Analytical Methods

$UQ^T = TP^T B+E$

$X = USV^T$

PLS-DA vs SIMCA

ROYAL SOCIETY
OF CHEMISTRY

PAPER
R. Augusti *et al.*
Forensic discrimination between authentic and counterfeit perfumes using
paper spray mass spectrometry and multivariate supervised classification

Check for updates

Forensic discrimination between authentic and counterfeit perfumes using paper spray mass spectrometry and multivariate supervised classification†

J. A. R. Teodoro,[a] H. V. Pereira,[a] D. N. Correia,[b] M. M. Sena,[a] E. Piccin[a] and R. Augusti*[a]

Perfumes are cosmetic products with high added value and worldwide consumption, which make them a potential target for counterfeiting. A novel, simple and rapid method was developed for the differentiation of samples of authentic and counterfeit perfumes by employing paper spray mass spectrometry (PS-MS) combined with multivariate supervised classification models: Partial Least Squares Discriminant Analysis (PLS-DA) and Soft Independent Modelling of Class Analogies (SIMCA). Samples of authentic ($n = 29$, consisting of 10 different brands from several batches and from the same producer) and seized counterfeit ($n = 31$) perfumes were analysed by PS-MS in the positive ionization mode and within a mass range of m/z 150–1000. An initial unsupervised exploratory model (Principal Component Analysis – PCA) provided a rough visual separation between the two classes. In contrast, PLS-DA and SIMCA provided good predictions, with low false positive and false negative rates for both models. The interpretation of informative vectors, *i.e.* regression coefficients and Variable Importance in Projection (VIP) scores obtained from the PLS-DA model allowed the detection of diagnostic ions for authentic and counterfeit samples. Some of the most discriminant ions for counterfeit perfumes were suggested to be attributed to compounds with allergenic properties.

Introduction

The word perfume is derived from the Latin *per fumen*. This term means "through smoke", because initially perfumes were manufactured as incense. This denotation also emphasizes the relevance of volatile compounds to the formation of their sensorial properties. Besides fragrances, other additives such as fixatives, solvents and antioxidants are also used for their manufacturing.[1]

There is historical proof indicating the use of perfumes in the Middle West, specifically in Assyria, since biblical times. In ancient Egypt, perfumes were associated with religious ceremonies and also with some distinctive activities of the Pharaohs. In Greece and Rome, technological improvements allowed for the fabrication of large amounts of good quality perfumes. In the Middle Ages, manufacturing techniques

(extraction, mixture and formulation) were developed by alchemists for cosmetic and medicinal applications. Since then, remarkable industrial innovations have expanded the access to perfumes throughout the whole world.[1]

Nowadays, this product has high added value and its industry moves billions of dollars per year.[2,3] For example, the market for global fragrance, the main input in perfume production, was estimated to be worth about US$ 38.8 billion in 2017.[4] In Brazil, the perfumery and cosmetics segment presented an average liquid revenue of almost R$ 43 billion (about US$ 13 billion) in the last few years. Just in taxes, the Brazilian perfumery segment has moved more than US$ 5 billion in 2015.[5] Due to its economic impact, the perfumery industry has become a potential worldwide target for counterfeiting practices.

In 2014, Brazil lost about R$ 100 billion (US$ 35 billion) in injuries caused by such illegal actions.[5] In addition to the financial losses, some forgery products, such as food, cosmetics and medicines, represent a risk to consumers' health because they are not submitted to any quality control. In particular, counterfeit perfumes can contain toxic and/or allergenic compounds.[6–8] Thus, the large amounts of seized fake samples reinforce the need for developing analytical techniques that combine simplicity, reproducibility, efficiency and speed aimed at identifying such illegal products.

[a]*Departamento de Química, Instituto de Ciências Exatas, Universidade Federal Minas Gerais, 31270-901, Belo Horizonte, MG, Brazil. E-mail: augusti.rodinei@gmail.com; Tel: +55 31 3409-5734*

[b]*Superintendência de Polícia Técnica-Científica, Instituto de Criminalística Dr Octávio Eduardo de Brito Alvarenga, 05507-06, São Paulo (SP), Brazil*

Some methods based on different analytical techniques for perfume analysis have been reported. For instance, the quality control and the quantification of allergens in perfumes have been established by applying GC-MS (gas chromatography coupled to mass spectrometry), the most commonly used technique for the analysis of volatile compounds.[9–11] Sensors, such as olfactory electronic systems (EOS)[3] and electronic noses,[2] have also been applied to differentiate between original and counterfeit perfume samples.

Recently, modern mass spectrometry (MS) methods that allow for the attainment of specific chemical profiles (fingerprints) have been used as alternatives to separation techniques, providing the advantages of faster analyses with minimum sample pre-treatment. Perfume fingerprints can be obtained by MS techniques based on atmospheric pressure sources, such as electrospray ionization (ESI-MS),[12] easy ambient sonic-spray ionization (EASI-MS),[6] extractive electrospray ionization (EESI-MS)[13] and electrostatic-spray ionization (ESTASI-MS).[14]

In MS analysis, sample pre-treatment steps (*e.g.*, clean-up and extractions) are laborious and time-consuming. A promising alternative to overcome these drawbacks is paper spray, an ambient ionization method proposed in 2010 by Wang and co-workers.[15] Paper spray mass spectrometry (PS-MS) has been used for the analysis of complex matrices with simplicity and low cost. This ionization methodology is highly versatile and demands minimum sample pre-treatment. PS-MS has been recently applied to the direct analysis of several complex samples, such as in the detection of different types of compounds in biological tissues,[16] determination of contaminants in foodstuffs,[17] quality control of teas,[18] examination of questioned documents,[19–21] discrimination among different bacterial species,[22,23] quantification of therapeutic drugs in blood,[24–26] classification of coffees according to their origin,[27] monitoring of chemical reactions,[28,29] and analysis of samples of forensic interest.[30–32]

For most complex samples, PS-MS fingerprints generate a high amount of multivariate data. The interpretation of this kind of information can be greatly simplified by the application of chemometric methods. These methods allow the extraction of latent information that cannot be evaluated by univariate analysis. Principal Component Analysis (PCA), an exploratory and unsupervised classification method, has been the main approach employed for processing MS data obtained for perfume analysis.[6,12,33] In this type of multivariate method, class information (for instance, counterfeit or authentic) cannot be previously incorporated into the model. Thus, unknown samples can be assigned to a specific group only manually/visually. On the other hand, supervised methods permit classification in an automated and more systematic manner, by validating the model with an independent test set.[34]

Partial Least Squares Discriminant Analysis (PLS-DA) has been the most used method for supervised classification in chemistry. The fundamentals of this method are the simultaneous decomposition and the subsequent correlation between the matrix X that contains the MS data and the class membership vector y, which contains the categorical or dummy variables (in this work, 0 for counterfeit or 1 for authentic samples).[35] Since the predictions provided by PLS-DA are not exactly the values 1 or 0, a Bayesian threshold should be established for the decision of class attribution. Finally, a multivariate analytical validation can be performed for the developed method by estimating appropriate figures of merit.[36,37]

In recent years, some papers have pointed out specific limitations associated with the application of discriminant analysis (*e.g.* PLS-DA) to authentication problems and characterization of counterfeit products (mainly foodstuffs). These suggested limitations are related to inherent difficulties in obtaining an adequate number of representative samples for modelling counterfeit classes, which certainly have an impact on the reliability of predictions and the robustness of the model. Alternative classification methods able to overcome this limitation are denominated one-class (or class modelling) classifiers. Usually, these methods just model the authentic class irrespective of the other classes and as such suspect samples are classified as belonging or not to this specific group. The most applied method for this purpose is SIMCA, Soft Independent Modelling Class Analogies.[38,39]

The basis of SIMCA relies on the construction of independent PCA models for each class of interest, while defining a confidence limit (usually 95%) for the multivariate space delineated by the scores. Analytical data from unknown samples are then inserted into the sampling matrix and the distance between these samples and each clustering class is used to classify them. This method classifies suspect samples as belonging or not to a determined class. The validation of SIMCA models is similar to that of other classification methods, being performed by the calculation of appropriate figures of merit.[34,40]

In the present work, PS-MS was combined with multivariate methods of supervised classification (PLS-DA and SIMCA) to produce a rapid and simple method that is directly applied to a typical forensic issue, *i.e.* discrimination between authentic and counterfeit perfumes. The results from both supervised models were then compared to provide additional elements on the discussion of this recent and controversial subject. In addition, the information arising from the informative vectors achieved from the PLS-DA model was investigated aiming at the detection of the main variables (ions) that most contribute to such discrimination.

Experimental

Materials and samples

HPLC grade methanol was purchased from J. T. Baker Chemicals (Center Valley, USA). Chromatographic grade 1 paper was purchased from Whatman International Ltd. (Maidstone, England). Isosceles triangle pieces (10 × 5 mm) of this chromatographic paper were used as the support for the PS-MS analysis.

All the authentic perfume samples evaluated herein were produced by the Boticário Group (São José dos Pinhais, PR, Brazil). The products from this company are among the main targets of perfume counterfeiting in Brazil because it is the largest perfumery and cosmetics franchise in the world, and the

Table 1 Description of the perfume samples analysed by PS-MS

Brand	Brand codes	Authentic samples	Counterfeit samples
Arbo	AB	3	2
Egeo dolce	ED	3	4
Egeo Woman	EW	3	1
Egeo Man	EM	3	3
Floratta in blue	FB	3	3
Floratta in rose	FR	2	2
Malbec	MB	3	8
Portinari	PT	3	3
Quasar azul	QA	3	4
Quasar vermelho	QV	3	1
TOTAL		29	31

leader of the Brazilian cosmetics market.[41] A total of 31 seized counterfeit perfume samples were provided by the Civil Police of the State of São Paulo (Brazil). A total of 29 samples of authentic perfumes (different brands and batches) produced by the Boticário Group were purchased at local stores, in Belo Horizonte (Brazil). The samples consisted of the same brands for the authentic and counterfeit perfumes. Both original and fake samples were stored in closed flasks at ambient temperature up to the moment of analysis. Table 1 presents the perfume brands and the number of samples analysed per brand.

PS-MS analysis

PS-MS analyses were performed in the positive ion mode on a mass spectrometer, Thermo Scientific LCQ Fleet (San Jose, USA), with an ion trap mass analyser. The following instrumental conditions were used: voltage applied to the paper, +4.5 kV; capillary temperature, 275 °C; capillary voltage, 35 V, and tube lens voltage, 65 V. The mass spectra were acquired in the full scan mode with a mass range of 150–1000 m/z. Each mass spectrum presented refers to an average of 30 scans, with each one requiring 1 s.

The triangular papers were positioned 1.0 cm away from the entrance of the mass spectrometer. They were supported on a mobile platform using a metal clip. An aliquot of 10 μL of perfume was transferred to the central position of the paper and a volume of 30 μL of methanol was applied with the aid of a micropipette. In sequence, the voltage source of the mass spectrometer was turned on to acquire the PS(+)-MS fingerprints of each perfume. The mass spectra were recorded in duplicate and in a randomized order.

Data analysis

The mass spectra were processed using the software Thermo Scientific Xcalibur 2.1.0 (San Jose, USA). PCA, SIMCA and PLS-DA models were built using Mathworks MATLAB 7.9.0.529 (Natick, EUA) combined with PLS toolbox 5.2.2 (Eigenvectors Research Inc., Manson, USA) software.

For each sample, average MS spectra of the duplicates were estimated and normalised by the ion intensity. These mass spectra were arranged in a data matrix X (60 × 851), containing

all the m/z data. This matrix was pre-processed by mean centering and used to build a PCA model. For the PLS-DA model, a vector y (60 × 1) was built containing the class attributions (1 for authentic and 0 for counterfeit perfumes). The y vector was also pre-processed by mean centering. The samples were divided as following: two thirds for the training set (36) and one third for the test set (24), using the Kennard–Stone algorithm for each class.[42] The number of latent variables (LV) for the model was chosen by venetian blinds (6 splits) cross-validation, based on the smallest cross-validation classification error (CVCE).

As for PLS-DA, PS(+)-MS data for SIMCA were divided using the Kennard–Stone algorithm. Hence, the training set was composed of 19 authentic samples from the 10 brands available and 17 counterfeit samples. The data were mean centered and the number of principal components for the authentic class was chosen based on cross-validation. One-class SIMCA was employed. The remaining samples were used to assess the quality of the model. PLS-DA and SIMCA models were validated by estimating the following qualitative figures of merit: sensitivity (true positive rate), specificity (true negative rate), false negative rate (FNR), false positive rate (FPR) and reliability rate (RLR). The area under the receiver characteristic curve (AUROC) was also calculated for the PLS-DA model.[36,37]

Results and discussion

PS(+)-MS fingerprints

Preliminary studies working with the negative ion mode demonstrated poor signal stability. Therefore, all the results and discussions described in the present manuscript pertain to measurements carried out by exclusively focusing on the mass spectra acquired in the positive ion mode. MS profiles (fingerprints) of authentic samples were compared with the respective
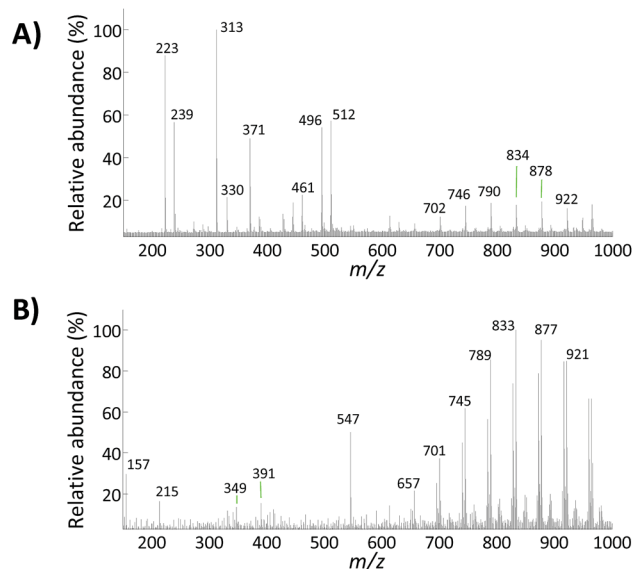

Fig. 1 Examples of PS(+)-MS fingerprints of the Arbo brand perfumes: (A) authentic; (B) counterfeit.

counterfeit perfumes. Fig. 1 illustrates the representative mass spectra of the original and counterfeit perfumes of the brand Arbo (AB). The differences between these two mass spectra are clear as the most intense ions in the mass spectra of the authentic and counterfeit samples do not match.

Hence, in the mass spectrum of the authentic Arbo (AB) perfume (Fig. 1A), prominent ions of m/z 223, 239, 313, 371, 496 and 512 are clearly visualized. Moreover, a series of ions separated by 44 u in the 600–1000 m/z region is also distinctive in this mass spectrum. Fig. 1B illustrates the PS(+)-MS of a counterfeit sample of the Arbo perfume. The most important ions in this mass spectrum are the ones of m/z 157, 215, 349, 391, and 547. Note also the presence of the same series of ions separated by 44 u in the 600–1000 m/z region. Similar comparisons among the PS(+)-MS fingerprints of authentic and counterfeit perfumes were performed for all brands (Table 1). Clear differences among the mass spectra of both types of samples were equally observed for all the perfume brands. The mass spectra of all authentic and counterfeit samples are provided in the ESI, Fig. S1 and S2,† respectively.

The mass spectra of the replicates within the set of samples of a given authentic perfume brand appear to be quite alike in spite of the assorted batches analysed (see the Experimental section for further details). Only minute differences in the relative intensities of the diagnostic ions could be recognised. This finding thus exposes the rigidly controlled conditions employed in all steps involved in the production of the authentic perfumes and demonstrates that the present methodology can be successfully applied to correctly classify a suspect sample. In great contrast to the observed mass spectra of the authentic samples, the PS(+)-MS of the counterfeit perfumes labelled with the same brand name were quite dissimilar. This diversity could be due to a high variability in the manufacturing practices used to produce them, such as different factories, formulations and inputs, in addition to the absence of any quality control guidelines. This can be clearly visualized in Fig. 2 that shows the fingerprints of an authentic sample (Malbec) and three different fake samples with the same nominal brand (Fig. 2A–D, respectively).

Considering the high diversity in the MS profiles of the authentic and counterfeit samples, a systematic method that allows for the unambiguous discrimination between these types of samples is mandatory. In addition, the selection of potential markers (ions with distinct m/z values) to diagnose counterfeit perfumes is also important from the forensic point of view. Thus, multivariate classification models are presented in the next section for discriminating authentic and fake samples encompassing ten different brands (Table 1) from a unique manufacturer (O Boticário). Broadly speaking, this study is conducted by modelling the common variance of such authentic samples in contrast to the usual variance related to counterfeiting (Fig. S1 and S2†). In spite of the fact that some of the counterfeit samples can be directly discriminated by simple spectral comparison, the variability of the fake samples is very large, including samples whose mass spectra are similar to those of authentic samples. As is typical for other types of counterfeit matrices, a broad range of counterfeiting is
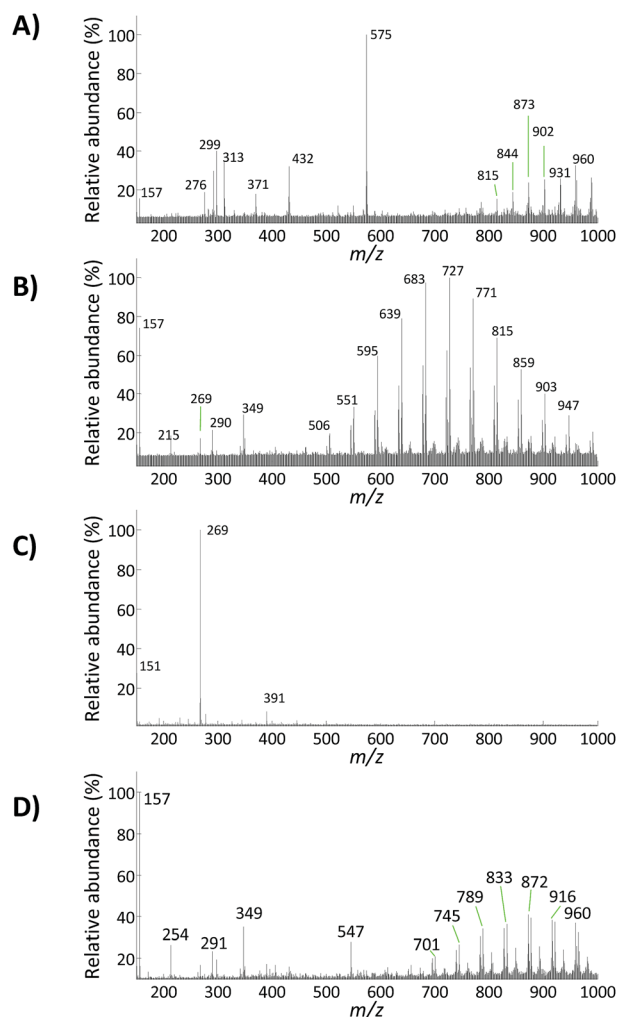

Fig. 2 PS(+)-MS fingerprints of the Malbec brand perfume: (A) authentic; (B–D) three different forgeries.

expected, varying from crude forgeries to high quality fake samples. Therefore, the use of chemometric methods is necessary for the systematic detection of counterfeit perfumes.

Chemometric analysis

In spite of the advantages of supervised methods, it is always advisable to previously build an unsupervised PCA model, searching for natural differences and patterns among samples.[34] Thus, a preliminary PCA model was constructed. The first two principal components (PCs) accounted for 67.09% of the data variance. No significant trend was noted in the remaining PCs. The scores plot of PC1 versus PC2 (Fig. 3A) shows that PC1 was not able to discriminate between authentic and counterfeit samples. However, authentic perfumes (squares) presented a trend of more negative values on PC2, in comparison with those of counterfeit samples (circles). It was also observed that counterfeit samples are more dispersed on PC2 than authentic samples. This trend corroborates the previous observation indicating a large variety of counterfeiting, including samples very different (crude forgeries) from and very similar (high quality
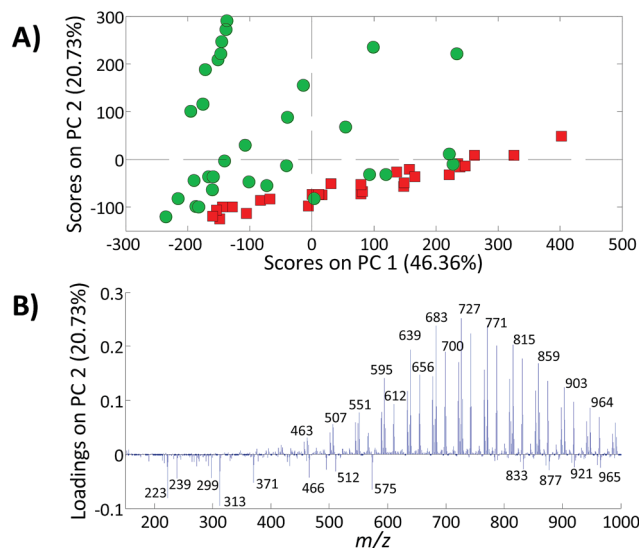
Fig. 3 PCA model built from the PS-(+)-MS data obtained for the perfume samples: (A) scores of PC1 *versus* PC2; (B) loadings of PC2. Green circles and red squares represent counterfeit and authentic perfumes, respectively.

forgeries) to authentic perfumes. Fig. S3 (ESI†), which displays the same PC1 × PC2 plot with the sample scores labelled by brand, indicates that PC1 describes a rough trend of discrimination of the authentic samples as a function of their brands.

Loadings of PC2 (Fig. 3B) point to the variables that are more related with each group of samples. The regular series of ions in the range of *m/z* 600–100 shows positive loadings and is associated with positive scores on PC2 (counterfeit perfumes). On the other hand, the ions of *m/z* 223, 239, 273, 299, 313, 371, 467 and 575 present negative loadings and are mainly associated with the authentic samples. Loadings of PC1 (Fig. S4†) did not contribute to the authentic *versus* counterfeit discrimination.

The best PLS-DA model was built with two latent variables (LV), accounting for 58.52% and 75.86% of the total variance in the *X* and *y* blocks, respectively. Fig. 4A illustrates the predicted values obtained from the PLS-DA model for the training and independent test samples. A Bayesian threshold of 0.628 was estimated and, as can be seen in Fig. 4a, all the samples were correctly predicted.

The figures of merit estimated for this PLS-DA model showed perfect values. The FNR and FPR were both equal to 0%. The sensitivity, selectivity and RLR were all 100%, whereas the AUROC was 1.0. Therefore, this model was able to correctly discriminate the authenticity of samples from all ten different perfume brands in the presence of a variety of counterfeit seized samples.

As previously mentioned in the Introduction, some criticisms have been found in the literature regarding the prevalence of discriminant methods applied to authentication problems.[38,39] These authors have pointed out the inherent difficulties in the acquisition of a sample set representative of all of the possible types of frauds, limiting the representativeness of the modelled adulterated class and the robustness of the model. An alternative to reduce these limitations is performing outlier detection combined with PLS-DA models.[43] As
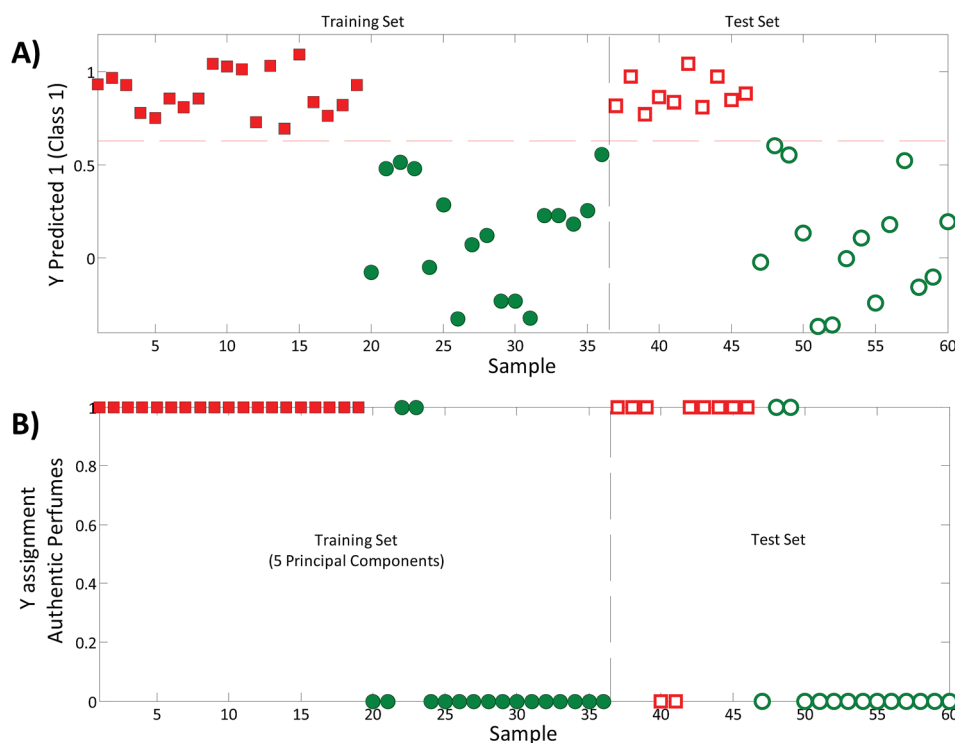

Fig. 4 PLS–DA (A) and SIMCA (B) predictions. Authentic (red squares) and counterfeit (green circles) samples are represented by full symbols for training samples and empty symbols for test samples. The horizontal dashed line indicates the estimated threshold for the PLS-DA model, whereas vertical dashed lines indicate the separation between the training and test sets.

a consequence, future samples very different from the original modelled variance, which cannot be attributed to the pre-defined classes, can be detected. Thus, an outlier detection based on Hotelling's T^2 (variance included in the model) and Q residues (variance left out of the model) was performed for our PLS-DA model. As can be observed in Fig. S5 (ESI†), no outlier was detected at the 95% confidence level. On the other hand, the main alternative for discriminant methods is the use of class modelling methods. In the sequence, a SIMCA model will be presented.

A one-class SIMCA model was constructed for modelling authentic perfumes. The best SIMCA model was built with 5 PCs, accounting for 88.96% of the variance. The results for this model are shown in Fig. 4B, providing only two false positives in the training set, and two false positives and two false negatives in the test set. Analytical validation of this model as a whole was performed by estimating an FPR of 12.9%, an FNR of 6.9%, a sensitivity of 93.1%, a specificity of 87.1% and an RLR of 80.2%. It is interesting to note that all four false positive predictions are related to the counterfeit samples of the same perfume brand (Malbec), while the two false negatives refer to two different brands of authentic perfumes (Floratta in blue and Egeo Man) predicted as counterfeits. Considering the possible use of this methodology for developing a rapid screening method, the purpose of primarily minimizing false negatives should be stressed, since detected positive samples can go further through the conventional reference method (*e.g.*, GC-MS) for confirmation.

Between the two supervised classification models, PLS-DA provided better predictions, with no misclassification. The worse SIMCA results were already expected. As a one-class modelling method, SIMCA only utilized information originating from the authentic samples to predict future classifications. This aspect makes the former model more realistic and robust in relation to new samples of different origins, such as those produced through new types of counterfeiting. In addition, it must be clearly stated that any chemometric/multivariate model is basically a local model. Therefore, our methodology was constructed for detecting counterfeits of this specific company (O Boticário), which is the main target of perfume counterfeiting in Brazil. However, it is possible to expand the scope of this method by including samples of other companies in sufficient number and updating the model.

A few papers have reported the attainment of MS fingerprints of perfumes. Usually the main ions are associated with specific chemical compounds. This has been reinforced by reason of trade secrets in relation to the formulations from the cosmetics industry. Nevertheless, the interpretation of informative vectors obtained from PLS-DA models allowed improving MS spectral attributions taking into account the multivariate structure of the data. Two informative vectors should be emphasised. First, the VIP (variable importance in projection) scores measure the importance of each variable in the projection used by the weights of a particular PLS-DA model, in absolute values.[44] In addition to the VIP scores, a second informative vector should also be inspected, which contains the regression coefficients.[34] For this model, positive regression coefficients were associated
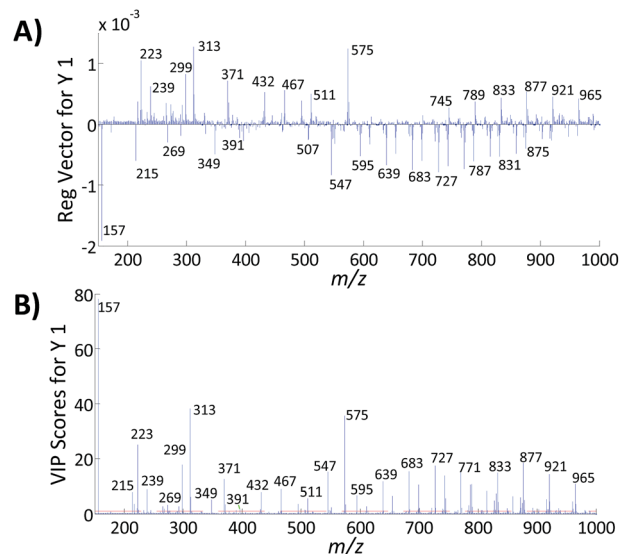

Fig. 5 Informative vectors resulting from PLS-DA (A) regression coefficients; (B) VIP scores.

with authentic perfumes while negative coefficients were related to counterfeit samples (Fig. 5A). Complementarily, variables with VIP scores higher than 1.0 (Fig. 5B) are considered relevant for the class discrimination. The analysis of Fig. 5B indicates that the ion of m/z 157 is the most important for this model, since it presented the highest VIP score. Simultaneously, this ion showed the largest negative regression coefficient thus indicating that it is the most important variable for classifying counterfeit perfumes. The chemical structures of some of the discriminating ions could be attributed by searching the literature and the results are displayed in Table 2.

The most discriminating ion for counterfeit samples may be attributed to citronellol (m/z 157).[11,45] This compound is considered of restricted use in the cosmetics manufacturing industry by the European Union.[51] This same regulation also prohibits the use of the compounds musk ambrette (m/z 269)[46] and DEHP (m/z 391)[9,46] in cosmetic formulations. Note that the ion of m/z 269 was associated with counterfeit perfumes in the model (negative regression coefficient in Fig. 5A).

Damascene (m/z 215),[11] acetylcedrene (m/z 269)[11] and deca-methylcyclopentasiloxane (m/z 371)[47] are compounds recently included in the attention list of the Scientific Committee on Consumer Safety (SCCS).[52] Decamethylcyclopentasiloxane (m/z 371) was associated with authentic samples in accordance with the PCA loadings (Fig. 3B) and PLS-DA regression vector (Fig. 5A). DNOP (m/z 391),[45] although not classified as a banned compound, is on the list of controlled phthalates by the US Environmental Protection Agency (EPA).[53] Benzophenone-2 (m/z 215)[48,49] is a common sun filter used in cosmetic formulations. The Brazilian National Health Surveillance Agency (ANVISA) allows the presence of this substance in cosmetic formulations at a maximum concentration of 10% m m^{-1}.[47]

The ion of m/z 313 was attributed to ethylhexyl methox-ycinnamate, a molecule that acts as a sun filter. This is the most discriminant variable for the classification of authentic

Table 2 Proposed structures of some diagnostic ions detected by PS(+)-MS and highlighted in the informative vectors of the PLS-DA model

m/z	Ion species	Suggested attributions	Perfumery application
157	[M + H]$^+$	Citronellol[11,45]	Fragrance
215	[M + Na]$^+$	Damascene[11]	Fragrance
	[M + Na]$^+$	Dimethyl benzyl carbonyl acetate (DMBCA)[11]	Fragrance
269	[M + Na]$^+$	Benzophenone-2 (ref. 48 and 49)	UV radiation filter
	[M + H]$^+$	2-*tert*-Butyl-4,6-dinitro-5 methylanisole (musk ambrette)[46]	Fragrance
	[M + Na]$^+$	Methyl cedryl ketone (acetylcedrene)[11]	Fragrance precursor
313	[M + Na]$^+$	Ethylhexyl methoxycinnamate[50]	UV radiation filter
371	[M + H]$^+$	Decamethylcyclopentasiloxane[47]	Solvent
391	[M + H]$^+$	Bis(2-ethylhexyl)phthalate (DEHP)[9,46]	Solvent/fixative
	[M + H]$^+$	Di-*n*-octylphthalate (DNOP)[46]	Solvent/fixative

perfumes, since it presented the largest positive regression coefficient and the second highest VIP score (Fig. 5). In general, most of the compounds listed in Table 2 and related to the counterfeit samples could be linked to possible allergic reactions. This observation reinforces the concern about health problems caused by counterfeit perfumes.

Besides the suggestions of attributions shown in Table 2, other ions were important for discrimination. In addition to the ions of m/z 313 and 371, ions of m/z 223, 239, 299, 467 and 575 presented higher VIP scores and larger positive regression coefficients (Fig. 5). Therefore, these ions were directly related to the classification of authentic perfumes and can be considered potential markers or diagnostic ions. This was corroborated by their negative loadings in the PCA model (Fig. 3A). The ions of m/z 349 and 547 presented higher VIP scores and larger negative coefficients. Thus, these two ions can also be considered potential markers of counterfeit perfumes.

The series of ions detectable in the mass range of m/z 600–1000 in the mass spectra of both authentic and counterfeit samples suggests the presence of a polyethoxylated polymer as a constituent of the perfumes. These ions probably refer to sodium/potassium adducts of polyethylene glycol (PEG). The presence of polyethoxylated polymers have been reported in the literature, suggesting an excessive utilization of PEG as an emulsifier in formulations of low quality perfumes.[6] The high intensity of these ions in the mass spectra of the counterfeit perfumes indicates the use of an excessive amount of this raw material during the manufacturing of such types of samples. These findings thus reinforce the complete disinterest of some counterfeiters in manufacturing a product with a minimally desirable quality.

Finally, potential attributions of the most discriminant ions were limited by the utilisation of a low resolution mass spectrometer. For performing more accurate and complete attributions, a high resolution mass spectrometer is definitely required.

Conclusion

A novel and direct method was developed by the combination of PS-MS and multivariate supervised classifications with PLS-DA and SIMCA for the differentiation between authentic and seized counterfeit perfumes. This method was simple and rapid, with the analytical procedure being less than one minute per sample. In addition, no sample pre-treatment was required and insignificant amounts of solvents were consumed. The models were developed exclusively with samples from only one producer (O Boticário), the largest cosmetics group in Brazil, and the main target for perfume counterfeiting. Both models are robust and were generated with 60 samples from 10 different brands and distinct batches. In spite of the great variance included in the models, both chemometric methods were able to yield high rates of correct classification for the samples in the training and test sets.

Another important aspect of this work was the possibility of detecting the ions (variables) that contribute most to the correct classification. It is important to highlight that the most abundant ions in the original MS spectra are not necessarily the most discriminant variables. Diagnostic ions for authentic and counterfeit perfume samples were proposed based on the interpretation of the informative vectors (VIP scores and regression coefficients) obtained from the PLS-DA model. Some of the diagnostic ions for counterfeit perfumes were suggested to be compounds of allergenic potential, reinforcing the health risks for consumers of these counterfeit products. Studies on other perfume brands are underway in our laboratory.

References

1 M. Aftel, *Essence and Alchemy – A Natural History of Perfume*, Gibbs Smith Publishers, Layton, 2004.
2 J. Gebicki, B. Szulczynski and M. Kaminski, *Meas. Sci. Technol.*, 2015, **26**, 125103.
3 M. Cano, V. Borrego, J. Roales, J. Idigoras, T. Lopes-Costa, P. Mendoza and J. M. Pedrosa, *Sens. Actuators, B*, 2011, **156**, 319–324.

4 Yahoo Finance, Global fragrances industry to be worth $38.8 billion by 2017, https://uk.finance.yahoo.com/news/global-fragrances-industry-worth-38-000000901.html, accessed February 2017.

5 Associação Brasileira da Indústria de Higiene Pessoal Perfumaria e Cosméticos, Anuário ABIHPEC 2015, http://abihpec.org.br/anuario-2015/, accessed February 2017.

6 R. Haddad, R. R. Catharino, L. A. Marques and M. N. Eberlin, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 3662–3666.

7 Fox News, Counterfeit fragrances on the rise, could be hazardous to health, 2014, http://www.foxnews.com/health/2014/02/06/counterfeit-fragrances-on-rise-could-be-hazardous-to-health.html, accessed February 2017.

8 R. Parascandola, *Five busted in a multi-million dollar counterfeit perfume sting*, Daily News, New York, USA, 2016, http://www.nydailynews.com/new-york/nyc-crime/busted-multi-million-dollar-counterfeit-perfume-sting-article-1.2650216, accessed February 2017.

9 M. Lopez-Nogueroles, J. L. Benede, A. Chisvert and A. Salvador, *Anal. Methods*, 2013, **5**, 409–415.

10 C. Debonneville and A. Chaintreau, *J. Chromatogr. A*, 2004, **1027**, 109–115.

11 A. Rey, E. Corbi, C. Peres and N. David, *J. Chromatogr. A*, 2015, **1404**, 95–103.

12 L. A. Marques, R. R. Catharino, R. E. Bruns and M. N. Eberlin, *Rapid Commun. Mass Spectrom.*, 2006, **20**, 3654–3658.

13 K. Chingin, G. Gamez, H. Chen, L. Zhu and R. Zenobi, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 2009–2014.

14 E. Tobolkina, L. Qiao, G. Xu and H. H. Girault, *Rapid Commun. Mass Spectrom.*, 2013, **27**, 2310–2316.

15 J. J. Liu, H. Wang, N. E. Manicke, J. M. Lin, R. G. Cooks and Z. Ouyang, *Anal. Chem.*, 2010, **82**, 2463–2471.

16 H. Wang, N. E. Manicke, Q. A. Yang, L. X. Zheng, R. Y. Shi, R. G. Cooks and O. Y. Zheng, *Anal. Chem.*, 2011, **83**, 1197–1201.

17 Z. P. Zhang, R. G. Cooks and Z. Ouyang, *Analyst*, 2012, **137**, 2556–2558.

18 J. Deng and Y. Yang, *Anal. Chim. Acta*, 2013, **785**, 82–90.

19 M. Jurisch and R. Augusti, *Anal. Methods*, 2016, **8**, 4543–4546.

20 P. S. Ferreira, D. F. A. Silva, R. Augusti and E. Piccin, *Analyst*, 2015, **140**, 811–819.

21 V. S. Amador, H. V. Pereira, M. M. Sena, R. Augusti and E. Piccin, *J. Am. Soc. Mass Spectrom.*, 2017, DOI: 10.1007/S13361-017-1686-Z.

22 A. M. Hamid, A. K. Jarmusch, V. Pirro, D. H. Pincus, B. G. Clay, G. Gervasi and R. G. Cooks, *Anal. Chem.*, 2014, **86**, 7500–7507.

23 A. M. Hamid, P. Wei, A. K. Jarmusch, V. Pirro and R. G. Cooks, *Int. J. Mass Spectrom.*, 2015, **378**, 288–293.

24 N. E. Manicke, Q. Yang, H. Wang, S. Oradu, Z. Ouyang and R. G. Cooks, *Int. J. Mass Spectrom.*, 2011, **300**, 123–129.

25 N. E. Manicke, P. Abu-Rabie, N. Spooner, Z. Ouyang and R. G. Cooks, *J. Am. Soc. Mass Spectrom.*, 2011, **22**, 1501–1507.

26 R.-Z. Shi, E. T. M. El Gierari, N. E. Manicke and J. D. Faix, *Clin. Chim. Acta*, 2015, **441**, 99–104.

27 R. Garrett, C. M. Rezende and D. R. Ifa, *Anal. Methods*, 2013, **5**, 5944–5948.

28 C. C. A. Paula, A. Valadares, M. Jurisch, E. Piccin and R. Augusti, *Rapid Commun. Mass Spectrom.*, 2016, **30**, 1176–1180.

29 S. F. Resende, J. A. R. Teodoro, I. Binatti, R. L. Gouveia, B. S. Oliveira and R. Augusti, *Int. J. Mass Spectrom.*, 2017, DOI: 10.1016/j.ijms.2016.10.019.

30 H. V. Pereira, V. S. Amador, M. M. Sena, R. Augusti and E. Piccin, *Anal. Chim. Acta*, 2016, **940**, 104–112.

31 C. C. A. Paula, R. A. Lordeiro, E. Piccin and R. Augusti, *Anal. Methods*, 2015, **7**, 9145–9149.

32 R. D. Espy, S. F. Teunissen, N. E. Manicke, Y. Ren, Z. Ouyang, A. van Asten and R. G. Cooks, *Anal. Chem.*, 2014, **86**, 7712–7718.

33 M. Asadollahi-Baboli and A. Aghakhanib, *Anal. Methods*, 2014, **6**, 4279–4287.

34 F. Marini, *Chemometrics in Food Chemistry*, Elsevier, Amsterdam, 2013.

35 R. G. Brereton and G. R. Lloyd, *J. Chemom.*, 2014, **28**, 213–225.

36 M. Isabel López, M. Pilar Callao and I. Ruisánchez, *Anal. Chim. Acta*, 2015, **891**, 62–72.

37 B. G. Botelho, N. Reis, L. S. Oliveira and M. M. Sena, *Food Chem.*, 2015, **181**, 31–37.

38 P. Olivieri and G. Downey, *TrAC, Trends Anal. Chem.*, 2012, **35**, 74–86.

39 O. X. Rodionova, A. V. Titova and A. L. Pomerantsev, *TrAC, Trends Anal. Chem.*, 2016, **78**, 17–22.

40 S. Wold, *Pattern Recognit.*, 1976, **8**, 127–139.

41 O Boticário, http://www.grupoboticario.com.br/pt/grupo-boticario/Paginas/Inicial.aspx, accessed February 2017.

42 R. W. Kennard and L. A. Stone, *Technometrics*, 1969, **11**, 137–148.

43 A. G. Martins, M. Talhavini, M. L. Vieira, J. J. Zacca and J. W. B. Braga, *Food Chem.*, 2017, **229**, 142–151.

44 T. N. Tran, N. L. Afanador, L. M. C. Buydens and L. Blanchet, *Chemom. Intell. Lab. Syst.*, 2014, **138**, 153–160.

45 B. Desmedt, M. Canfyn, M. Pype, S. Baudewyns, V. Hanot, P. Courselle, J. O. De Beer, V. Rogiers, K. De Paepe and E. Deconinck, *Talanta*, 2015, **131**, 444–451.

46 L. Sanchez-Prado, M. Llompart, J. Pablo Lamas, C. Garcia-Jares and M. Lores, *Talanta*, 2011, **85**, 370–379.

47 Y. Horii and K. Kannan, *Arch. Environ. Contam. Toxicol.*, 2008, **55**, 701–710.

48 Agência Nacional de Vigilância Sanitária (ANVISA), Resolução RDC no. 79, de 28 de agosto de 2000, Brazil, http://www.anvisa.gov.br/cosmeticos/guia/html/79_2000.pdf, accessed February 2017.

49 A. Gimenez-Arnau, E. Gimenez-Arnau, E. Serra-Baldrich, J. P. Lepoittevin and J. G. Camarasa, *Contact Dermatitis*, 2002, **47**, 345–352.

50 W. Uter, M. Gonçalo, K. Yazar, E. Kratz, G. Mildau and C. Lidén, *Contact Dermatitis*, 2014, **71**, 162–169.

51 The European Parliament and the council on cosmetic products, Regulation (EC) No. 1223/2009, *Official Journal of the European Union*, 2009, **59**, 59–209.

52 The Scientific committee on cosmetic products and non-food products (SCCNFP), The 1ST update of the inventory of ingredients employed in cosmetics products. Section II: Perfume and aromatic raw materials, 2000, http://www.leffingwell.com/cosmetics/out131_en.pdf, accessed February 2017.

53 U.S. Environmental Protection Agency (EPA), Phthalates Action Plan, 2012, http://www.epa.gov/sites/production/files/2015-09/documents/phthalates_actionplan_revised_2012-03-14.pdf, accessed February 2017.