



# Anthocyanins as markers for the classification of Argentinean wines according to botanical and geographical origin. Chemometric modeling of liquid chromatography–mass spectrometry data



Pablo L. Pisano<sup>a,\*</sup>, María F. Silva<sup>b</sup>, Alejandro C. Olivieri<sup>a</sup>

<sup>a</sup>Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Instituto de Química Rosario (IQUIR-CONICET), Suipacha 531, Rosario S2002LRK, Argentina

<sup>b</sup>Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo, Instituto de Biología Agrícola de Mendoza (IBAM-CONICET), Alte. Brown 500, Chacras de Coria, 5505 Mendoza, Argentina

## ARTICLE INFO

### Article history:

Received 8 August 2014

Received in revised form 20 October 2014

Accepted 20 November 2014

Available online 27 November 2014

### Keywords:

Liquid chromatography–mass spectrometry

Multivariate curve resolution

Discriminant partial least-squares

Wine discrimination

## ABSTRACT

A study of Argentinean red wines was performed by direct injection of untreated wine samples into a liquid chromatography–mass spectrometry system, processing the collected data with two chemometric algorithms: multivariate curve resolution with alternating least-squares and discriminant unfolded partial least-squares (D-UPLS). The objectives were: (1) the chemometric resolution of profiles in the modes represented by elution time and  $m/z$  ratio, (2) the discrimination of samples according to *varietal* and/or geographical origin, and (3) the identification of key compounds helping to perform sample discrimination. The results indicate that all wine *varietals* can be adequately discriminated, and also three wine producing regions (located in the east, south and north of the Cuyo region) were differentiated from the remaining regions. The applied chemometric models allowed the tentative identification of anthocyanin compounds as responsible for both type of discriminations, in the case of D-UPLS by employing the concept of variables importance in the projection.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

Phenolic compounds, also known as polyphenols, constitute an essential part of the human diet. These compounds have been reported to display multiple biological health-promoting properties, such as antioxidant, lower risk for certain cancers, urinary tract health, improved memory, and normal ageing. These findings strongly suggest that polyphenols are absorbed and display several physiological activities and health benefits (Nile & Park, 2014). On the other hand, phenolic composition is one of the most important quality parameters of wines, and depends on both variety and *terroir*. The French term *terroir* is used in oenology to define a geographical and environmental origin where the grapes used for the vintage were grown. This term has been used worldwide to describe the influence of the climate, soil, and human intervention on the chemical and metabolic wines composition (Ribéreau-Gayon, Glories, Maujean, & Dubourdieu, 2006).

Anthocyanins are a subgroup of polyphenols that are commonly found in grapes, and are the original red coloring molecules in red

wine. They are very reactive and their transformation into anthocyanin-derived pigments (more complex oligomers, *e.g.*, pyranoanthocyanins) begins as soon as they are extracted to fermenting must. Therefore, the final anthocyanin red wine composition is the result of the combination of many kinds of anthocyanins and anthocyanin-derived pigments. Major anthocyanins present in *Vitis vinifera* grapes are the 3-monoglucosides of malvidin, cyanidin, petunidin, peonidin, and delphinidin; being malvidin-3-*O*-glucoside and its acylated esters the most dominant anthocyanin present in red wine (Fig. 1) (Blanco-Vega, Gómez-Alonso, & Hermosín-Gutiérrez, 2014).

Since only a few reference standards are available for anthocyanins or anthocyanin-related pigments derived from wine, a detailed description of red wine pigments seems to be still an analytical challenge, as it involves at least the previous fractionation of different kinds of pigments and the use of high-performance liquid chromatography with tandem mass spectrometric detection (HPLC-MS/MS) for the accurate MS/MS fragmentation analysis of separated fractions. In addition, HPLC analysis of red wine pigments based on the use of C18 and other reversed-phase stationary phases only offers good separations for anthocyanins and low molecular weight anthocyanin-derived pigments, whereas the

\* Corresponding author. Tel./fax: +54 341 4372704.

e-mail address: [pisano@iquir-conicet.gov.ar](mailto:pisano@iquir-conicet.gov.ar) (P.L. Pisano).

polymeric pigments appear as an unresolved broad peak; thereby, identification of these compounds is often tentative and is usually based on the combination of reversed phase-LC (RP-LC) elution order, UV-visible spectra and mass spectral information (Abad-García, Berrueta, Garmón-Lobato, Gallo, & Vicente, 2009; Alberts, Stander, & de Villiers, 2012).

Although anthocyanins have been postulated as chemical markers to differentiate grape cultivars and red wine provenance, in most reported works using elution profiles of hyphenated techniques [LC-diode array detection (LC-DAD), LC-MS, gas chromatography-MS (GC-MS), capillary electrophoresis (CE), etc.] as a fingerprint for wine classification, it is rare to find works that are able to identify the key compounds within that profile that contributed most to achieve the classification (Alves, Nascimento, & Nogueira, 2005; Ballabio, Skov, Leardi, & Bro, 2008; Cuadros-Inostroza et al., 2010; de Villiers, Vanhoenacker, Majek, & Sandra, 2004; Fanzone, Peña-Neira, Jofré, Assof, & Zamora, 2010; Garrido-Delgado, López-Vidal, Arce, & Valcárcel, 2009; Jaitz et al., 2010; Pazourek et al., 2005; Tredoux et al., 2008; Vaclavik, Lacina, Hajslova, & Zweigenbaum, 2011). This is true in part due to compounds identification being beyond the scope of the work when classification is the principal objective; and second and mostly because it is very difficult identifying a particular compound in a complex matrix as wine when no pre-treatment was applied to the sample.

Previously in our working group we performed an exploration of Argentinean red wines by direct injection HPLC-DAD without sample pre-treatment coupled to multivariate curve resolution-alternating least-squares (MCR-ALS) as a chemometric data processing algorithm (Pisano, Silva, & Olivieri, 2014). We also attempted wine classification by grape *varietal* and geographical origin, achieving Malbec *varietal* discrimination from the remaining ones, and partial success in discriminating samples according to their geographical origin. Furthermore, a thorough analysis of the results showed that anthocyanin compounds present in wine were crucial to perform both discriminations.

Encouraged by these results, we carried out a similar study, using a more selective chromatographic detection technique such as mass spectrometry, with the objective of achieving better resolution of chromatographic and *m/z* profiles, attempting classification of the studied wine samples, and identifying key compounds that help to discriminate samples. Nowadays, the use of LC-MS assisted with chemometric algorithms as a strategy both for classification or quantification seems to be a promising methodology (Bertoncelj, Polak, Kropf, Korošec, & Golob, 2011; Gu et al., 2014; Kulakowski, Wu, Balick, & Kennelly, 2014; Taamalli, Arráez Román, Zarrouk, Segura-Carretero, & Fernández-Gutiérrez, 2012; Xia, Zhou, Tan, Ding, & Xu, 2014). Specifically, in this study we directly injected the wine samples into the LC-MS system, and processed the data with two chemometric algorithms: multivariate curve resolution with alternating least-squares (MCR-ALS) (Tauler, 1995), and unfolded partial least-squares (UPLS) (Wold, Geladi, Esbensen, & Öhman, 1987) employed in its discriminant mode D-UPLS (Arancibia, Boschetti, Olivieri, & Escandar, 2008). The models were designed to discriminate the samples according

to *varietal* and geographical origin, respectively. The results of this data exploration indicate that: (1) with MCR-ALS algorithm all wine *varietals* can be adequately discriminated from each other and D-UPLS models adequately discriminated Malbec, Merlot, and Cabernet Sauvignon *varietals* from the remaining samples; and (2) three geographical regions were differentiated from the remaining regions with both chemometric algorithms. Furthermore, analysis of the specific MCR-ALS results, and the so-called variables importance in the projection (VIP) of the D-UPLS models allowed the tentative identification of anthocyanin compounds as responsible for both types of discrimination.

## 2. Experimental section

### 2.1. Reagents and standards

LC-MS grade water and acetonitrile were purchased from Carlo Erba (Milan, Italy), and formic acid from Sigma-Aldrich (Buenos Aires, Argentina) was LC-MS grade and used directly.

### 2.2. Wine samples

The wine samples were obtained from red grapes of *V. Vinifera* L. of eight varieties [Aspiran (A), Bonarda (B), Cabernet Sauvignon (C), Malbec (Ma), Merlot (Me), Sangiovese (Sa), Syrah (Sy) and Tempranillo (T)], harvested in 2012 from thirteen collaborating wineries of Mendoza and San Juan (Argentina), including an experimental winery from Facultad de Ciencias Agrarias (FCA), Universidad Nacional de Cuyo, Mendoza, Argentina. The thirteen wineries were: Galán (east Mendoza, A, B, C, Ma, Me, T), CoViTu (Valle de Uco, south Mendoza city, B, C, Ma, Me, T), experimental winery FCA (Luján, C, Ma, Me), San Rafael (south Mendoza, Ma, Sy), Agrelo (Ma, Me, Sa), San Juan (north Mendoza, Cs, two samples of Ma, Sy), Mayor Drummond (Cs), La Consulta (Sy), Plantago (Ma), and Albahaca (Ma). The wine samples from each winery were collected directly from fermentation tanks at the end of malolactic fermentation, transferred under nitrogen to completely filled amber glass bottles, and stored at 4 °C to ensure their preservation until their analysis in the laboratory.

### 2.3. LC operating conditions

The optimization of HPLC method was based on the work developed by De Villiers et al. (Alberts et al., 2012; De Villiers et al., 2004). Prior to analysis, wine samples were filtered through a 0.22- $\mu$ m pore size nylon membrane (Gamafil, Buenos Aires, Argentina) without further treatment, and 5  $\mu$ L of every sample were injected directly into the chromatographic system, consisting of an Agilent 1200 series HPLC equipped with a binary pump model G1312B SL, an autosampler model G1367D SL + WP, a thermostated column compartment model G1316B, and a variable wavelength detector G1314C (Agilent Technologies, Santa Clara, CA). Separation was performed on a reversed-phase column Zorbax Eclipse XDB-C18 (50 mm  $\times$  3 mm, 1.8  $\mu$ m particle size; Agilent Technologies) at 25 °C. Two mobile phases were employed for elution: **A** (water/formic acid, 99.9:0.1, v/v) and **B** (acetonitrile/formic acid, 99.9:0.1, v/v), and the gradient profile was as follows: 3% **B** (0 min); 20% **B** (10 min); 30% **B** (15 min); 50% **B** (20 min); 95% **B** (25 min); 95% **B** (27 min); and 3% **B** (30 min). The flow rate was 0.2 mL min<sup>-1</sup>. Each sample was run in triplicate, and good repeatability was observed. No changes were detected in chromatographic parameters, such as retention time, and peak shapes and areas in a reference sample that was run at the beginning and at the end of the analysis. Wavelength detector was set to 280 nm to monitor most phenolic compounds. All the analyses were

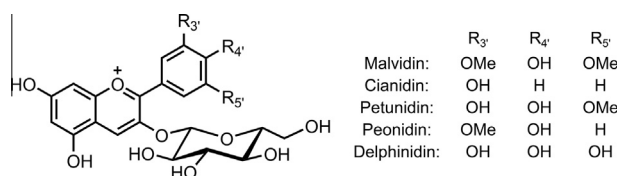


Fig. 1. Structures of the major anthocyanidin-3-O-glucosides present in *Vitis vinifera* grapes.

conducted with the same column, keeping the maximum working pressure in the range 100–110 bar, 600 bar being the maximum recommended working pressure for the column used in this study.

#### 2.4. MS operating conditions

After passing through the flow cell of the HPLC wavelength detector, the column eluate was directed to the MS system, consisting of a Bruker micrOTOF-Q II time-of-flight mass spectrometer coupled with an Apollo II ion funnel ESI Electropray source and controlled by Data Analysis software (Bruker Daltonics, Billerica, MA). Operating conditions of the ESI interface in the positive-ion mode were: collision energy 8 eV, drying gas ( $N_2$ ) temperature 180 °C, drying gas flow 4.0 L min<sup>-1</sup>, nebuliser gas ( $N_2$ ) pressure 1.5 bar and capillary voltage 4500 V. MS spectra were obtained by scanning  $m/z$  ratios from 50 to 1500.

### 3. Chemometric models and software

#### 3.1. Multivariate curve resolution – alternating least-squares

Multivariate curve resolution refers to a group of techniques which intend the recovery of the pure response profiles (UV–visible or mass spectral profiles, elution profiles, etc.) of the chemical constituents (variables) of an unresolved mixture. To process three-way data (e.g., second-order LC–MS data for several samples), the MCR-ALS algorithm builds an augmented data matrix by placing all individual sample matrices adjacent to each other in a column-wise augmentation mode. This allows one to model, via suitable constraints during the fitting phase, the varying time profiles of the sample components in the various samples. For description of the model and further details, see [Supplementary Data](#).

#### 3.2. Discriminant unfolded partial least-squares

Partial least-squares is a regression technique which finds a regression model by projecting the predicted variables (Y) and the observable variables (X) to a new space. PLS regression has proven to be a very versatile method for multivariate data analysis (Escandar et al., 2007). Discriminant partial least-squares (D-PLS) is a variant of PLS in which the predicted variables (Y) are categorical values or codes, allowing the separation of samples into different classes. When working with second-order data, a useful alternative is to rearrange them into vectors and apply a first-order PLS algorithm for discrimination (Olivieri & Escandar, 2014), leading to D-UPLS (see [Supplementary Data](#) for more details).

#### 3.3. Software and data pre-processing

All calculations were made using MATLAB (version 7.0; The Mathworks Inc., Natick, MA). MCR-ALS was implemented using the graphical interface provided by Tauler on his web page (<http://www.mcrals.info>; Jaumot, Gargallo, de Juan, & Tauler, 2005). Principal component analysis (PCA) was run using an in-house MATLAB code. The D-UPLS analysis was performed with MVC2, a new version of the MATLAB toolbox already reported in the literature (Olivieri, Wu, & Yu, 2009) and available at [www.iquir-conicet.gov.ar/descargas/mvc2.rar](http://www.iquir-conicet.gov.ar/descargas/mvc2.rar). All programs were run on an HP Pavilion dv5-2043la microcomputer with an Intel Pentium P6000, 1.86 GHz microprocessor and 6 GB of RAM. LC–MS data (files of 350 MB approximately for each sample) were exported as ChemStation file extension (25 MB approximately for each sample) with Data Analysis software and re-opened with OpenChrom® software (Community Edition 0.8.0 Dempster) in order to reduce

the size of the original data before chemometric analysis. OpenChrom is an open-source software for handling chromatography and mass spectrometry data files without prior conversion. LC–MS data files were exported from OpenChrom to finally obtain .csv files (comma separated values, ca. 13 Mb for each sample) for subsequent data processing under MATLAB. Therefore, each sample subjected to analysis consisted of an array of 133 × 1001 data points (3–25 min taken in steps of 10 s and  $m/z$  100–1100 taken in steps of 1 amu, respectively).

### 4. Results and discussion

#### 4.1. General considerations

Grape native anthocyanins are the original red coloring molecules in red wine, but they are very reactive and their transformation into anthocyanin-derived pigments begins as soon as they are extracted to fermenting must. In this sense, pyranoanthocyanins are formed as early as during alcoholic fermentation, and also over the ageing of the wine. In this work, therefore, we have performed the analysis of wine samples collected directly from fermentation tanks at the end of malolactic fermentation and prior to the bottling process, to avoid the natural condensation of anthocyanins to more complex polymers during the wine ageing, which would result in an even more complicated analysis.

#### 4.2. MCR-ALS analysis

Considering the complexity of the raw analytical data obtained by LC–MS measurements of wine samples without any experimental pre-treatment, visual inspection of this type of data sets would be considerably time-demanding and inefficient. As an example, Fig. 2 shows the three-dimensional surface obtained by ESI in the positive ion mode for a specific sample (Malbec varietal, CoViTu winery from Valle de Uco, south Mendoza city) after injection into the LC–MS system. Multiple signals can be generated during ESI ionization of a single compound and, as a consequence, extremely large amounts of data are obtained when multiple metabolites are simultaneously fragmented, as in wine metabolomic analysis. Therefore, addressing this problem requires the selection of suitable data processing algorithms to extract patterns that allow the resolution of profiles in terms of their chromatograms and mass spectra.

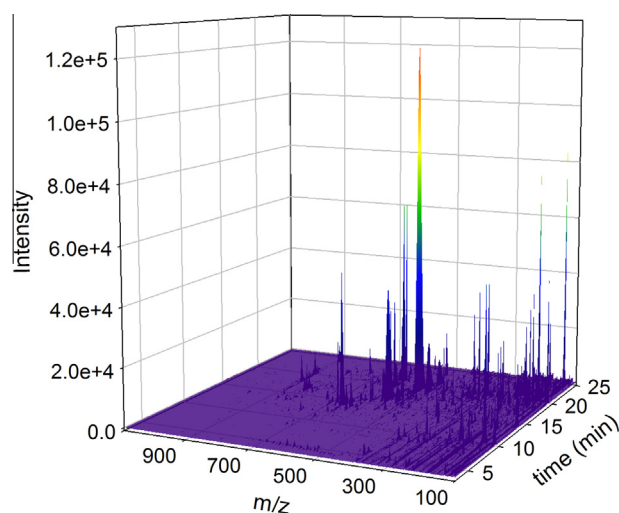


Fig. 2. A typical liquid chromatographic-mass spectrum landscape of a red wine sample.

In a first phase, MCR-ALS was applied to the LC–MS data measured for the studied wine samples. To do so, data matrices from each sample were organized by joining the elution time–spectral data matrices on top of each other sharing the spectral subspace (i.e., by column-wise augmentation), creating the so-called augmented data matrix before MCR-ALS decomposition. Subsequently, MCR-ALS analysis was applied to the column-wise augmented data matrix, namely, an array of  $3591 \times 1001$  data points, as explained in Sections 3.1 and 3.3. The number of components was estimated by principal component analysis of the augmented data matrix, inspecting a plot of singular values as a function of increasing number of trial components (Jolliffe, 2002). In this way, 20 components were selected, which explained 93.02% of the data variance. In order to achieve successful resolution, non-negativity in both spectra and chromatograms was applied during the least-squares fit, until successive changes in residual fit were smaller than 0.1%. This typically required 25 iterations. In addition to the resolved spectral and mass profiles, MCR-ALS resolution of the LC–MS data renders the area under the resolved chromatographic profile for each component in a particular sample, i.e., the so-called MCR-ALS scores. This resolution was obtained with good quality parameters, namely, fitting error (L.O.F.) of 4.86% and 6.90% (regarding PCA and experimental, respectively) and 93.09% of explained variance.

The fingerprint information obtained by the MCR-ALS resolution was arranged into a matrix of size  $27 \times 20$  (27 samples and 20 constituent scores). This latter matrix was submitted to principal component analysis (PCA) for discrimination purposes, in order to study the relationship among the MCR-ALS fingerprint

information with the eight different *varietals* and the geographical origin of the wine samples. Fig. 3A shows the score plot of first vs. second principal component (67.03% and 15.86% of variance retained by PC1 and PC2, respectively). In this figure, we can observe the successful discrimination between the eight different *varietals* from each other. On the other hand, Fig. 3B shows a plot of first vs. third principal component (PC3, 6.05% of explained variance) in which we can observe discrimination by geographical origin of the samples corresponding to San Juan (North Mendoza), and San Rafael (South Mendoza) from the rest of samples. Therefore, MCR-ALS resolution of the LC–MS data, and subsequent processing with PCA allows discrimination of all *varietal* wine samples, and also the discrimination of some samples corresponding to northern and southern regions of the Mendoza province.

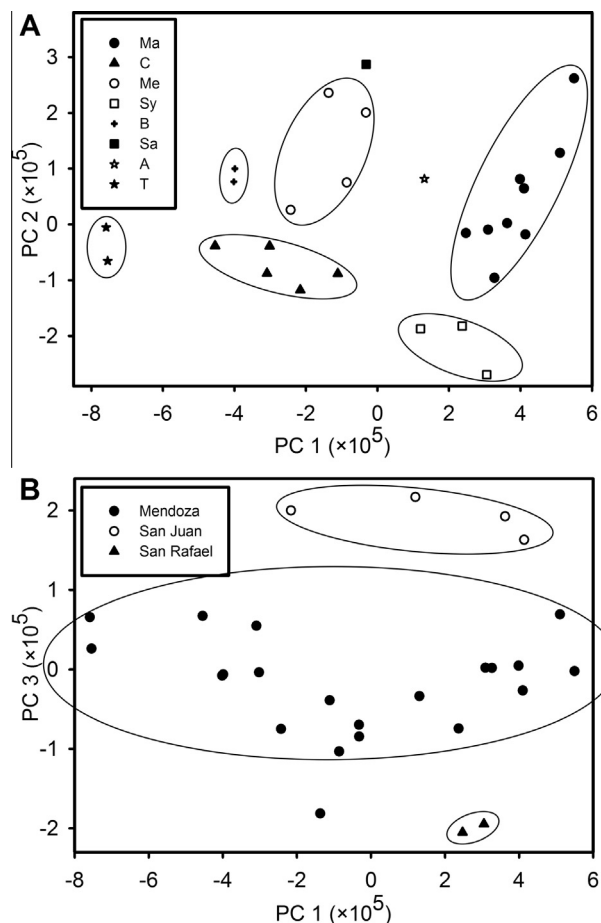
Examination of the contribution of the twenty constituents resolved by MCR-ALS in each principal component reveals which compounds were decisive for wine discrimination. Therefore, constituents No. 2, 4 and 5 displayed the largest contributions to PC1, constituents No. 2, 4 and 8 were the largest contributors to PC2, and constituents No. 1, 4 and 8 to PC3 (see Supplementary Data). Fig. 4A shows the overlapped resolved mass spectra of the five relevant constituents resolved by MCR-ALS, in which it can be observed that base peaks of 129, 493, 535, 618, and  $m/z$  639 ratios (from the constituents 8, 2, 4, 1, and 5, respectively) can be assigned to different anthocyanin compounds (*vide infra*).

#### 4.3. D-UPLS analysis

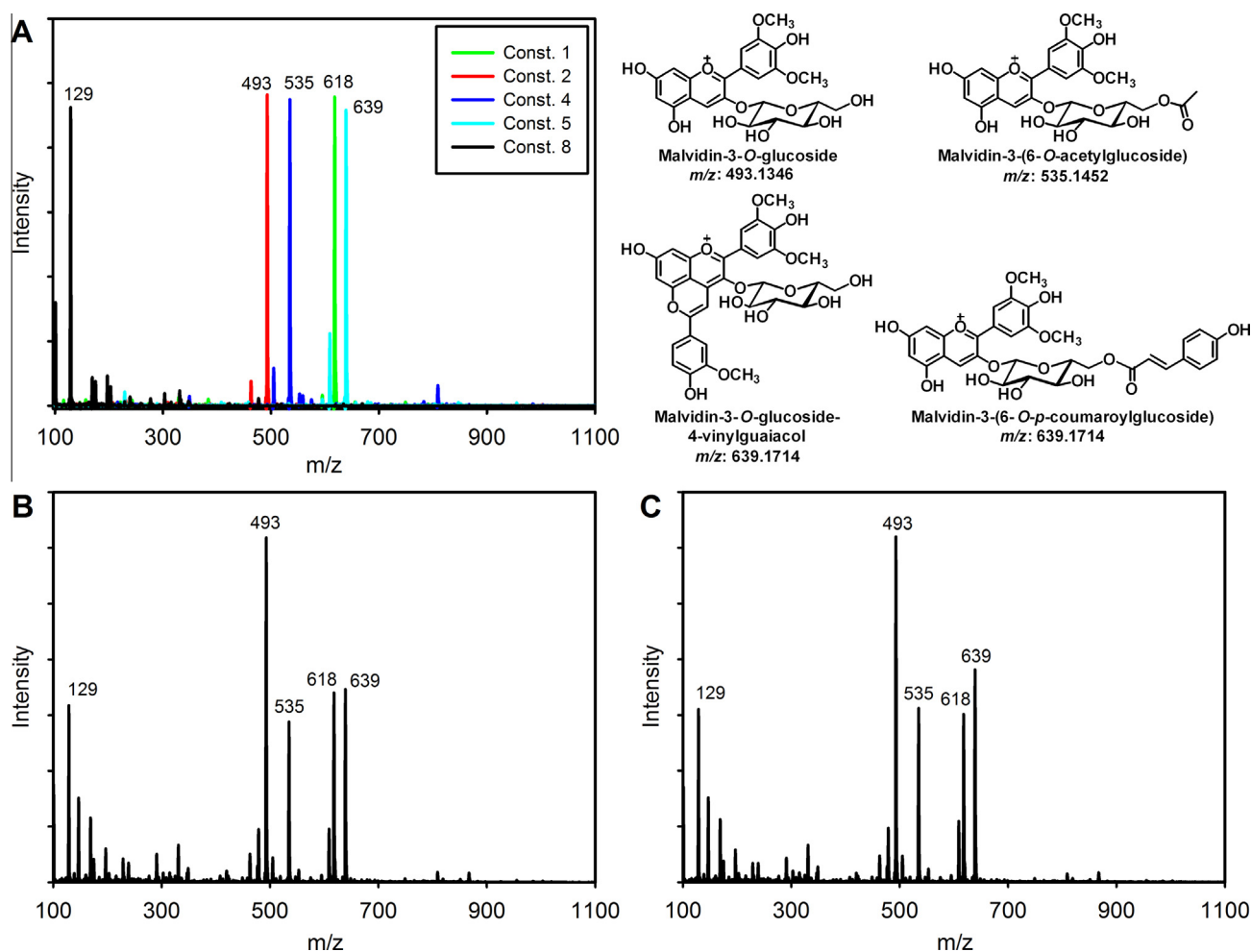
In this section we describe the application of a second, complementary technique to the discrimination of wine samples based on LC–MS data. Due to the high variability in the data for both *varietal* and geographical origin, selection of D-UPLS models was directed to form well-defined groups for each classification model. Regarding geographical origin of the wine samples, Galán winery (east Mendoza) is located in a low altitude zone, and therefore less radiation is involved in the grape maturing process, wines from San Juan (north Mendoza) come from a completely different *terroir*, with a much more arid climate and higher average temperatures, while CoViTu winery is from Valle de Uco (south Mendoza city) and produce wines from a higher altitude zone, with more radiation and larger thermal amplitude than the remaining classes. Therefore, two classification models were designed (Table 1), with the intention of discriminating wine samples according to: (1) *varietal* (D-UPLS-1), and (2) geographical origin (D-UPLS-2). In the D-UPLS-1 model, four classes were considered: 1 (Malbec), 2 (Cabernet Sauvignon), 3 (Merlot), and 4 (remaining samples not belonging to the former three *varietals*), whereas in the D-UPLS-2 model, four classes were considered: 1 (Galán, east Mendoza), 2 (CoViTu, south Mendoza), 3 (San Juan, north Mendoza), and 4 (remaining samples not belonging to the former three geographical origins). Table 1 also shows the discrimination potential offered by the D-UPLS models, representing the prediction abilities in percentage of the samples correctly classified. The prediction ability was 100% (no misclassified samples); therefore an excellent separation among the wine samples was obtained by the designed D-UPLS models for *varietal* and geographical origin discrimination respectively. For the D-UPLS modeling, two PLS latent variables, estimated from leave-one-out cross-validation, accounted for most of the variability of the data, capturing 88% and 85% of the X and Y variables, respectively.

#### 4.4. Variables importance in the projections of D-UPLS

Once discrimination of the classes was achieved by each D-UPLS model, our interest was to reveal which compounds effectively helped to perform this classification. To do so, we explored a



**Fig. 3.** Discrimination of wine samples from principal component analysis. (A) PC2 vs PC1. (B) PC3 vs PC1. A: Aspiran; B: Bonarda; C: Cabernet Sauvignon; Ma: Malbec; Me: Merlot; Sa: Sangiovese; Sy: Syrah; T: Tempranillo.



**Fig. 4.** (A) Mass spectra of the five relevant constituents resolved by MCR-ALS. (B) Variables importance in the projections (VIP) obtained for mass spectra of D-UPLS-1 model (*varietals*). (C) VIPs obtained for mass spectra of D-UPLS-2 model (*geographical origin*).

**Table 1**

Discrimination models for wine samples according *varietal* (D-UPLS-1) and *geographical origin* (D-UPLS-2).

Class	D-UPLS-1 model			D-UPLS-2 model		
	<i>Varietal</i>	No. samples	Classified samples	<i>Geographical origin</i>	No. samples	Classified samples
1	Cabernet Sauvignon	5	5 (100%)	East Mendoza	6	6 (100%)
2	Malbec	9	9 (100%)	South Mendoza	5	5 (100%)
3	Merlot	4	4 (100%)	San Juan	4	4 (100%)
4	Other	9	9 (100%)	Other	12	12 (100%)

variable selection technique for PLS regression, such as variables importance in the projection (Mehmood, Liland, Snipen, & Sæbø, 2012) (see Supplementary Data).

The calculation of VIPs for the wine dataset furnishes a three-dimensional surface defined in the time and  $m/z$  modes, in which peaks for specific sites are highlighted from the rest. These elution times and  $m/z$  values correspond to the variables that contributed most to the classification models. Fig. 4B and C shows a plot of the five overlapped VIPs obtained for mass spectra of the models D-UPLS-1 and D-UPLS-2, respectively. As can be seen, the base peaks found in these five VIPs exactly match between the two D-UPLS models, and also with the mass spectra of the five relevant constituents resolved by MCR-ALS (Fig. 4A). This means that the outcome of two different chemometric modeling of LC–MS wine data not only yields concordant results, but also allows to conclude that

the classification models for wine *varietals* and *geographical origin* were correctly designed.

Comparison of these five  $m/z$  ratios with original LC–MS raw data, and with literature reported data on retention times and  $m/z$  values (Alberts et al., 2012; Blanco-Vega et al., 2014; De Villiers et al., 2004; Flamini & Traldi, 2010; Fraige, Pereira-Filho, & Carrilho, 2014) about anthocyanin compounds found in wine allow the tentative identification of three of the five compounds, as described in Table 2. Structures of these compounds are shown in Fig. 4. The compound with retention time of 10.28 min was assigned to malvidin-3-*O*-glucoside with a molecular ion of 493.1346, and the compound with retention time of 13.09 min was assigned to malvidin-3-(6-*O*-acetylglucoside) with a molecular ion of 535.1452. Finally, the compound corresponding to a retention time of 15.24 min could be assigned to

**Table 2**

Tentative identification of compounds selected by VIPs analysis with their retention times of TIC (Total Ion Current) chromatograms and  $m/z$  values.

RT (min)	Compound	VIPs values	Molecular ion [M] <sup>+</sup>
6.47	Unknown	618	
9.12	Unknown	129	
10.28	Malvidin-3- <i>O</i> -glucoside	493	493.1346
13.09	Malvidin-3-(6- <i>O</i> -acetylglucoside)	535	535.1452
15.24	Malvidin-3- <i>O</i> -glucoside-4-vinylguaiacol or Malvidin-3-(6- <i>O</i> - <i>p</i> -coumaroylglucoside)	639	639.1714
18.71	Unknown	129	

malvidin-3-*O*-glucoside-4-vinylguaiacol or malvidin-3-(6-*O*-*p*-coumaroylglucoside) as their molecular ion  $m/z$  ratios are equal to 639.1714 for both compounds.

As mentioned above, in most reported works using elution profiles as a fingerprint for wine classification, it is very difficult to identify key compounds within that profile that contributed most to achieve the classification when no pre-treatment was applied to the wine sample. In concordance with our previous work (Pisano et al., 2014), in which we conclude that different anthocyanin compounds contributed to wine discrimination, in this work we found that anthocyanin compounds that contribute to discrimination were three malvidin-derived anthocyanins, considering the remarkable fact that the study was performed by direct injection of untreated wine samples.

## 5. Conclusions

Direct injection of untreated wine samples into an LC–MS system and chemometric data processing allowed the discrimination of red wine samples according to *varietal* and geographical origin. Whereas MCR–ALS resolution of LC–MS data allows the successful discrimination of eight different *varietals*, D–UPLS models adequately discriminated Malbec, Merlot, and Cabernet Sauvignon *varietals* from the remaining samples, and also three geographical regions (east Mendoza, south Mendoza and San Juan) were differentiated from the remaining regions. Furthermore, analysis of the variables importance in the projection for the D–UPLS models allowed identification of malvidin-3-*O*-glucoside, malvidin-3-(6-*O*-acetylglucoside), and malvidin-3-*O*-glucoside-4-vinylguaiacol (or malvidin-3-(6-*O*-*p*-coumaroylglucoside)) as the malvidin-derived anthocyanins responsible of both types of discrimination.

## Acknowledgements

We acknowledge financial support from Universidad Nacional de Rosario, Universidad Nacional de Cuyo, CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) and ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, Project PICT 2010-0084). P.L.P. thanks CONICET for a postdoctoral fellowship.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.foodchem.2014.11.124>.

## References

Abad-García, B., Berrueta, L. A., Garmón-Lobato, S., Gallo, B., & Vicente, F. (2009). A general analytical strategy for the characterization of phenolic compounds in fruit juices by high-performance liquid chromatography with diode array detection coupled to electrospray ionization and triple quadrupole mass spectrometry. *Journal of Chromatography A*, 1216(28), 5398–5415.

- Alberts, P., Stander, M. A., & de Villiers, A. (2012). Advanced ultra high pressure liquid chromatography–tandem mass spectrometric methods for the screening of red wine anthocyanins and derived pigments. *Journal of Chromatography A*, 1235, 92–102.
- Alves, R. F., Nascimento, A. M. D., & Nogueira, J. M. F. (2005). Characterization of the aroma profile of Madeira wine by sorptive extraction techniques. *Analytica Chimica Acta*, 546(1), 11–21.
- Arancibia, J. A., Boschetti, C. E., Olivieri, A. C., & Escandar, G. M. (2008). Screening of oil samples on the basis of excitation–emission room-temperature phosphorescence data and multiway chemometric techniques. Introducing the second-order advantage in a classification study. *Analytical Chemistry*, 80(8), 2789–2798.
- Ballabio, D., Skov, T., Leardi, R., & Bro, R. (2008). Classification of GC–MS measurements of wines by combining data dimension reduction and variable selection techniques. *Journal of Chemometrics*, 22(8), 457–463.
- Bertoncelj, J., Polak, T., Kropf, U., Korošec, M., & Golob, T. (2011). LC–DAD–ESI/MS analysis of flavonoids and abscisic acid with chemometric approach for the classification of Slovenian honey. *Food Chemistry*, 127(1), 296–302.
- Blanco-Vega, D., Gómez-Alonso, S., & Hermosín-Gutiérrez, I. (2014). Identification, content and distribution of anthocyanins and low molecular weight anthocyanin-derived pigments in Spanish commercial red wines. *Food Chemistry*, 158, 449–458.
- Cuadros-Inostroza, A., Giavalisco, P., Hummel, J., Eckardt, A., Willmitzer, L., & Peña-Cortés, H. (2010). Discrimination of wine attributes by metabolome analysis. *Analytical Chemistry*, 82(9), 3573–3580.
- de Villiers, A., Vanhoenacker, G., Majek, P., & Sandra, P. (2004). Determination of anthocyanins in wine by direct injection liquid chromatography–diode array detection–mass spectrometry and classification of wines using discriminant analysis. *Journal of Chromatography A*, 1054(1–2), 195–204.
- Escandar, G. M., Olivieri, A. C., Faber, N. M., Goicoechea, H. C., Muñoz de la Peña, A., & Poppi, R. J. (2007). Second- and third-order multivariate calibration: Data, algorithms and applications. *TrAC Trends in Analytical Chemistry*, 26(7), 752–765.
- Fanzone, M., Peña-Neira, Á., Jofré, V., Asof, M., & Zamora, F. (2010). Phenolic characterization of Malbec wines from Mendoza province (Argentina). *Journal of Agricultural and Food Chemistry*, 58, 2388–2397.
- Flamini, R., & Traldi, P. (2010). Grape and wine polyphenols. In: *Mass spectrometry in grape and wine chemistry* (pp. 163–225). John Wiley & Sons Inc.
- Fraige, K., Pereira-Filho, E. R., & Carrilho, E. (2014). Fingerprinting of anthocyanins from grapes produced in Brazil using HPLC–DAD–MS and exploratory analysis by principal component analysis. *Food Chemistry*, 145, 395–403.
- Garrido-Delgado, R., López-Vidal, S., Arce, L., & Valcárcel, M. (2009). Differentiation and identification of white wine varieties by using electropherogram fingerprints obtained with CE. *Journal of Separation Science*, 32(21), 3809–3816.
- Gu, H.-W., Wu, H.-L., Yin, X.-L., Li, Y., Liu, Y.-J., Xia, H., et al. (2014). Multi-targeted interference-free determination of ten  $\beta$ -blockers in human urine and plasma samples by alternating trilinear decomposition algorithm-assisted liquid chromatography–mass spectrometry in full scan mode: Comparison with multiple reaction monitoring. *Analytica Chimica Acta*, 848, 10–24.
- Jaitz, L., Siegl, K., Eder, R., Rak, G., Abranko, L., Koellensperger, G., et al. (2010). LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. *Food Chemistry*, 122(1), 366–372.
- Jaumot, J., Gargallo, R., de Juan, A., & Tauler, R. (2005). A graphical user-friendly interface for MCR–ALS: A new tool for multivariate curve resolution in MATLAB. *Chemometrics and Intelligent Laboratory Systems*, 76(1), 101–110.
- Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). New York: Springer.
- Kulakowski, D. M., Wu, S.-B., Balick, M. J., & Kennelly, E. J. (2014). Merging bioactivity with liquid chromatography–mass spectrometry-based chemometrics to identify minor immunomodulatory compounds from a Micronesian adaptogen, *Phaleria nisidai*. *Journal of Chromatography A*, 1364, 74–82.
- Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 118, 62–69.
- Nile, S. H., & Park, S. W. (2014). Edible berries: Bioactive components and their effect on human health. *Nutrition*, 30(2), 134–144.
- Olivieri, A. C., & Escandar, G. M. (2014). *Practical three-way calibration*. Amsterdam: Elsevier.
- Olivieri, A. C., Wu, H.-L., & Yu, R.-Q. (2009). MVC2: A MATLAB graphical interface toolbox for second-order multivariate calibration. *Chemometrics and Intelligent Laboratory Systems*, 96(2), 246–251.
- Pazourek, J., Gajdošová, D., Spanilá, M., Farková, M., Novotná, K., & Havel, J. (2005). Analysis of polyphenols in wines: Correlation between total polyphenolic content and antioxidant potential from photometric measurements: Prediction of cultivars and vintage from capillary zone electrophoresis fingerprints using artificial neural network. *Journal of Chromatography A*, 1081(1), 48–54.
- Pisano, P. L., Silva, M. F., & Olivieri, A. C. (2014). Exploration of liquid chromatographic–diode array data for Argentinean wines by extended multivariate curve resolution. *Chemometrics and Intelligent Laboratory Systems*, 132, 1–7.
- Ribéreau-Gayon, P., Glories, Y., Maujean, A., & Dubourdieu, D. (2006). Phenolic compounds. In: *Handbook of Enology* (pp. 141–203). John Wiley & Sons Ltd.
- Taamalli, A., Arráez Román, D., Zarrouk, M., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2012). Classification of ‘Chemlali’ accessions according to the geographical area using chemometric methods of phenolic profiles analysed by HPLC–ESI–TOF–MS. *Food Chemistry*, 132(1), 561–566.

- Tauler, R. (1995). Multivariate curve resolution applied to second order data. *Chemometrics and Intelligent Laboratory Systems*, 30(1), 133–146.
- Tredoux, A., de Villiers, A., Májek, P., Lynen, F. d. r., Crouch, A., & Sandra, P. (2008). Stir bar sorptive extraction combined with GC–MS analysis and chemometric methods for the classification of South African wines according to the volatile composition. *Journal of Agricultural and Food Chemistry*, 56(12), 4286–4296.
- Vaclavik, L., Lacina, O., Hajslova, J., & Zweigenbaum, J. (2011). The use of high performance liquid chromatography–quadrupole time-of-flight mass spectrometry coupled to advanced data mining and chemometric tools for discrimination and classification of red wines according to their variety. *Analytica Chimica Acta*, 685(1), 45–51.
- Wold, S., Geladi, P., Esbensen, K., & Öhman, J. (1987). Multi-way principal components-and PLS-analysis. *Journal of Chemometrics*, 1(1), 41–56.
- Xia, B., Zhou, Y., Tan, H. S., Ding, L. S., & Xu, H. X. (2014). Advanced ultra-performance liquid chromatography–photodiode array–quadrupole time-of-flight mass spectrometric methods for simultaneous screening and quantification of triterpenoids in *Poria cocos*. *Food Chemistry*, 152, 237–244.