J. Sci. Food Agric.* **1990**, *51*, 555–560.
(20) Boscolo, M.; Bezerra, C. W. B.; Cardoso, D. R.; Lima-Neto, B. S.; Franco, D. W. Identification and dosage by HRGC of minor alcohols and esters in Brazilian sugar-cane spirit. *J. Braz. Chem. Soc.* **2000**, *11*, 86–90.
(21) Silva, A. A. Compostos fenólicos em madeiras Brazileiras. Tese (Mestrado em Química Analítica), Instituto de Química de São Carlos, Universidade de São Paulo, 2006. http://www.teses.usp.br/teses/disponiveis/75/75132/tde-16042007–110050/.
(22) Nascimento, R. F.; Marques, J. C.; Lima-Neto, B. S.; De Keukeleire, D.; Franco, D. W. Qualitative and quantitative high-performance liquid chromatographic analysis of aldehydes in Brazilian sugar cane spirits and other distilled alcoholic beverages. *J. Chromatogr., A* **1997**, *782*, 13–23.
(23) Cardoso, D. R.; Bettin, S. M.; Reche, R. V.; Lima-Neto, B. S.; Franco, D. W. HPLC-DAD analysis of ketones as their 2.4-dinitrophenylhydazones in Brazilian sugar-cane spirits and rum. *J. Food Compos. Anal.* **2003**, *16*, 563–573.
(24) Nascimento, R. F.; Bezerra, C. W. B.; Furuya, S. M. B.; Schultz, M. S.; Polastro, L. R.; Lima-Neto, B. S.; Franco, D. W. Mineral profile of Brazilian cachaça and other International dpirits. *J. Food Compos. Anal.* **1999**, *12*, 17–25.
(25) Boscolo, M.; Andrade-Sobrinho, L. G.; Lima-Neto, B. S.; Franco, D. W. Spectrophotometric determination of caramel content in spirits aged in oak casks. *J. AOAC Int.* **2002**, *85*, 744–750.
(26) Ferreira, M. M. C.; Antunes, A. M.; Melgo, M. S.; Volpe, P. L. O. Quimiometria I: calibração multivariada, um tutorial. *Quím. Nova* **1999**, *22*, 724–731.
(27) Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* **2003**, *17*, 166–173.
(28) Maia, A. B. Componentes secundários da aguardente. *Tec./Pesq.* **1994**, *12*, 29–34.
(29) Diéguez, S. C.; De La Peña, M. L. G.; Gómez, E. F. Volatile composition and sensory characters of commercial Galician orujo spirits. *J. Agric. Food Chem.* **2005**, *53*, 6759–6765.
(30) Cortés, S.; Gil, M. L.; Férnandez, E. Volatile composition of traditional and industrial orujo spirits. *Food Control* **2005**, *16*, 383–388.
(31) Reche, R. V.; Leite Neto, A. F.; Silva, A. A.; Galinaro, C. A.; Osti, R. Z.; Franco, D. W. Influence of type of distillation apparatus on chemical profiles of brazillian cachaças. *J. Agric. Food Chem.* **2007**, *55*, 6603–6608.

Received for review September 10, 2007. Revised manuscript received December 18, 2007. Accepted December 25, 2007. This work was supported by CNPq, FAPESP (2001/12934-0), and CAPES.

JF0726841