



**Universidade de São Paulo**

**Biblioteca Digital da Produção Intelectual - BDPI**

---

Outros departamentos - IQSC/Outros

Artigos e Materiais de Revistas Científicas - IQSC/Outros

---

2012

# Quantitative chemical profile and multivariate statistical analysis of alembic distilled sugarcane spirit fractions

---

J. Braz. Chem. Soc., v.23, n.8, p.1506-1514, 2012

<http://www.producao.usp.br/handle/BDPI/39235>

*Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo*

## Quantitative Chemical Profile and Multivariate Statistical Analysis of Alembic Distilled Sugarcane Spirit Fractions

Felipe A. T. Serafim,<sup>a</sup> Carlos A. Galinaro,<sup>a</sup> Alexandre A. da Silva,<sup>a</sup> Silmara F. Buchviser,<sup>a</sup> Eduardo S. P. do Nascimento,<sup>a</sup> Fernando V. Novaes<sup>b</sup> and Douglas W. Franco<sup>\*a</sup>

<sup>a</sup>Instituto de Química de São Carlos, Universidade de São Paulo, Avenida Trabalhador São-Carlense 400, CP 780, 13560-970 São Carlos-SP, Brazil

<sup>b</sup>Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ), Universidade de São Paulo, Av. Pádua Dias, 11, 13418-900 Piracicaba-SP, Brazil

As concentrações de 39 compostos orgânicos foram determinadas em três frações (cabeça, coração e cauda) obtidas da destilação em alambique do caldo de cana fermentado. Os resultados foram avaliados utilizando-se análise de variância (ANOVA), teste de Tukey, análise de componentes principais (PCA), agrupamento hierárquico (HCA) e análise discriminante linear (LDA). De acordo com PCA e HCA, os dados experimentais conduzem à formação de três agrupamentos. As frações de cabeça deram origem a um grupo mais definido. As frações coração e cauda apresentaram alguma sobreposição coerente com sua composição em ácidos. As habilidades preditivas de calibração e validação dos modelos gerados pela LDA para a classificação das três frações foram de 90,5 e 100%, respectivamente. Este modelo reconheceu como coração doze de treze cachaças comerciais (92,3%) com boas características sensoriais, apresentando potencial para a orientação do processo de cortes.

Concentrations of 39 organic compounds were determined in three fractions (head, heart and tail) obtained from the pot still distillation of fermented sugarcane juice. The results were evaluated using analysis of variance (ANOVA), Tukey's test, principal component analysis (PCA), hierarchical cluster analysis (HCA) and linear discriminant analysis (LDA). According to PCA and HCA, the experimental data lead to the formation of three clusters. The head fractions give rise to a more defined group. The heart and tail fractions showed some overlap consistent with its acid composition. The predictive ability of calibration and validation of the model generated by LDA for the three fractions classification were 90.5 and 100%, respectively. This model recognized as the heart twelve of the thirteen commercial *cachaças* (92.3%) with good sensory characteristics, thus showing potential for guiding the process of cuts.

**Keywords:** sugarcane spirits, alembic fractions, multivariate analysis

### Introduction

According to the Brazilian Cachaça Institute, the volume of sugarcane spirit (*cachaça*) exported was around 9.8 million liters in 2011 and generated US\$17.28 million in revenue. Currently, Germany, the United States, Portugal and France are the main importers.<sup>1</sup>

With the increasing demand for high-standard quality *cachaça*, the adoption of a more efficient production technique is a challenge that needs to be overcome by producers. *Cachaça* chemical composition and its sensory

quality depend mainly on sugarcane juice, alcoholic fermentation and the subsequent wine distillation procedures.<sup>2-7</sup>

Usually, in small-scale production (artisanal sugarcane spirits), the distillation process occurs in alembics (pot stills). During this process, the distillate is separated into three different fractions, through operations named cuts, to improve the beverage quality. The alcohol content in the distillate collected from the first cut (head fraction) reaches the range of 70.0 to 55.0% (v/v) or a volume equivalent to 5-10% of the total distillate volume.<sup>8</sup> The second fraction, the heart (the noble distillate), begins to be collected when the alcohol content reaches 55% (v/v)

\*e-mail: douglas@iqsc.usp.br

and ends up around 38% (v/v), which corresponds to 75-80% of the total distillate volume. The last fraction, the tail, is collected when the alcohol content of distillate is below 38% (v/v), which corresponds to about 10% of the total distillate volume.<sup>8</sup> Usually the head and tail fractions are discarded or used by some producers in the next distillation batch. All the alcohol content data are collected at a temperature around  $20 \pm 3$  °C. This is a traditional separation procedure in Brazil and was probably inspired by the colonizer experience with grappa, bagaceira and whiskey production. In small distilleries, each ton of sugarcane processed can produce 80 to 120 distillate (heart fraction) liters.

It is known for different distilled alcoholic beverages that the cutting processes, applied during alembic distillation, can influence the chemical composition and consequently affect the product quality.<sup>9-15</sup> A part from a few examples, in case of sugarcane spirits obtained through alembic distillation, a quantitative analysis of the cutting process fractions has not received the deserved attention compared to what is described for other distilled spirits. Cardeal *et al.*<sup>16</sup> monitored the alembic distillation process by fingerprint chromatogram analysis without using chemometrics tools. Alcarde *et al.*<sup>17</sup> evaluated the secondary components of sugarcane spirits during double distillation process in a rectifying still in order to verify the cut point according to the ethanol concentration between head and tail fractions. Scanavini *et al.*<sup>18</sup> presented the application of a differential distillation model for the simulation of *cachaça* production in alembic.

The study described herein aims to learn more about the spirit secondary compound composition in the three different fractions produced in the alembic distillation process, and thus eventually infer how the cuts procedure would influence the *cachaça* quality with a different approach.

## Experimental

### Samples

The chemical profiles of the samples were determined for the three fractions of the distillate (head, heart and tail). Altogether, 14 alembic distilled *cachaça* samples from different Sao Paulo State producers were evaluated in the present study. The still operations were performed by the producers following their own tradition. None of the distilled samples were submitted to the aging processes and they were all stored in glass bottles under refrigeration (4 °C) and analyzed within three to four months. Thirteen highly ranked non aged commercial *cachaças* previously

evaluated through chemical and sensory analysis<sup>7</sup> were used for the LDA model evaluation.

### Reagents

All reagents were of analytical grade standards (Fluka, Sigma-Aldrich, St. Louis, MO, USA). All solvents (methanol, ethanol, dichloromethane and acetonitrile) were of HPLC grade (Mallinckrodt Baker, Xalostoc, Tlaxcala, México), and all water used was deionized using a Milli-Q system (Millipore, Bedford, MA, USA).

### Distillation apparatus

The copper alembic apparatus were purchased from Brazilian manufacturers and operated by the distillery staff following the respective technical specifications. The alembic distillation process was carried out in single pot stills with similar geometries with capacities ranging from 180-250 L. Direct fire was used for heating alembics where the temperature ranged from 75-90 °C. The fractions were collected at the rate of  $40 \pm 9$  mL *per* min. The cutting process was carried out according to the alcoholic content, which was monitored during the sample distillation.

### Analytical method description

#### Higher alcohols and acetic acid<sup>3</sup>

The presence of methanol, propanol, isobutanol, 1-butanol, 2-butanol, isoamyl alcohol and acetic acid was determined through direct injection of 1.0 µL aliquots of the sample (spiked with 4-methyl-1-propanol, internal standard - 126 mg L<sup>-1</sup>) into a gas chromatograph system (Hewlett-Packard, HP 5890-A GC) using a flame ionization detector (FID) and a HP-FFAP column (cross-linked polyethylene glycol esterified 50 m × 0.20 mm × 0.33 µm film thickness). The inlet and detector temperatures were 250 °C. The injected volume was 1.0 µL, and the split ratio 1:20. The oven temperature program was 55 °C (5 min); 2 °C min<sup>-1</sup> to 100 °C (3 min), 5 °C min<sup>-1</sup> to 190 °C (30 min); 5 °C min<sup>-1</sup> to 220 °C (15 min); oven temperature program (HP-20M) 45 °C; 1 °C min<sup>-1</sup> to 100 °C (3 min), 5 °C min<sup>-1</sup> to 190 °C (30 min); 5 °C min<sup>-1</sup> to 220 °C (15 min). The compound identification and quantification were made through authentic standard addition method.

#### Aldehydes and ketones<sup>5</sup>

Acetylacetone, formaldehyde, 5-hydroxymethylfurfural (5-HMF), acetaldehyde, acrolein, furfuraldehyde, propionaldehyde, butyraldehyde, benzaldehyde, isovaleraldehyde, valeraldehyde and 2,3-butanedione

(diacetyl) were analyzed as their 2,4-dinitrophenylhydrazones (aldehyde-DNPHs) using an HPLC Shimadzu model LC-10AD equipped with a UV-Vis diode array detector (Shimadzu SPD M6A, wavelength = 365 nm). The HPLC separation was performed with a Shimadzu Shim-Pak C18 column (25 cm × 4.6 mm i.d. × 5.0- $\mu$ m particle size) and gradient system of water and methanol/acetonitrile (80:20, v/v) solution. The injection volume was 20.0  $\mu$ L and the following gradient (methanol/acetonitrile)-water was used: (methanol/acetonitrile) (8:2)-water 60:40 (v/v) isocratic for 9 min (1.0 mL min<sup>-1</sup>), from 60:40 to 95:5 in 16 min (1.1 mL min<sup>-1</sup>), from 95:5 to 60:40 in 9 min (1.0 mL min<sup>-1</sup>), 60:40 isocratic for 15 min (1.0 mL min<sup>-1</sup>). The compound identification and quantification were carried out through relative retention time comparison and authentic standard addition.

#### Ethyl carbamate<sup>2</sup>

The determination of the ethyl carbamate concentration was performed through sample direct injection without previous treatment into a gas chromatograph model GC17A (Shimadzu, Tokyo, Japan) hyphenated to a mass selective detector model QP 5050A (Shimadzu, Tokyo, Japan) using electron impact (70 eV) as the ionization source. The mass spectrometer detector operated in SIM mode (*m/z* 62), and propyl carbamate was used as an internal standard (150  $\mu$ g L<sup>-1</sup>). The inlet and detector interface temperature were 250 and 230 °C, respectively. The oven program temperature used was: 90 °C (2 min); 10 °C min<sup>-1</sup> for 150 °C (0 min); 40 °C min<sup>-1</sup> for 230 °C (10 min). The injected volume was 1.0  $\mu$ L in the splitless mode. The ethyl carbamate quantification was carried out through authentic standard addition.

#### Esters<sup>19</sup>

Ethyl acetate, ethyl butyrate, ethyl hexanoate, ethyl lactate, ethyl octanoate, ethyl nonanoate, ethyl decanoate, ethyl laurate and isoamyl octanoate were analyzed by direct sample injection. The volume of 1  $\mu$ L was injected into a gas chromatography model GC17A (Shimadzu, Tokyo, Japan) hyphenated to a mass selective detector model QP 5050A (Shimadzu, Tokyo, Japan) using electron impact (70 eV) as the ionization source, using 4-methyl-2-pentanol as internal standard. The target analytes were separated through a capillary column coated with an esterified polyethylene glycol phase (HP-FFAP; 50 m × 0.20 mm × 0.33  $\mu$ m film) (Hewlett-Packard, Palo Alto, CA). The temperature of the injector and detector interface was 220 °C. The oven temperature was programmed from 35 to 180 °C at a rate of 5 °C min<sup>-1</sup> and then raised at 20 °C min<sup>-1</sup> increments from 180 to 220 °C (5 min), using split mode (1:15). The esters were quantified through standard calibration curve.

#### Organic acids<sup>20</sup>

Nine organic acids (lactic, glycolic, pyruvic, succinic, capric, citramalic, lauric, myristic and palmitic) were analyzed in distilled samples. The methodology was based on the evaporation of 20 mL *cachaça* to dryness at room temperature and the subsequent addition of 200  $\mu$ L derivatizing solution, which contained 100  $\mu$ L of N-methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) and 100  $\mu$ L of nonanoic acid (internal standard, 100 mg L<sup>-1</sup>) in an acetonitrile solution. A Hewlett-Packard 5890 model gas chromatograph (GC) equipped with flame ionization detector (FID) was used with a capillary column DB-5 (5%-phenyl-methylpolysiloxane) with dimensions of 50 m × 0.20 mm × 0.33  $\mu$ m. The oven temperature program used was: 60 °C (2 min) to 100 °C at a programming rate of 25 °C min<sup>-1</sup> and raised at 10 °C min<sup>-1</sup> increments from 100 to 300 °C (5 min), using split mode (1:15). The acids were identified by authentic standard addition method and quantified through standard calibration curve.

#### Multivariate and statistical analysis

Principal components analysis (PCA) reduces the dimensionality of the data through linear combinations of the original independent variables. It allows the maximum amount of variance in the data set to transform into a smaller number of components and allows the data to be observed graphically by grouping samples according to their similarities.<sup>21</sup> For hierarchical cluster analysis (HCA), an agglomerative hierarchical method is used to join the clusters, indicating the level of similarity between them. In this procedure, Ward's linkage method was used to determine the distance between clusters and the Euclidian distance for their amalgamation. Linear discriminant analysis (LDA) is one of the parametric classification methods of pattern recognition that uses linear boundaries to define the groups.<sup>22</sup> The generated LDA calibration model was composed of the 21 samples. The predictive ability of the LDA model was evaluated through cross-validation, using 9 unknown fractions as a test samples.

All the multivariate analysis was performed in the data matrix for the 3 alembic fractions that was structured using all sets of data of Table S1 (Supplementary Information section). The matrix contained 42 rows, which represented the number of *cachaça* samples, and 39 columns, which corresponded to the analyzed chemical variable concentrations. The ND (not detected) in data set matrix was fixed at zero (0.00). As pre-processing, data set in the X-matrix were autoscaled prior to analysis.

Variance analyze (ANOVA) was previously used for the identification of the statistically significant differences

among the secondary compound mean concentration values in the alembic fractions. The ANOVA results were checked using Tukey's test. PCA, HCA, LDA, ANOVA and Tukey's test were applied using Minitab 15.1.1 release software (MINITAB® and the MINITAB logo™ are trademarks of Minitab Inc.). Since many variables are involved along the all production process, the median values were considered more reliable than the mean values for comparison purposes of the three fractions chemical profile.<sup>23</sup>

## Results and Discussion

### Quantitative chemical profile

Figure 1 presents the median values for secondary composition concentration ( $\text{mg L}^{-1}$ ) and alcoholic content ( $\%$ , v/v) for the three spirit fractions produced by alembic distillation. The median concentration values for the 39 organic compounds analyzed in each alembic fraction are presented in Table S2 (Supplementary Information section). The results shown here indicated a quantitative difference for the major compounds analyzed in the three spirit fractions produced during the alembic distillation process.

According to Figure 1, the head fraction presents higher concentration of alcohols (total alcohols in Figure 1) when compared to the other fractions. The methanol, propanol, isobutanol, 1-butanol and isoamyl alcohol is presented higher concentrations in the head fractions. Isoamyl alcohol was the most abundant compound among the alcohols analyzed.

The ethyl acetate and ethyl lactate are the major ester compounds present in *cachaças* according to Nascimento *et al.*<sup>19</sup> All esters, other than ethyl lactate, were found in higher concentrations in the head fraction rather than in the other fractions. Ethyl lactate, the most soluble

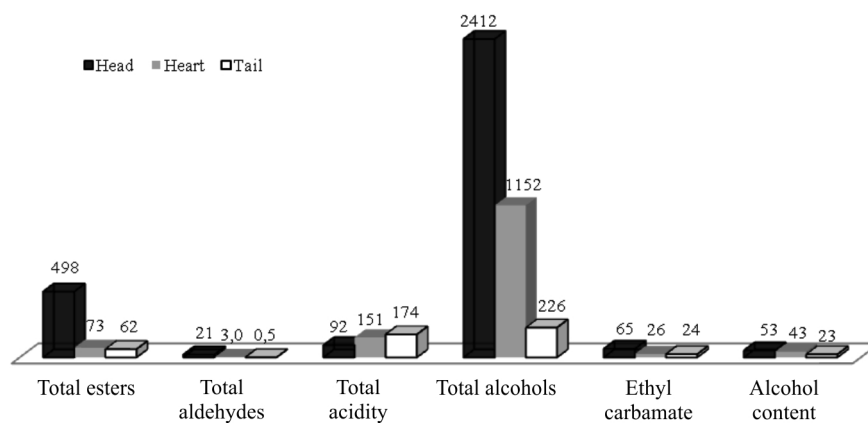
in water among the esters evaluated, was predominant in the heart and tail fraction. Therefore, the increase observed for the total ester concentration in tail fraction in Figure 1 is due to the presence of ethyl lactate.

As can be seen in Figure 1, aldehydes are present in higher concentrations in head fraction. All aldehydes analyzed, except 5-HMF and furfuraldehyde, were predominant in the head fraction, in which acetaldehyde was found to be the major compound. The 5-HMF and furfuraldehyde, which are more soluble in water than in ethanol, differently from the other aldehydes, presented higher concentrations in the heart and tail fractions. The aldehyde acetylacetone was not observed in any sample analyzed.

Ethyl carbamate is more soluble in ethanol than in water and was mainly found in the head fraction (Figure 1). The median values for ethyl carbamate concentrations were  $65.0$ ,  $26.0$  and  $24.0 \mu\text{g L}^{-1}$  for the head, heart and tail fractions, respectively. The results suggest that a tuning in the first cut could be used to reduce the concentration of this unwanted compound in the heart fraction.<sup>24</sup>

The sum of acetic and lactic acid concentrations accounts for more than 90% of the total organic acid content in the alembic fractions. The sum acetic, lactic, glycolic, succinic and citramalic acids, more soluble in water than in ethanol and with boiling points ranging from  $112$ - $235 \text{ }^\circ\text{C}$ , presented higher concentrations in the tail fraction than in the head fraction. The acetic acid in heart fraction presented similar concentration to the tail fraction, showing the difficulty in remove it through the cutting process.

An opposite behavior was observed for capric, lauric, myristic and palmitic acids. Despite its higher boiling points (ranging from  $225$ - $270 \text{ }^\circ\text{C}$ ), regarding to the other acids, the later compounds were predominant in the head fraction since they are more soluble in ethanol than in water. These findings are in similarity with the described by Léauté<sup>11</sup> for cognac.



**Figure 1.** Median concentration ( $\text{mg L}^{-1}$ ) of secondary composition, ethyl carbamate ( $\mu\text{g L}^{-1}$ ) and alcoholic content ( $\%$ , v/v) in the three fractions produced in alembic distillation.

It would be very interesting to correlate the data of sugarcane distilled spirit fractions with the other spirits such as grappa, bagaceira, orujo and rum. However, this is not an easy task since only *cachaça* and rum are sugarcane products, being the others originated from grapes. Unfortunately, at least as far as we know, there are not at the moment available data to perform a reliable comparison. There are few accessible reports which despite well conducted, were not performed in the same experimental conditions and therefore not allow comparisons. Furthermore, besides being originated from diverse raw material, the spirit production process follows different procedures dictated by local practices, including different yeast, fermentation process and distillation apparatus design. The comparison between *cachaça* and rum, considering just the commercial product (heart fraction) points out that despite being sibling spirits, both have their own identity due to the all production process by itself.<sup>25</sup>

Despite of all these difficulties, as example and with information purposes, Table S3 (SI section) exhibits some chemical concentration data for select compounds in *cachaça* and bagaceira for the head, heart and tail fraction and for the heart fraction of orujo, bagaceira, grappa and rum.<sup>15, 25, 26</sup> According to these results, except for capric acid, the concentrations of lauric acid, ethyl acetate, ethyl butyrate, ethyl hexanoate, ethyl octanoate, ethyl decanoate, ethyl dodecanoate, methanol, acetaldehyde, propanol, 1-butanol and 2-butanol on general, independent of the fraction (head, heart and tail) are present in higher concentration in bagaceira than in *cachaça*. Capric acid concentration is only higher in the *cachaça* than in bagaceira when the head fraction is concerned. Similar trend is observed comparing the heart fraction (commercial product) chemical composition of the other spirits. Exceptions for runs (Cuban and non Cuban) which presented the smaller propanol concentration than all the other distilled spirits.

#### ANOVA and multivariate statistical analysis

Before the multivariate statistical analysis were performed, ANOVA test (p-value) was used to observe if there are any significant statistical differences among the secondary compounds mean concentration from the head, heart and tail fractions at the 95% confidence intervals. Tukey's multiple comparison method was used to corroborate with ANOVA test results. Table S4 in the SI section shows the p-value results ( $p < 0.005$ ) for the mean concentration values of each chemical compound for each pair of alembic fractions. The high standard

deviation values observed, which cannot be attributed to experimental analytical errors, strongly suggest that the *cachaça* production process is not uniform by itself. This is not unexpected since in the production there are many independent variables whose strict control is very difficult to assure.

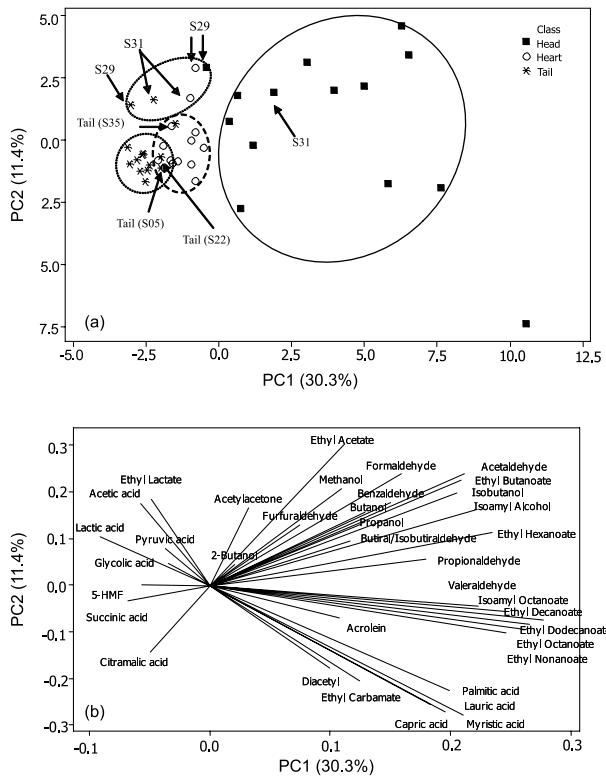
For the thirty nine compounds analyzed, only ethyl decanoate, ethyl octanoate, ethyl nonanoate, isoamyl alcohol and the alcoholic content presented statistical differences when comparing their concentrations among the head, heart and tail. The other remaining compounds presented significant differences only between two of the fractions (see Table S4 in the SI section).

The ANOVA test results here reported would be useful to classify the forty two samples dealt with in this work. However, the production process by itself is not uniform and new variables could be introduced. It would be desirable a more general approach in which not only compounds with significant differences in their concentrations were considered.

PCA and HCA were applied to the database in the Table S1 in the SI section to observe the similarities among the three alembic fractions. The PCA score plot (Figure 2a) shows the clustering between the samples according to the similarities of their chemical composition. The head fraction cluster, despite being more disperse than the other two, does not overlap with the heart fraction. Some overlap was observed between the clusters corresponding to the heart and tail fractions mainly due to the similarity in their acid composition. So, the PCA analysis indicates that the cutting between these two fractions was not so efficient to remove the acidity excess from the heart fraction.

The loading plot (Figure 2b) illustrates the behavior of the 39 analyzed organic compounds in the head, heart and tail fractions. The total variance obtained by the sum of the first seven principal components was 74.8%. The PC1 (30.3%) showed that esters (except ethyl lactate), alcohols, aldehydes (except 5-HMF and furfuraldehyde), ethyl carbamate and fatty acids were the most representative variables responsible for head fraction definition. The acetic, lactic, glycolic, pyruvic, succinic and citramalic acids, ethyl lactate and 5-HMF correlated negatively with the other compounds in PC1, which may be responsible for the partial tail fraction separation.

According to Figure 2a, three tail fractions (S05, S22 and S35) exhibited heart fraction characteristics. From these three fractions, two (S05 and S22) were collected above the recommended alcoholic content level (39.7 and 44.16% (v/v), respectively) and exhibit lactic and acetic acid concentrations below the median concentration



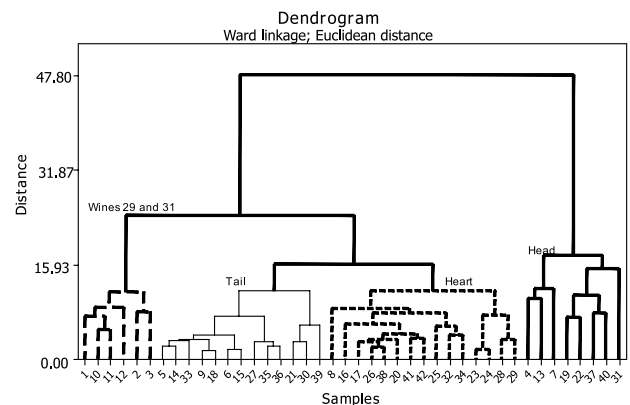
**Figure 2.** PCA of head, heart and tail fractions from alembic distillation. (a) Score plot; (b) loading plot.

values observed for heart fractions (41 and 108 mg L<sup>-1</sup>, respectively). The third tail fraction (S35) was collected at 31% (v/v); however, it presented a lower acid concentration than the expected on the basis of the other samples, and thus exhibiting in these aspect characteristics which resemble more a heart than a tail fraction. Sample S34 from the heart cluster presented higher lactic and acetic acid concentrations (116 and 120 mg L<sup>-1</sup>, respectively) and alcohol content (32.8%, v/v) below the expected value. From the fractions of distillates from wines S31 and S29, only one (head S31) was correctly classified in the head cluster (Figure 2a); the other five define a fourth cluster which mix head, heart and tail fractions. This can be explained considering the composition similarity of fractions regarding to acetic, lactic acid, ethyl acetate and ethyl lactate (Table S1 in the SI section), suggesting that these samples had some problem during the fermentation and distillation process.<sup>19, 27, 28</sup>

The acid concentration may compromise the *cachaças* quality. Indeed, the heart samples (S33 and S34) that are borderline between heart and tail exhibit alcoholic content lower than 38% (v/v) (see Table S1 in the SI section). Therefore, the cutting process during this step should be anticipated before the final alcohol content in the distilled sample reaches 26% (v/v), which is the median value found for the analyzed tail samples (see Table S2 in the SI section).

This procedure will reduce the final product volume, but a balance between quality and productivity should dictate the action of the producers. Thus, the relationship between the cutting process and the sensorial effect of acidity should also be evaluated. As can be observed in Table S1 (SI section) the presence in the heart fraction of compounds that affect the *cachaça* quality as ethyl acetate, methanol, acetaldehyde, ethyl carbamate, propanol (depending on their concentrations) and others with head and tail fraction characteristics can be reduced to reasonably compatible concentrations after the cutting application.

Figure 3 shows the dendrogram analysis for the three distillate fractions based on their chemical profile similarities. As observed in the PCA analysis, the samples from the head did not exhibit similarities with the other two fractions, whereas there was some similarity between the heart and tail fractions. The similarities regarding to the chemical profile for distillates from wines S31 (samples: 1, 2, 3) and S29 (samples: 10, 11, 12) confirm the information obtained from the PCA treatment. The same is true for the heart fractions S33 and S34 (samples 5 and 14, respectively), which were included in the tail sample cluster and for the tail fraction S35 (sample 42) included in the heart fraction cluster.



**Figure 3.** HCA dendrogram for head, heart and tail fractions from alembic distillation.

### Statistical model construction e application to commercial *cachaça* samples

Aiming to contribute in the improvement of the characterization of spirit fractions and consequently the cutting in the alembics distillation, the samples and variables that better represented their respective fractions, chosen from the PCA and HCA plots (Figures 2 and 3) and from ANOVA analysis, were used to build a more representative model. For the model construction, the variables no highly correlated were selected by LDA

analysis. Lactic acid, ethyl acetate, alcoholic content, isoamyl alcohol, capric acid, lauric acid, myristic acid and palmitic acid were used as standard variables. LDA was then applied to this data set to generate a classification model for the three alembic fractions. A data cross-validation was used to avoid the generation of an optimistic model and thus increase its prediction ability.<sup>21, 29</sup> The calibration model to identify head, heart and tail fractions was built up with 21 samples being seven samples for each fraction. According to Table 1, this model presented efficient prediction ability for 90.5% of the samples and was validated with other nine different fractions whose origins were 100% correctly fitted.

In a subsequent step, the self consistency of this model was checked using thirteen high quality commercial *cachaças* (non aged), representative of a “very good heart fraction” whose sensory and chemical qualities have been previously established by a trained panel of sensory analysts and by a group of *cachaça* consumers.<sup>7</sup> Twelve samples (92.3%) were correctly fitted in the heart fraction cluster (Table 1). Only the sample SA was misclassified due to its high isoamyl alcohol concentration (1190 mg L<sup>-1</sup>, Table S5

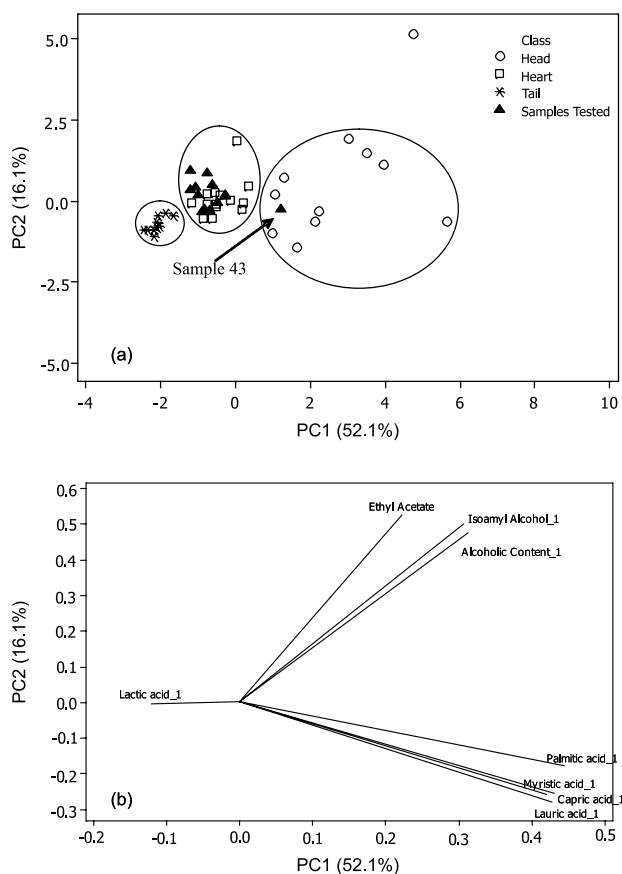
**Table 1.** Classification of brazilian *cachaça* fractions distilled in copper alembics (pot stills) using linear discriminant analysis (LDA)

Model construction	True group		
	Head	Heart	Tail
Total number of samples	7	7	7
Samples with correct classification	6	7	6
Individual correct proportion	0.857	1.0	0.857
Number of samples = 21			
Number of samples correct = 19			
Proportion correct = 90.5%			
Model validation (unknown samples)	True group		
	Head	Heart	Tail
Total number of samples	3	3	3
Samples with correct classification	3	3	3
Individual correct proportion	1.0	1.0	1.0
Number of samples = 9			
Number of samples correct = 9			
Proportion correct = 100%			
Test of model with commercial <i>cachaças</i> (heart fractions)	True group		
	Head	Heart	Tail
Total number of samples	0	13	0
Samples with correct classification	0	12	0
Individual correct proportion	-	0.923	-
Number of samples = 13			
Number of samples correct = 12			
Proportion correct = 92.3%			

Variables: lactic acid, ethyl acetate, alcoholic content, isoamyl alcohol, capric acid, lauric acid, myristic acid, palmitic acid.

in the SI section).

To visualize the fitting of the thirteen high quality commercial samples in the model generated by LDA, PCA plots (Figure 4) were built up using analytical data for lactic acid, ethyl acetate, alcoholic content, isoamyl alcohol, capric acid, lauric acid, myristic acid and palmitic acid for 11 heads, 11 tails and 24 (11+13) heart fractions. It must be recalled however that the head samples (S31, S29, S22), heart fractions (S31, S29, S34, S09) and tail fractions (S35, S08, S05, S22, S29 and S31), which were misclassified due to distillation or fermentation errors, were removed from the PCA analysis. As can be observed in Figure 4, the three clusters were quite well defined and the previous overlap between tail and heart fractions (Figure 2) now was not observed. Table 2 was built up taking into account the well grouped samples in the score plot of Figure 4 and displays for each compound its respective concentration range in each one of the three fractions. These concentration range could be used as a proposition of a guide for the fractions quality evaluation. Therefore, all together, these results strongly suggested that the model generated here is a conceivable pattern reference for the cutting process guidance.



**Figure 4.** PCA of samples and fractions used in the LDA model. (a) Score plot; (b) loading plot.



**Table 2.** Median and range of concentration for organic compounds (mg L<sup>-1</sup>) in head, heart and tail fractions

Compound*	Head fraction		Heart fraction		Tail fraction	
	Median	Range	Median	Range	Median	Range
Ethyl acetate	422	1080-262	63.4	226 -19	26	80-15
Ethyl butanoate	1.14	5.4-0.64	< LOD	0.61- < LOD	< LOD	< LOD
Ethyl hexanoate	2.16	7.2 -1.03	< LOD	1.3-0.53	< LOD	< LOD
Ethyl lactate	13.8	22-3.8	27.3	226-7.2	28.5	270-20
Ethyl octanoate	10.9	42-0.84	1.4	5.6-0.8	< LOD	< LOD
Ethyl nonanoate	0.15	0.74-0.31	< LOD	< LOD	< LOD	< LOD
Ethyl decanoate	16.9	46.8-3.1	3.2	3.1-1.2	< LOD	< LOD
Isoamyl octanoate	0.08	0.51-0.07	< LOD	< LOD	< LOD	< LOD
Ethyl dodecanoate	2.91	7.23-1.25	0.79	6.24-0.32	< LOD	< LOD
Lactic acid	13.5	63.3-1.5	48.3	400-2.14	89.5	174-25
Succinic acid	0.03	0.07-0.16	0.08	0.5-0.08	0.10	1.67-0.05
Capric acid	5.85	54.8-1.1	0.09	1.7-0.09	< LOD	1.1-0.3
Lauric acid	1.06	8.54-0.25	0.15	0.5-0.02	0.08	0.08-0.06
Myristic acid	0.60	1.3-0.30	0.11	0.40-0.04	0.10	0.3-0.1
Palmitic acid	1.97	8.4-1.0	0.61	2.7-0.03	0.35	0.72-0.10
Citramalic acid	< LOD	0.3-0.1	< LOD	0.4-0.05	0.13	0.75-0.13
Acetic acid	49.0	83-35	103	250-24	94.7	328-69
Methanol	< LOD	414-5.3	15.6	22-3.2	12.1	20-8.8
Propanol	232	623-142	187	762-93	49	530-29
Isobutanol	415	554-107	182	307- 50	19	72.6-5.3
Butanol	6.63	17.6-2.8	3.45	8.9-1.87	< LOD	3.4-0.9
Isoamyl alcohol	1814	2070-1174	735	1224-394	70.9	168-20
Ethyl carbamate (ppb)	30	580-65	< LOD	75-52	< LOD	94-21
Alcoholic content	52.9	65-50	41.3	47.4-39	21.6	31-15.5
Acetylacetone	< LOD	0.8-0.0	< LOD	0.34	< LOD	0.04-0.0
Formaldehyde	< LOD	0.8-0.05	< LOD	0.34-0.22	0.01	0.02-0.004
5-HMF	0.12	0.25-0.05	1.21	20.4-0.03	0.08	0.5-0.01
Acetaldehyde	0.08	1.2-0.02	0.5	3.6-0.05	0.2	0.5-0.06
Acrolein	19.3	51-9.5	61.8	236-0.6	0.01	0.01-0.004
Propionaldehyde	0.08	0.44-0.06	0.04	1.37-0.06	0.03	0.05-0.006
Butiral/Isobutiraldehyde	0.19	0.5-0.03	0.02	0.16-0.001	0.03	0.04-0.009
Furfuraldehyde	0.46	3.43-0.16	0.12	4.78-0.004	0.01	0.03-0.02
Benzaldehyde	0.11	1.55-0.06	< LOD	0.15-0.001	< LOD	0.04-0.01
Valeraldehyde	0.4	1.25-0.1	1.72	8.6-0.04	< LOD	< LOD
Diacetyl	0.65	2.64-0.2	< LOD	5.75-0.09	0.01	0.26-0.001

\*Compound selected according to LDA and PCA plot. The heart fraction includes data from the commercial and alembic heart fraction. LOD = Limit of Detection

## Conclusions

The quantitative chemical profile of head, heart and tail fractions of alembic *cachaças* is described using a considerable volume of analytical data. The cutting process markedly influences the alembic sugarcane spirit secondary compound concentrations. The dispersion observed in the PCA and HCA plots of the three alembic fractions showed that the process deserves more attention regarding cutting optimization and standardization. Chemical compounds produced in excess during the fermentation step as ethyl acetate and ethyl carbamate,

which can affect the sensorial and chemical *cachaça* qualities, can be efficiently controlled by the judicious cut application between head and heart fractions through alcoholic content control. The way in which cuts between heart and tail fractions are usually collected, may lead to an increase in acidity in the heart fraction, thus negatively affecting the spirit quality. For the samples that exhibited higher acidity, the cut between heart and tail fractions could be anticipated.

The PCA, HCA data and the LDA model here described showed to be useful tools to discriminate and recognize the alembic fractions conveniently cut and thus could be used to

help on improving the alembic distillation process quality. On principle this procedure could be adapted to alembic distillation of other spirits than *cachaça*.

## Supplementary Information

Complete analytical data for the analyzed samples are available free of charge at <http://jbcs.sbq.org.br> as PDF file.

## Acknowledgments

The authors would like to thank FAPESP, CNPq and CAPES for the financial support and the Professor Dr. Milan Trsic of Departamento de Química e Física Molecular do Instituto de Química de São Carlos (IQSC), Universidade de São Paulo (USP) for his helpful discussions.

## References

1. Brazilian Cachaça Institute; *Cachaça Exportation News*, 2010, [http://www.ibrac.net/index.php?option=com\\_content&view=article&id=47&Itemid=44](http://www.ibrac.net/index.php?option=com_content&view=article&id=47&Itemid=44) accessed in April 2012.
2. Andrade-Sobrinho, L. G.; Boscolo, M.; Lima-Neto, B. S.; Franco, D. W.; *Quim Nova* **2002**, *25*, 1074.
3. Boscolo, M.; Bezerra, C. W. B.; Cardoso, D. R.; Lima-Neto, B. S.; Franco, D. W.; *J. Braz. Chem. Soc.* **2000**, *11*, 86.
4. Nascimento, R. F.; Cardoso, D. R.; Lima-Neto, B. S.; Franco, D. W.; *Chromatographia* **1998**, *48*, 751.
5. Cardoso, D. R.; Bettin, S. M.; Reche, R. V.; Lima-Neto, B. S.; Franco, D. W.; *J. Food Comp. Anal.* **2003**, *16*, 563.
6. De Souza, M. D. C. A.; Vasquez, P.; Del Mastro, L. N.; Acree, T. E.; Lavin, E. H.; *J. Agric. Food Chem.* **2006**, *54*, 485.
7. Seixas, F. R. F.; Da Silva, A. A.; Galinaro, C. A.; Odello, L.; Braceschi, G. P.; Franco, D. W.; *Quim. Nova* **2009**, *32*, 1839.
8. Lima, U. A.; *Aguardente - Fabricação em Pequenas Destilarias*, 1<sup>st</sup> ed.; Fundação de Estudos Agrários Luiz de Queiroz: Piracicaba, 1999.
9. Apostolopoulou, A. A.; Flouros, A. I.; Demertzis, P. G.; Akrida-Demertzi, K.; *Food Control.* **2005**, *16*, 157.
10. Osório, D.; Perez-Correa, J. R.; Biegler, L. T.; Agosin, E.; *J. Agric. Food Chem.* **2005**, *53*, 6326.
11. Léauté, R.; *Am. J. Enol. Vitic.* **1990**, *41*, 90.
12. García-Llobodanin, L.; Achaerandio, I.; Ferrando, M.; Güell, C.; López, F.; *J. Agric. Food Chem.* **2007**, *55*, 3462.
13. Prado-Ramírez, R.; Gonzáles-Alvarez, V.; Pelayo-Ortiz, C.; Casillas, N.; Estarrón, M.; Gómez-Hernandez, H. E.; *Int. J. Food Sci. Technol.* **2005**, *40*, 701.
14. Silva, M. L.; Malcata, F. X.; *Z. Lebensm. Unters. Forsch* **1999**, *208*, 134.
15. Silva, M. L.; Macedo, A. C.; Malcata, F. X.; *Food Sci. Technol. Int.* **2000**, *6*, 285.
16. Cardeal, Z. L.; De Souza, P. P.; Da Silva, M. D. R. G.; Marriott, P. J.; *Talanta* **2008**, *74*, 793.
17. Alcarde, A. R.; Souza, P. A.; Belluco, A. E. S.; *Sci. Agric.* **2010**, *67*, 280.
18. Scanavini, H. F. A.; Ceriani, R.; Cassini, C. E. B.; Souza, E. L. R.; Filho, F. M.; Meirelles, A. J. A.; *J. Food Proc. Eng.* **2010**, *33*, 226.
19. Nascimento, E. S. P.; Cardoso, D. R.; Franco, D. W.; *J. Agric. Food Chem.* **2008**, *56*, 5488.
20. Serafim, F. A. T.; Buchviser, S. F.; Galinaro, C. A.; Novaes, F. V.; Franco, D. W.; *Quim. Nova* **2011**, *34*, 28.
21. Beebe, K. R.; Pell, R. J.; Seasholtz, M. B.; *Chemometrics: A Practical Guide*, 1<sup>st</sup> ed.; John Wiley Sons: New York, 1998.
22. Coomans, D.; Massart, D. L.; *Anal. Chim. Acta* **1979**, *112*, 97.
23. Miller, J. C.; Miller, J. N. In *Statistics for Analytical Chemistry*, 3<sup>rd</sup> ed.; Ellis Horwood PTR Prentice Hall: London, 1993, ch. 6.
24. Nóbrega, I. C. C.; Pereira, J. A. P.; Paiva, J. E.; Lachenmeier, D. W.; *Food Chem.* **2011**, *127*, 1243.
25. 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.; *J. Agric. Food Chem.* **2004**, *52*, 3429.
26. Sampaio, O. M.; Reche, R. V.; Franco, D. W.; *J. Agric. Food Chem.* **2008**, *56*, 1661.
27. Skinner, K. A.; Leathers, T. D.; *J. Ind. Microbiol. Biotechnol.* **2004**, *31*, 401.
28. Narendranath, N. V.; Hynes, S. H.; Thomas, K. C.; Ingledew, W. M.; *Appl. Environ. Microbiol.* **1997**, *63*, 4158.
29. Gonzalez-Arjona, D.; Lopez-Perez, G.; Gonzalez-Gallero, V.; Gonzalez, A. G.; *J. Agric. Food Chem.* **2006**, *54*, 1982.

Submitted: January 18, 2012

Published online: July 19, 2012

FAPESP has sponsored the publication of this article.