**UHPLC-HRMS (Orbitrap) Fingerprinting in the Classification and Authentication of Cranberry-based Natural Products and Pharmaceuticals using Multivariate Calibration Methods**

Sergio Barbosa[a], Naiara Pardo-Mates[a], Miriam Hidalgo-Serrano[a], Javier Saurina[a,b], Lluis Puignou[a,b] and Oscar Núñez[a,b,c,*]

[a] Department of Chemical Engineering and Analytical Chemistry, University of Barcelona. Martí i Franquès 1-11, E08028 Barcelona, Spain.

[b] Research Institute in Food Nutrition and Food Safety, University of Barcelona. Av. Prat de la Riba 171, Edifici Recerca (Gaudí), E-08901 Santa Coloma de Gramanet, Barcelona, Spain.

[c] Serra Hunter Fellow. Generalitat de Catalunya (Spain).

* Corresponding author: Oscar Núñez

Department of Chemical Engineering and Analytical Chemistry, University of Barcelona
Martí i Franquès 1-11, E-08028, Barcelona, Spain.
Phone: 34-93-403-3706
Fax: 34-93-402-1233
e-mail: oscar.nunez@ub.edu

**Abstract**

UHPLC-HRMS (Orbitrap) fingerprinting in negative and positive H-ESI mode was applied to the characterization, classification and authentication of cranberry-based natural and pharmaceutical products. HRMS data in full scan mode (*m/z* 100-1500) at a resolution of 70,000 full-width at half maximum was recorded and processed with MSConvert software to obtain a profile of peak intensities in function of *m/z* values and retention times. A threshold peak filter of absolute intensity ($10^5$ counts) was applied to reduce data complexity. Principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA) revealed patterns able to discriminate the analyzed samples according to the fruit of origin (cranberry, grape, blueberry and raspberry). Discrimination among cranberry-based natural and cranberry-based pharmaceutical preparations was also achieved. Both, UHPLC-HRMS fingerprints in negative and positive H-ESI modes, as well as the data fusion of both acquisition modes, showed to be good chemical descriptors to address cranberry extracts authentication. Validation of the proposed methodology showed a prediction rate of 100% of the samples. Obtained data was further treated by partial least squares (PLS) regression to identify frauds and quantify the percentage of adulterant fruits in cranberry-fruit extracts, achieving prediction errors in the range 0.17-3.86%.

**Keywords:** Non-targeted UHPLC-HRMS fingerprinting; Cranberry-based products; Food characterization; Food Authentication; UHPLC; High resolution mass spectrometry;

## 1. Introduction

Nowadays, it is a common practice worldwide to address the prevention of several chronic diseases by employing, together or not with the use of regulated medicines, plant-based and/or fruit-based pharmaceutical extracts. This is the case of American red cranberry (*Vaccinium macrocarpon*), a small evergreen shrub from the Ericaceae family that grows in acid swamps in humid forests. American cranberry fruits are composed mostly of water (>80%), and are a rich source of vitamin C and dietary polyphenols, such as flavonols, anthocyanins, organic acids and proanthocyanidins (PACs). Cranberries have been used for centuries as a flavoring or by sailors to prevent scurvy due to its high vitamin C content. Moreover, their consumption may be associated with reduced risk of chronic diseases such as cancer, although strong evidences have not been yet established in humans, and several berry-based extracts have shown antitumor activities. Cranberry-extracts enriched in polyphenolic contents have also shown enhanced antiproliferative activity. These extracts may also play an important role in the treatment of oral infections by reducing the pathogenesis of dental caries, the protection against cardiovascular diseases, and the prevention of oxidation of low density lipoproteins and platelet aggregation.[1–5]

Recently cranberries have attracted much attention due to their high content on PACs and the capacity of some of them to prevent urinary tract infections (UTIs). This activity is attributed to the inhibition of the adhesion of pathogenic bacteria, such as Escherichia coli and Helicobacter pylori, to the cells of the urinary tract tissues, thus preventing bacterial colonization and the proliferation of infections.[6–9] PACs, also known as condensed tannins, are flavan-3-ol polymeric structures mainly based on (epi)catechin oligomers, called procyanidins, but other forms can have (epi)gallocatechin units (i.e. prodelphinidins) or (epi)afzelechin (i.e. propelargonidins) units.[10]

PACs can be classified according to the linkage between their units. PACs linked through C4-C8 or C4-C6 bonds are known as B-type PACs. If these structures have an additional ether linkage between C2-C5 or C2-C7 they are known as A-type PACs.[11] As an example, Fig. S1 (electronic supplementary information) shows the structure of a trimeric PAC with A-type and B-type linkages. Nevertheless, only A-type PACs, very abundant in American red cranberries, exhibit the bioactive activity to prevent UTIs, while B-type PACs, which are found in other fruits such as grapes and blueberries, do not show this activity.[12–14] Many cranberry-based pharmaceutical

preparations have recently appeared in the market to prevent UTIs, and there is the suspicion that some of them do not contain the necessary bioactive PACs. The fact that only A-type PACs have the required bioactive capacity and that pharmaceutical laboratories frequently assess the total content of PACs by non-selective colorimetric methods,[15,16] unable to differentiate among A- and B-type PACs, demonstrates the importance on developing analytical methods to characterize cranberry fruit-based extracts and pharmaceutical preparations to authenticate the fruit of origin employed in these processed extracts and to prevent frauds. Moreover, due to food trade globalization and the increased complexity of supply chains, the need for effective systems to protect consumers from impure, contaminated and fraudulently presented food-processed products has increased. Current food labeling and traceability systems cannot strictly guarantee that the food is authentic, of good quality and safe. As a result, consumers are demanding verifiable traceability evidences as an important criterion of food quality and safety.[17]

Liquid chromatography coupled to mass spectrometry (LC-MS), tandem mass spectrometry (LC-MS/MS) and high resolution mass spectrometry (LC-HRMS), in combination with chemometric methods, emerge today as the best analytical tools to characterize, classify and authenticate food products.[18–23] These platforms result in one of the best ways to detect fraudulent practices derived from the substitution of the most valued components in the fruit-processed extracts by others of lower commercial value, with worse organoleptic characteristics, or without the intended beneficial properties for human health. Food fingerprinting, the non-targeted chemical analysis of food products with multivariate data analysis, is emerging as an innovative approach for food authentication.[24–26] This approach is based on the principle of metabolomics, which describes the scientific study of metabolites (small molecules below 1,500 Da), present in a biological system with the aim to detect as many components as possible. Although the main focus of metabolomics are in the field of pharmacology and toxicology, the use of these approaches in food science is gaining acceptance. However, in the food field, an important distinction is made between the concepts of food fingerprinting and food profiling in accordance to the corresponding definitions of metabolomics.[24,27] Researchers coming from the metabolomic field use “profiling” and “fingerprinting” on a different way to researchers who are devoted to food science. The arrival of a “foodomics” discipline was not enough to allay this terminological problem, since authors keep on using the terms with both meanings.[26] Food profiling focuses on the

analysis of a group of known selected metabolic chemicals, or a group of chemicals belonging to the same family or with a similar structural feature. The concentrations (or peak signals) of these targeted compounds are then used as food features (markers) to address food authentication. In contrast, food fingerprinting do not deal with the identification of metabolites, but on the recognition of patterns, the so-called "fingerprints" of the foodstuff.[28] After identification and mapping of the patterns to individual food matrices, the objective is usually to differentiate between various food fingerprints in terms of food features such as botanical species, geographical origin, or with respect to possible food adulterations.[24]

The fact that similar fruit extracts but with different properties are used as the ingredients in the preparation of food processed products and pharmaceutical preparations increases the difficulty of using targeted methods and may need to employ non-targeted approaches in order to obtain specific fingerprints of the original products. These so-called food fingerprinting approaches aim to capture as many compounds or features as technically possible to gain a comprehensive insight into the composition of the sample.[25] However, a large amount of chemical data is obtained making difficult its treatment. In this regard, several chemometric data processing software packages with different characteristics and algorithms have been introduced for MS users.[29] After data acquisition and processing, chemometric univariate and multivariate statistical methods are then used for sample characterization, classification and authentication.[30,31]

The aim of this work was to develop a suitable method to characterize, classify and authenticate natural and pharmaceutical cranberry-based products, employing ultra-high performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS) using a non-targeted fingerprinting approach with a Q-Exactive Orbitrap analyzer. Different classes of fruit-based (cranberry, blueberry, raspberry and grape) products including the raw fruit extracts, fruit juices and raisins, as well as commercial cranberry-based pharmaceutical preparations including raw extracts, powder capsules, syrups and sachets were analyzed after a simple sample extraction procedure. The hypothesis established in this work is that UHPLC-HRMS fingerprinting data, obtained in both positive and negative ESI mode, exploring also the possibility of data-fusion, can be considered as a source of potential chemical descriptors to be exploited for the characterization and classification of fruit-based natural products and pharmaceuticals by unsupervised principal component analysis (PCA) and supervised partial least squares-discriminant analysis (PLS-DA). Data was further treated by partial least square

(PLS) regression to quantify percentages of fruit extracts (grape, blueberry and raspberry) used for adulteration in cranberry extracts.

## 2. Experimental

### 2.1. Chemicals

Unless otherwise specified, the reagents and chemicals used were always of analytical grade. LC-MS grade water, methanol, acetonitrile, formic acid (98-100 %) and acetone were purchased from Sigma-Aldrich (Steinhein, Germany), and hydrochloric acid (98%) from Merck (Seelze, Germany).

### 2.2. Instrumentation and methods

The chromatographic fingerprints were obtained with an Accela UHPLC system (Thermo Fisher Scientific, San José, CA, USA) consisting of a quaternary pump and an Accela AS autosampler, coupled to a Q-Exactive Orbitrap HRMS system (Thermo Fisher Scientific) mass spectrometer equipped with a heated electrospray ionization source (HESI-II) operated in positive and negative ionization mode by means of two runs, respectively. Nitrogen (purity >99.98 %) was used as the nebulizer and desolvation gas. Specific parameters were as follows: sheath gas, 60 a.u. (arbitrary units); sweep gas, 0 a.u.; auxiliary gas, 10 a.u.; capillary temperature, 320 ºC; HESI-II probe temperature, 350 ºC; electrospray voltage, 2.5 kV; S-lens RF level, 50 V. Q-Exactive Orbitrap HRMS system was tuned and calibrated using Thermo Fisher calibration solutions every three days in both negative and positive modes to ensure a working mass accuracy error lower than 5 ppm. Mass spectra were acquired in full MS scan mode employing two runs, in both positive and negative, respectively, at a resolution of 70,000 full-width at half maximum (FWHM) at *m/z* 200 with a scan range of 100-1,500 *m/z*. The automatic gain control (AGC) target was set at $10^6$ and the maximum injection time (IT) was 200 ms.

A porous-shell Ascentis® Express C18 reversed-phase column (150 × 2.1 mm, 2.7 µm particle size) provided by Supelco (Bellefonte, PA, USA) was used for the separation. Gradient elution was performed with 0.1% formic acid aqueous solution (solvent A) and 0.1% formic acid in acetonitrile (solvent B) at a constant flow rate of 0.3 mL/min with the following gradient program: 0-1 min at 10% B; 1-20 min from 10 to 95% B; 20-23 min at 95% B, 23-24 min back to initial conditions at 10% B; and column re-equilibration for 6 min at initial conditions. Injection volume was 10 µL.

### 2.3. Sample treatment

A total of 106 natural and pharmaceutical products were analyzed in this work. Natural products from different brands were purchased from Barcelona markets and pharmaceutical preparations and raw extracts were provided by Deiters S.L. Company (Barcelona, Spain). The samples included 22 commercial pharmaceutical berry-based products (11 capsules, 2 syrups, 4 sachets and 5 extracts), 33 cranberry-based products (4 fruit samples, 8 raisin samples and 21 juice samples), 29 grape-based products (4 fruit samples, 8 raisin samples and 17 juice samples), 12 blueberry-based products (6 fruit samples and 6 juice samples) and 10 raspberry-based fruit samples. All fruits and raisins were grinded using an Ika Ultra-Turrax machine (Staufen, Germany) with different applicators. Water, in a 1:1 sample:water ratio, was added to raisins to improve the crushing. Then, all analyzed samples (10 g for solid sample, and 25 mL for juices and syrups) were freeze-dried to achieve fully lyophilized products. For this purpose, samples remained 24 h inside a lyophilizer (Telstar LyoQuest, Terrasa, Spain) with a gradient temperature ramp from -80 ºC to room temperature and then were kept for 6.5 h at 40 ºC. Once lyophilized extracts were obtained, sample treatment was carried out using a method previously described.[18,19] Briefly, 0.1 g of lyophilized sample were dispersed in 10 mL of an acetone:water:hydrochloric acid (70:29.9:0.1 *v/v/v*) solution by sonication for 10 min. Then, the samples were centrifuged for 15 min at 3500 rpm, and the supernatant extracts separated from the solid and stored at -4 ºC until analyzed. Before injection, extracts were filtered through 0.45 µm nylon filters (Whatman, Clifton, NJ, USA). Besides, a quality control (QC) consisting of a mixture of 50 µL of each sample extract was prepared and injected every 10 samples to evaluate the repeatability of the method and the robustness of the chemometric results.

For authentication studies by PLS regression, three cases were evaluated in which cranberry extracts were adulterated with different amounts of grape, blueberry or raspberry, respectively. For such a purpose, 3 cranberry, 3 grape, 3 blueberry- and 3 raspberry-fruit sample extracts were processed as indicated above. These series of extracts were used to prepare standard and unknown samples to be used for calibration and prediction sets. Hence, apart from those pure extracts, mixtures of cranberry and the adulterant fruits were prepared as follows: 50% adulterant (5 samples), 20% adulterant (3 samples), 12% adulterant (3 samples), 10% adulterant (3 samples), 7% adulterant (3 samples), 6% adulterant (3 samples), 5% adulterant (3 samples), 2.5% adulterant (3 samples), and 2% adulterant (3 samples), for each adulterant fruit. 2% adulterant was

established as the minimum adulterant level that can be quantified with the proposed methodology.

### 2.4. Data analysis

The data treatment for the untargeted analysis was carried out with *R* software (R Foundation, Vienna, Austria). Data chemometric analyses with PCA, PLS-DA and PLS regression were carried out with the Stand Alone Chemometrics Software (Solo), provided by Eigenvector Research Inc. (Manson, WA, USA).[32] A detailed description of the theoretical background of these chemometric methods is given elsewhere.[33]

The X-data matrices to be treated by PCA and PLS-DA consisted of the UHPLC-ESI-q-Orbitrap fingerprints (peak intensities as a function of retention times and *m/z* values) obtained in both H-ESI(+) and H-ESI(-), as well as the data-fusion combination of the two ionization modes. The Y-data matrix in the PLS-DA models consisted of the sample classes. Scatter plots of scores and loadings of the principal components (PCs), in PCA, and of the latent variables (LVs), in PLS-DA, were used to investigate the structure of maps of samples and variables, respectively. The quantification of the percentage of fruit-extract used for adulteration (grape, blueberry or raspberry extracts) in the adulterated cranberry-based extracts analyzed was based on PLS regression. Samples available were distributed among training and test sets as follows. Training set: 100% adulterant (3 samples), 50% adulterant (5 samples), 20% adulterant (3 samples), 10% adulterant (3 samples), 7% adulterant (3 samples), 5% adulterant (3 samples), 2% adulterant (3 samples), and 100% cranberry-fruit (3 samples). The remaining samples considered as unknown (12% adulterant, 6% adulterant, 2.5% adulterant, 3 samples each) were used for validation and prediction purposes.

## 3. Results and discussion

### 3.1. UHPLC-HRMS fingerprinting

In this work, a non-supervised UHPLC-HRMS fingerprinting analysis of fruit-based products and cranberry-based pharmaceuticals was evaluated in order to obtain proper chemical descriptors for sample classification and authentication. For that purpose, 106 samples were processed with a simple sample extraction method and the obtained extracts were analyzed with a C18 reversed-phase UHPLC-HRMS method

(see experimental section). The fingerprint of a fruit-based product will depend on both the fruit variety genotype and the product phenotype (food attributes determined by ambient conditions, agricultural practices, food-processing procedures, etc.). Thus, it is expected that these fingerprints will provide good chemical descriptors to achieve sample characterization and classification by means of chemometric methods. Accordingly, an untargeted strategy relied on UHPLC-HRMS fingerprints consisting of intensity peaks recorded as a function of *m/z* and retention time. Data was then registered in both negative and positive HRMS full scan mode (*m/z* 100-1500). As an example, Fig. 1 shows the total ion chromatograms (TIC) obtained in negative H-ESI mode for the four types of fruits analyzed (raspberry (a), grape (b), blueberry (c), and cranberry (d)). The figure also shows, as an arbitrary example, the full scan HRMS spectra obtained for each fruit extract at a retention time of 6.61 min. As can be seen, important differences in peak signals and abundances in both total ion chromatograms and HRMS spectra were obtained. Cranberry- and blueberry-based samples seemed to provide richer fingerprints (more signals), while those belonging to grape-based products were simpler.

### 3.2. Exploratory principal component analysis study

The obtained UHPLC-HRMS fingerprint raw data was processed with MSConvert software to obtain a profile of peak intensities in function of *m/z* values and retention times. In order to reduce the data complexity, a threshold peak filter of absolute intensity $10^5$ was applied. The converted data was then processed with R software to obtain a data matrix including the UHPLC-HRMS fingerprints of the 106 samples analyzed and the QCs. The dimension of the obtained data matrices were 118 × 469 for positive H-ESI mode and 118 × 641 for negative H-ESI mode. PCA required 5 and 6 principal components (PCs), for negative and positive H-ESI, respectively, to explain most of the variance observed within the analyzed samples. As an example, Fig. 2 shows the score plots that provided the best sample differentiation: (a) PC2 *vs* PC4 when using UHPLC-HRMS fingerprints in positive H-ESI mode and (b) PC1 *vs* PC3 when using UHPLC-HRMS fingerprints in negative H-ESI mode. As can be seen, complete discrimination between all the analyzed types of samples was not obtained. Nevertheless, samples tend to be grouped according to the fruit of origin. In general, UHPLC-HRMS fingerprinting in negative ionization mode seem to provide more discriminant chemical descriptors among samples which allow to concentrate the

samples in smaller regions within the score plot (see for instance the distribution of cranberry-based and, especially, for raspberry-based natural products). Although at this point complete discrimination among the four fruit types was not achieved, for example blueberry- and grape-based natural products tend to be overlapped in both score plots, cranberry-based natural products were completely separated from the other three types of fruits when using both positive and negative UHPLC-HRMS fingerprints. This is an interesting result because cranberry natural products should be the specific source of the raw extracts employed in the preparation of pharmaceuticals, and thus susceptible to adulteration with other fruit extracts as commented in the introduction.

Finally, in Fig. 2 it can also be observed that cranberry-based samples (both natural products and pharmaceutical preparations) tend to be grouped, more or less, in the same region of the score plots, although with certain discrimination depending on the pharmaceutical form (raw extract, capsules, syrups and sachets), and in some cases clearly differentiated from cranberry-based natural products. This is probably due to the fact that purification and preconcentration procedures followed by pharmaceutical companies in the preparation of raw extracts from cranberry-fruits enriched with bioactive compounds in comparison to non-treated cranberry-fruit natural products, thus providing different patterns even though the fruit of origin is the same.

Taking into account that raspberry, blueberry and grape extracts are expected to be used as potential adulterant of cranberry extracts, independent PCA among cranberry-based natural products and the other three fruit families were also evaluated. In this case, the dimensions of the obtained data matrices were 84 × 469 for positive H-ESI mode and 84 × 641 for negative H-ESI mode. As an example, Fig. 3 shows the score plots that provided the best sample differentiation: (a) PC1 *vs* PC2 in positive H-ESI mode and (b) PC1 *vs* PC3 in negative H-ESI mode. It can be observed that, except for some outliers that are expected when working with natural products, samples tend to be grouped in both cases according to the fruit of origin, although more overlapping between groups was observed in positive ionization mode. Up to this point, it seems that UHPLC-HRMS fingerprinting in negative ionization mode provided better discrimination among cranberry-fruit products and other adulterant fruits, as can be seen in Fig. 3b. PC1 clearly differentiated cranberry-based fruit products (clustered at the right of the plot) from those obtained with other fruits (distributed at the left of the plot). It was thus concluded from PCA that samples were reasonably distinguished as a

function of fruits of origin. Hence, data was expected to be of interest to tackle further classification studies by PLS-DA.

### 3.3. Supervised partial least squares-discriminant analysis study

UHPLC-HRMS fingerprints obtained in both H-ESI negative and positive acquisition modes were here evaluated as chemical descriptors to address sample classification by PLS-DA. For that purpose, no further data treatment from the employed PCA data matrices was required. Therefore, the same X-data matrix employed in PCA was submitted in PLS-DA, while the Y-data matrix coded the belonging of the samples to their corresponding classes (i.e., cranberry products, cranberry-based nutraceuticals, and grape, blueberry and raspberry products). Fig. 4 shows the 3D-scatter plots of scores of LV1 *vs* LV2 *vs* LV3 from UHPLC-HRMS fingerprints in negative H-ESI (Fig. 4a) and in positive H-ESI (Fig. 4b). In addition, data fusion of both negative and positive fingerprints were also evaluated as chemical descriptors for PLS-DA, and the obtained 3D-scatter plot of score of LV1 *vs* LV2 *vs* LV3 is depicted in Fig. 4c. As seen in the figures, in general very acceptable discrimination among the analyzed sample groups (samples tend to be grouped according to the fruit of origin) was obtained independently of the H-ESI acquisition mode employed, as well as when data fusion of both ionization modes was considered. In addition, cranberry-based products are differentiated into two groups, namely: fruit-based and pharmaceutical-based products, in agreement with purification and preconcentration procedures followed applied to nutraceuticals, as previously commented. Nevertheless, both sample groups tend to be distributed on the same area of the 3D-scatter plots of scores and opposed to the grape-, blueberry- and raspberry-fruit based samples. Therefore, either individual data sets from negative or in positive H-ESI mode, or even the data fusion set of both fingerprints, are adequate, a priory, for the characterization and authentication of cranberry-based natural products and pharmaceutical preparations.

In order to validate the proposed methodology, and taking into consideration that raspberry, blueberry and grape extracts are expected to be used as potential adulterants of cranberry extracts, PLS-DA models using UHPLC-HRMS fingerprints in negative H-ESI mode were built by pairs (cranberry vs grape, cranberry vs blueberry and cranberry vs raspberry). The optimum number of latent variables of each PLS-DA model was

stablished cross validation classification error average, being approximately the first minimum point the most appropriate one (Fig. S2, electronic supplementary information). As a good classification was obtained for the three studied pairs. Models were built by using a 70% of each group of samples as the calibration set and were validated with the remaining 30% of the samples. Fig. 5 depict the obtained PLS-DA plots of scores projected on LV1 *vs* LV2 as well as the classification plot of cranberry *vs* raspberry (a), blueberry (b) and grape (c), respectively. As seen in the figures, all samples were correctly assigned to its corresponding class, thus reaching a prediction rate of 100% in each studied case.

### 3.4. Adulteration studies by partial least square regression

UHPLC-HRMS fingerprints in negative H-ESI acquisition mode were also here considered for the authentication and quantitation of fraud levels of fruit extract adulterants in cranberry-based extracts. Thus, a cranberry fruit-extract was adulterated at different levels (from 2 to 50%) of the other three fruits studied (blueberry, raspberry and grapes). Adulterated samples were then processed with the proposed sample treatment procedure and the obtained extract solutions were analyzed by UHPLC-HRMS to obtain the corresponding fingerprints as chemical data for partial least square regression. A data set of calibration as indicated in the experimental section was first employed in order to establish the PLS model. The number of latent variables (LV) used for the assessment of the PLS model was estimated by venetian blinds cross validation considering 2 data splits. The PLS model was further applied to quantify the percentage of adulteration in the samples belonging to the test set. Fig. 6 shows the obtained PLS results when grape (a), blueberry (b), and raspberry (c) were the adulterants, showing the good performance of the obtained PLS models. Calibration errors were in all cases below 0.01% and, in general, small prediction errors were also obtained in the validation study, with values of 0.17% and 0.47% when grape and blueberry were used as adulterants, respectively, except when raspberry was used as adulterant where the prediction error increased up to 3.86%. However, taking into consideration that adulteration levels in nutraceuticals are expected to be high if an economical profit is intended, the proposed methodology showed a good performance for the authentication and quantitation of frauds, even at low adulteration levels.

## 4. Conclusions

A non-targeted UHPLC-HRMS fingerprinting method was employed for the characterization and classification of fruit-based natural products and cranberry-based pharmaceutical preparations. PCA results using both UHPLC-HRMS fingerprints in negative and positive H-ESI mode provided a reasonable discrimination among cranberry-based natural products and pharmaceuticals when compared to those obtained from other fruit extracts. Results improved when using PLS-DA, achieving an acceptable discrimination between the groups of samples analyzed both when using UHPLCHRMS fingerprints in negative and in positive acquisition mode, as well as when combining both acquisition modes in a data fusion matrix. The proposed methods were also able to distinguish among pharmaceutical preparations against cranberry-fruit based samples, probably due to the enhancement or reduction in bioactive compounds because of different extraction and purification steps used by the pharmaceutical companies when preparing their raw extracts.

PLS-DA models using UHPLC-HRMS fingerprints in H-ESI negative mode as chemical descriptors of cranberry-fruit based products against possible adulterant fruit extracts (grape, blueberry and raspberry) showed good discrimination among samples depending on the fruit of origin, with a prediction rate of 100% in each case study when 30% of the samples were employed as validation set. Moreover, the proposed chemical descriptors were also useful to achieve adulterant level quantification by PLS regression, showing calibration errors below 0.01% and prediction errors in the range of 0.17-3.86%.

As hypothesized in the introduction section, the results obtained in this work demonstrate that non-targeted UHPLC-HRMS fingerprinting methodology can be applied to the characterization, classification and authentication of cranberry-based products and pharmaceuticals adulterated with more economic fruit-based products such as grape, blueberry and raspberry extracts. Moreover, PLS results showed that adulterant levels below 2.5% can be quantified successfully with low enough calibration and validation errors.

Finally, the proposed methodology will allow to obtain in a fast way fingerprint chemical descriptors with lower data processing in comparison to targeted UHPLC-HRMS profiling approaches in which a known family of bioactive compounds need to be characterized, detected and their signal confirmed and quantified by employing standards. This will make non-targeted UHPLC-HRMS fingerprinting methods cheaper

than targeted approaches as chemical standards are not required to perfectly achieve sample classification.

## Conflict of interest declaration

The authors declare no conflict of interest.


## Acknowledgements

The authors gratefully acknowledge the financial support received from the Spanish Ministry of Economy and Competitiveness under the projects CTQ2014-56324-C2-1-P and CTQ2015-63968-C2-1-P and from the Agency for Administration of University and Research Grants (Generalitat de Catalunya, Spain) under the projects 2017SGR-171 and 2017SGR-310. Authors wished to thank Deiters S.L. Company for providing some cranberry-based raw material extracts and several commercial cranberry-based pharmaceutical preparations.


## References


1 K. Feghali, M. Feldman, V. D. La, J. Santos and D. Grenier, *J. Agric. Food Chem.*, 2012, **60**, 5728–5735.

2 I. Tarascou, J. P. Mazauric, E. Meudec, J. M. Souquet, D. Cunningham, S. Nojeim, V. Cheynier and H. Fulcrand, *Food Chem.*, 2011, **128**, 802–810.

3 N. P. Seeram, L. S. Adams, M. L. Hardy and D. Heber, *J. Agric. Food Chem.*, 2004, **52**, 2512–2517.

4 H. El Gharras, *Int. J. Food Sci. Technol.*, 2009, **44**, 2512–2518.

5 J. Côté, S. Caillet, G. Doyon, J. F. Sylvain and M. Lacroix, *Crit. Rev. Food Sci. Nutr.*, 2010, **50**, 666–679.

6 I. Iswaldi, A. M. Gómez-Caravaca, D. Arráez-Román, J. Uberos, M. Lardón, A. Segura-Carretero and A. Fernández-Gutiérrez, *J. Pharm. Biomed. Anal.*, 2012, **58**, 34–41.

7 F. Sánchez-Patán, B. Bartolomé, P. J. Martín-Alvarez, M. Anderson, A. Howell and M. Monagas, *J. Agric. Food Chem.*, 2012, **60**, 3396–3408.

8 E. Jungfer, B. F. Zimmermann, A. Ruttkat and R. Galensa, *J. Agric. Food Chem.*, 2012, **60**, 9688–9696.

9 Â. Luís, F. Domingues, L. Pereira and Â. Luís, *J. Urol.*, 2017, **198**, 614–621.

10 H. Parastar, M. Jalali-Heravi, H. Sereshti and A. Mani-Varnosfaderani, *J. Chromatogr. A*, 2012, **1251**, 176–187.

11 L. Z. Lin, J. Sun, P. Chen, M. J. Monagas and J. M. Harnly, *J. Agric. Food Chem.*, 2014, **62**, 9387–9400.

12 L. Y. Foo, Y. Lu, A. B. Howell and N. Vorsa, *Phytochemistry*, 2000, **54**, 173–181.

13 C. G. Krueger, J. D. Reed, R. P. Feliciano and A. B. Howell, *Anal. Bioanal. Chem.*, 2013, **405**, 4385–4395.

14 A. B. Howell, J. D. Reed, C. G. Krueger, R. Winterbottom, D. G. Cunningham and M. Leahy, *Phytochemistry*, 2005, **66**, 2281–2291.

15 R. P. Feliciano, C. G. Krueger, D. Shanmuganayagam, M. M. Vestling and J. D. Reed, *Food Chem.*, 2012, **135**, 1485–1493.

16 R. L. Prior, E. Fan, H. Ji, A. Howell, C. Nio, M. J. Paynef and J. Reed, *J. Sci. Food Agric.*, 2010, **90**, 1473–1478.

17 M. M. Aung and Y. S. Chang, *Food Control*, 2014, **39**, 172–184.

18 L. Puigventós, M. Navarro, É. Alechaga, O. Núñez, J. Saurina, S. Hernández-Cassou and L. Puignou, *Anal. Bioanal. Chem.*, 2015, **407**, 597–608.

19 L. Parets, É. Alechaga, O. Núñez, J. Saurina, S. Hernández and L. Puignou, *Anal. Methods*, 2016, **8**, 4363–4378.

20 P. Lucci, J. Saurina and O. Núñez, *TrAC Trends Anal. Chem.*, 2016, **88**, 1–24.

21 N. P. Kalogiouri, R. Aalizadeh and N. S. Thomaidis, *Food Chem.*, 2018, **256**, 53–61.

22 A. M. Knolhoff, J. A. Zweigenbaum and T. R. Croley, *Anal. Chem.*, 2016, **88**, 3617–3623.

23 S. Barbosa, N. Pardo-Mates, M. Hidalgo-Serrano, J. Saurina, L. Puignou and O. Nuñez, *J. Agric. Food Chem.*, 2018, **66**, 9353–9365.

24 S. Esslinger, J. Riedl and C. Fauhl-Hassek, *Food Res. Int.*, 2014, **60**, 189–204.

25 J. Riedl, S. Esslinger and C. Fauhl-Hassek, *Anal. Chim. Acta*, 2015, **885**, 17–32.

26 L. Cuadros-Rodríguez, C. Ruiz-Samblás, L. Valverde-Som, E. Pérez-Castaño and A. González-Casado, *Anal. Chim. Acta*, 2016, **909**, 9–23.

27 M. M. Koek, R. H. Jellema, J. van der Greef, A. C. Tas and T. Hankemeier, *Metabolomics*, 2011, **7**, 307–328.

28 J. P. Antignac, F. Courant, G. Pinel, E. Bichon, F. Monteau, C. Elliott and B. Le Bizec, *TrAC - Trends Anal. Chem.*, 2011, **30**, 292–301.

29 A. Alonso, S. Marsal and A. Julià, *Front. Bioeng. Biotechnol.*, 2015, **3**, 1–20.

30 L. Yi, N. Dong, Y. Yun, B. Deng, D. Ren, S. Liu and Y. Liang, *Anal. Chim. Acta*, 2016, **914**, 17–34.

31 S. Liu, Y. Z. Liang and H. tao Liu, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, 2016, **1015**–**1016**, 82–91.

32 SOLO software, Eigenvector Research, http://eigenvector.com/software/solo.htm (Accessed on 10th March 2019).

33 J. Massart, D. L. Vandeginste, B. G. M. Buydens, L. M. C. de Jong, S. Lewi and P. J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics, J. Chem. Inf. Comput. Sci.*, 1997.

**Figure captions**

**Fig. 1.** (a) UHPLC-HRMS total ion chromatograms of the four fruit sample types analyzed in negative ionization mode. (b) HRMS full scan spectra (*m/z* 100-1500) obtained for the four fruit sample types obtained at the retention time of 6.61 min.

**Fig. 2.** PCA score plots of (a) PC2 *vs* PC4 when using UHPLC-HRMS fingerprints of all the analyzed samples (fruit-based products and pharmaceutical preparations) in positive H-ESI mode, and (b) PC1 *vs* PC3 when data was acquired in negative H-ESI mode.

**Fig. 3.** PCA score plots of (a) PC1 *vs* PC2 when using UHPLC-HRMS fingerprints of natural fruit-based products in positive H-ESI mode, and (b) PC1 *vs* PC3 when data was acquired in negative H-ESI mode.

**Fig. 4.** 3D-scatter plots of scores of LV1 *vs* LV2 *vs* LV3 for the analyzed samples when using UHPLC-HRMS fingerprints in (a) H-ESI negative acquisition mode, (b) H-ESI positive acquisition mode, and (c) data fusion of both H-ESI negative and positive acquisition modes as sample chemical descriptors.

**Fig. 5.** PLS-DA plots of scores projected on LV1 *vs* LV2 as well as classification plots for cranberry *vs* (a) raspberry, (b) blueberry, and (c) grape, respectively. Dark and clear symbols correspond to calibration and validation sets, respectively. Dashed lines indicated the classification thresholds separating the classes.

**Fig. 6.** Quantitation of (a) grape-, (b) blueberry-, and (c) raspberry-fruit percentages on cranberry-fruit extracts adulterated when using UHPLC-HRMS fingerprints in negative H-ESI mode by PLS.

Figure 1

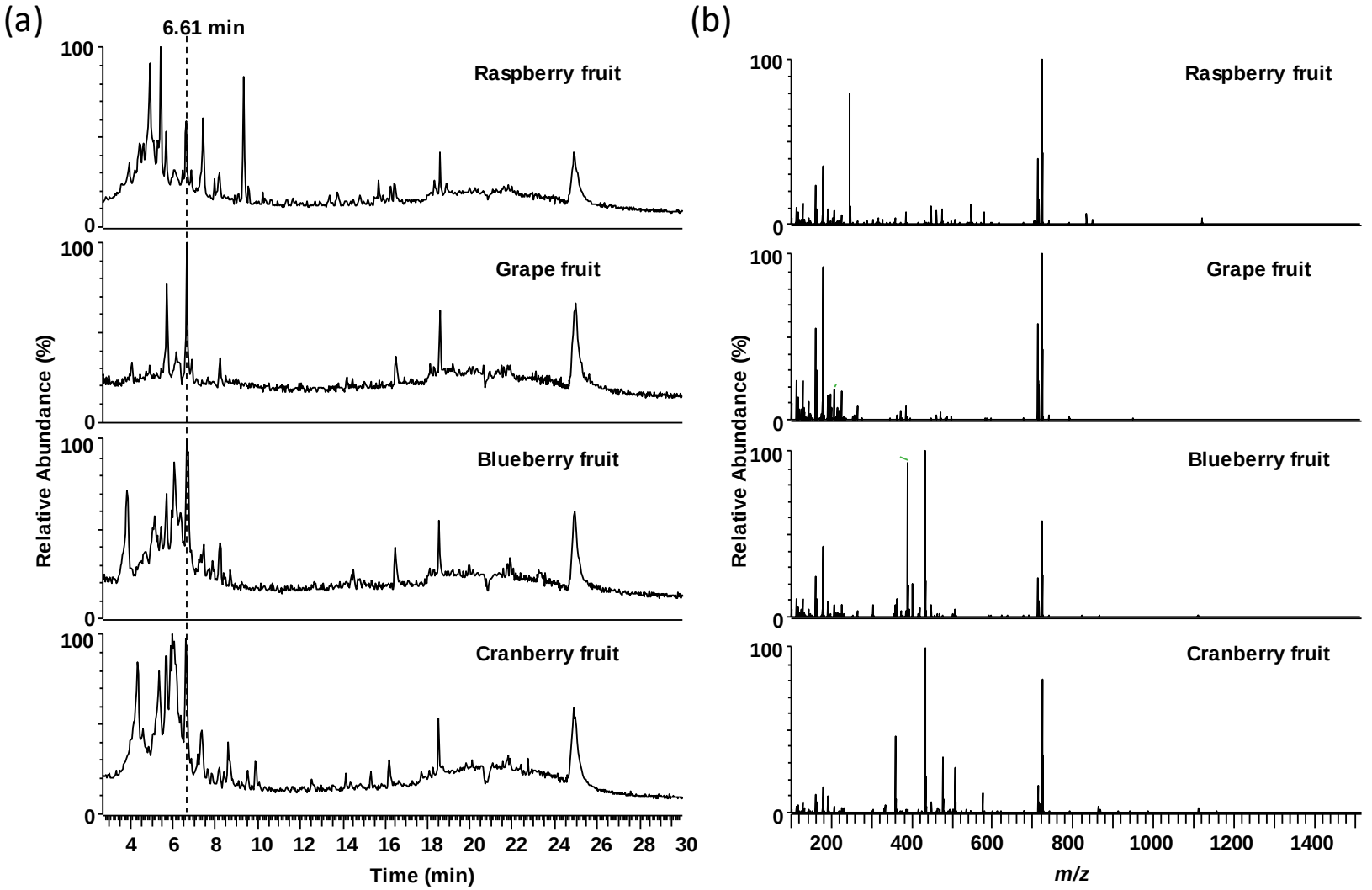
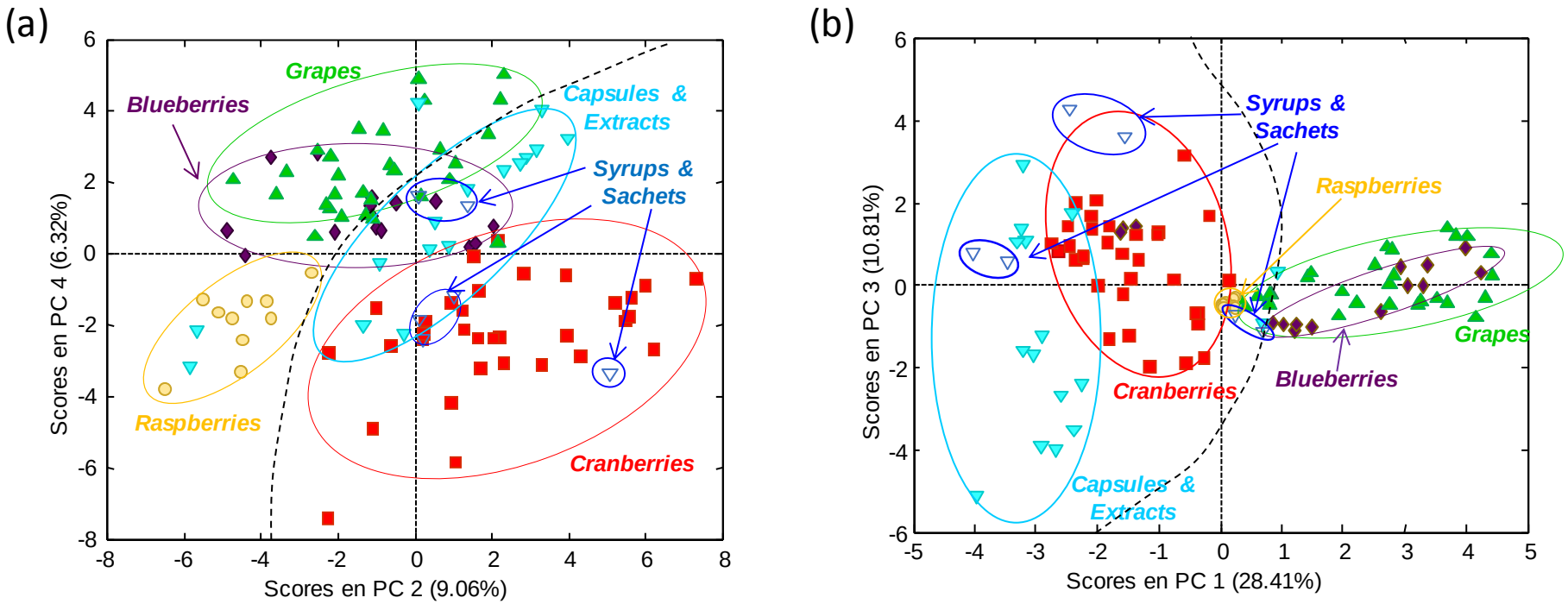

(a)
6.61 min
Raspberry fruit
Grape fruit
Blueberry fruit
Cranberry fruit
Relative Abundance (%)
100
0
4 6 8 10 12 14 16 18 20 22 24 26 28 30
Time (min)
(b)
Raspberry fruit
Grape fruit
Blueberry fruit
Cranberry fruit
Relative Abundance (%)
100
0
200 400 600 800 1000 1200 1400
m/z

Figure 2




- Cranberry-based natural products
- Cranberry-based pharmaceutical capsules & extracts
- Cranberry-based pharmaceutical syrups & sachets
- Grape-based natural products
- Blueberry-based natural products
- Raspberry-based natural products

Figure 3

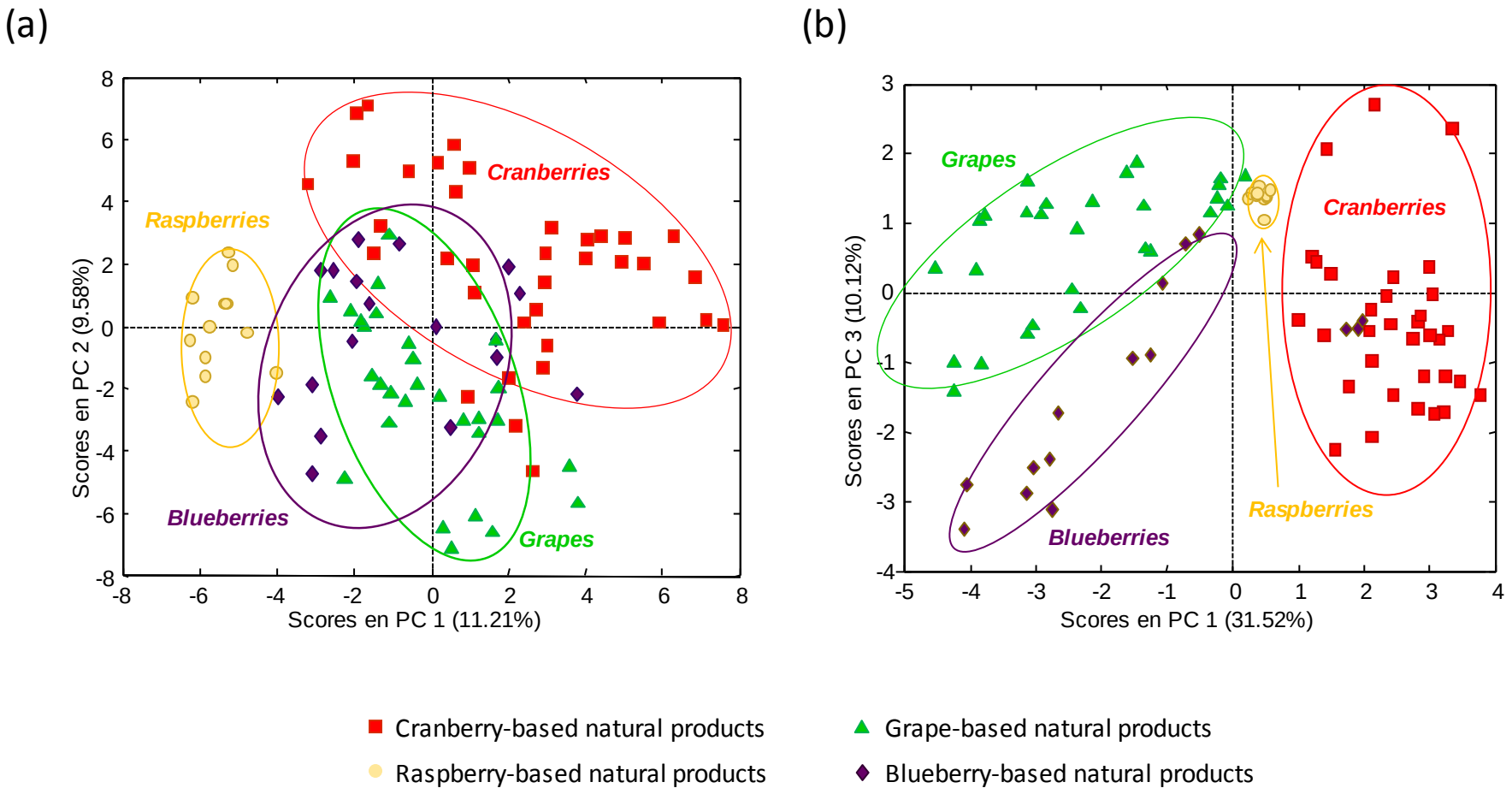

(a)
(b)
Cranberries
Raspberries
Blueberries
Grapes
Scores en PC 2 (9.58%)
Scores en PC 1 (11.21%)
Scores en PC 3 (10.12%)
Scores en PC 1 (31.52%)
Cranberry-based natural products
Raspberry-based natural products
Grape-based natural products
Blueberry-based natural products

Figure 4

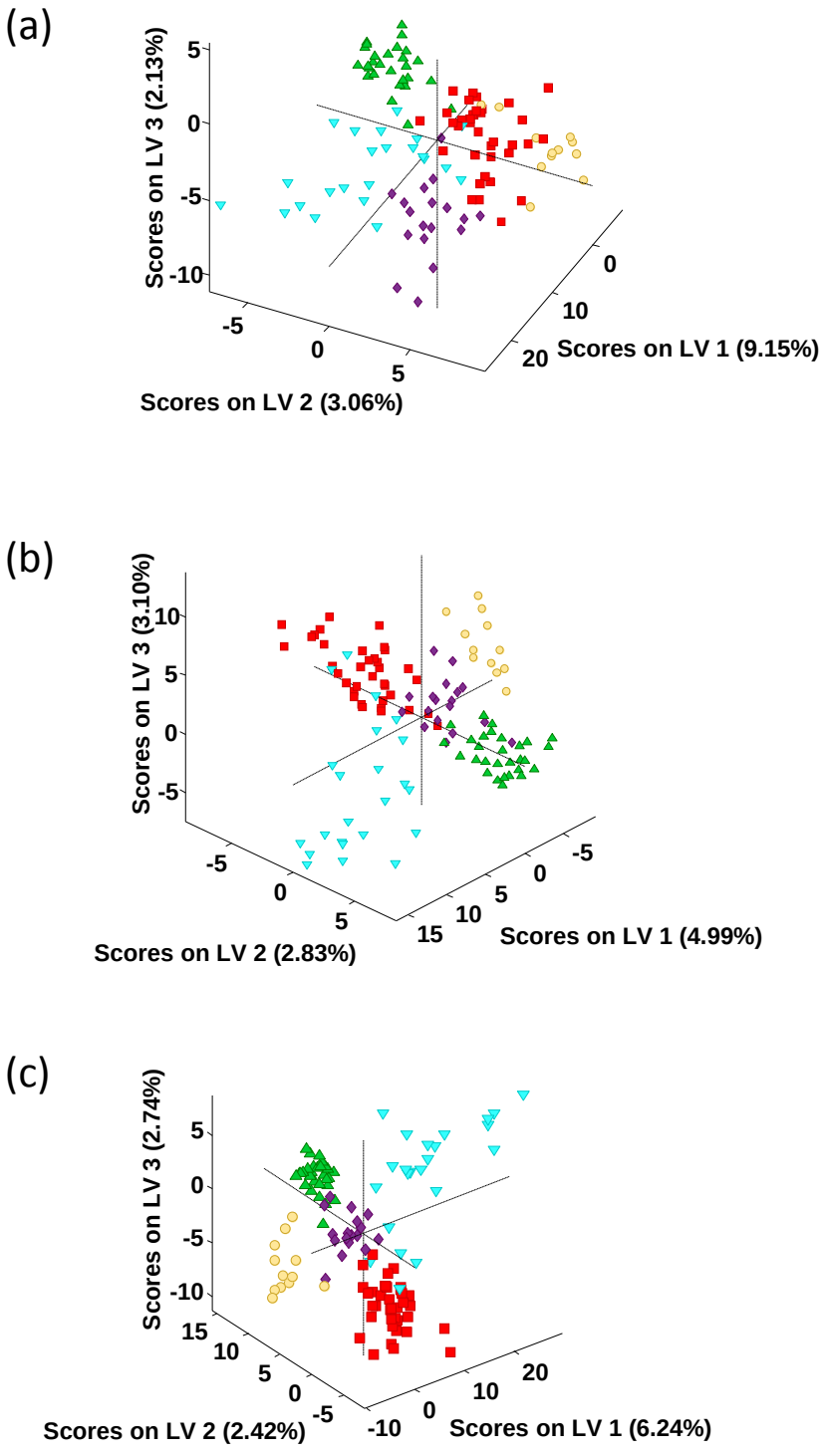

(a)
Scores on LV 3 (2.13%)
Scores on LV 2 (3.06%)
Scores on LV 1 (9.15%)
(b)
Scores on LV 3 (3.10%)
Scores on LV 2 (2.83%)
Scores on LV 1 (4.99%)
(c)
Scores on LV 3 (2.74%)
Scores on LV 2 (2.42%)
Scores on LV 1 (6.24%)
Cranberry-based natural products
Cranberry-based pharmaceutical products
Grape-based natural products
Blueberry-based natural products
Raspberry-based natural products

Figure 5

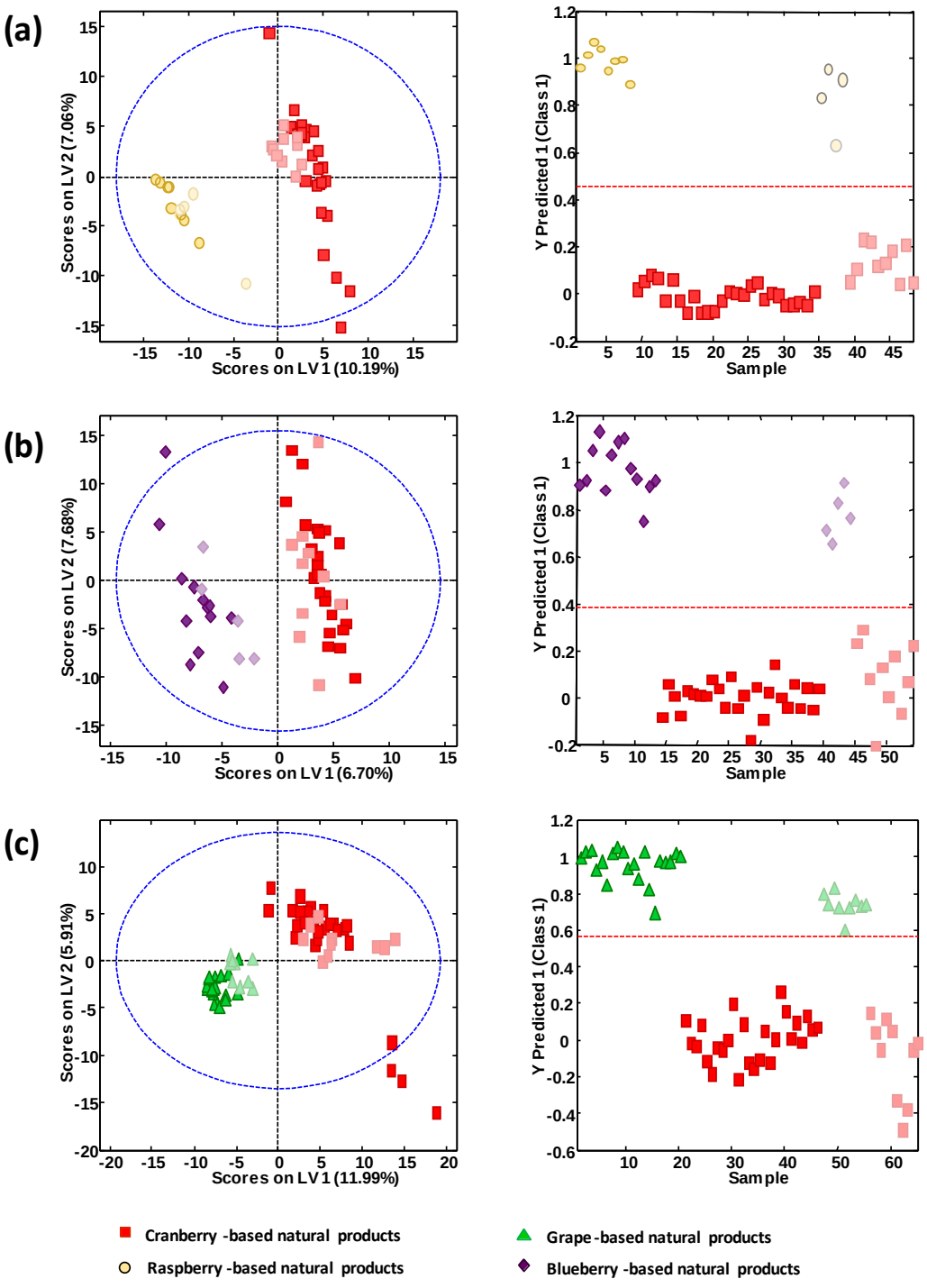
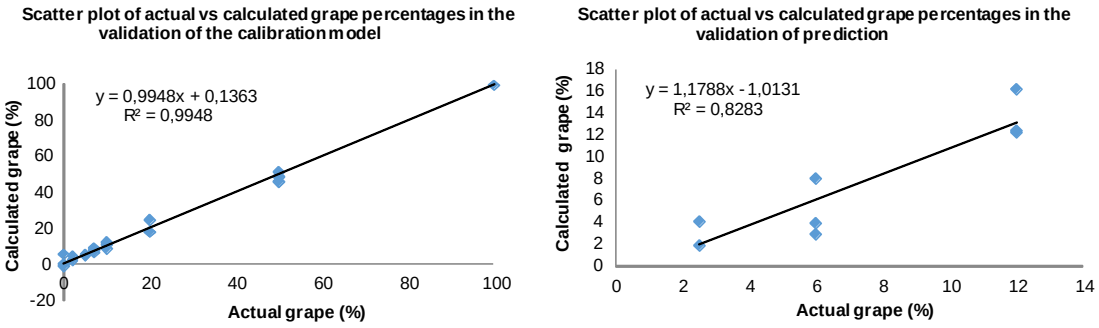
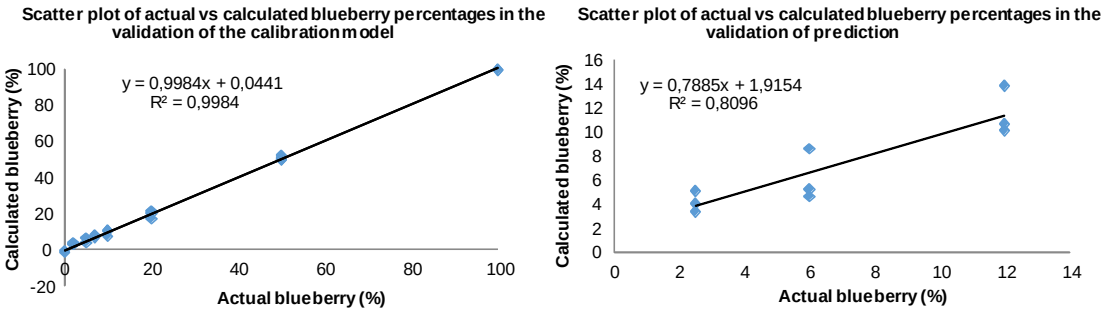

(a)
Scores on LV 2 (7.06%)
Scores on LV 1 (10.19%)
Y Predicted 1 (Class 1)
Sample
(b)
Scores on LV 2 (7.68%)
Scores on LV 1 (6.70%)
Y Predicted 1 (Class 1)
Sample
(c)
Scores on LV 2 (5.91%)
Scores on LV 1 (11.99%)
Y Predicted 1 (Class 1)
Sample
Cranberry -based natural products
Raspberry -based natural products
Grape-based natural products
Blueberry -based natural products

Figure 6

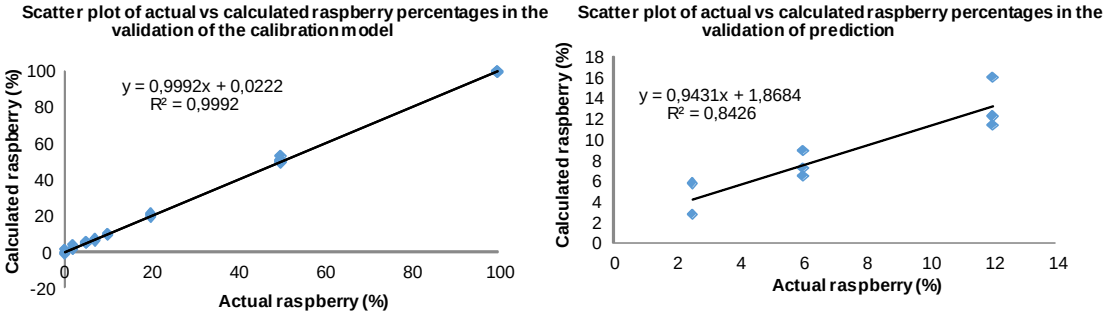
**(a)**


Scatter plot of actual vs calculated grape percentages in the validation of the calibration model
y = 0,9948x + 0,1363
R² = 0,9948
Calculated grape (%)
Actual grape (%)
Scatter plot of actual vs calculated grape percentages in the validation of prediction
y = 1,1788x - 1,0131
R² = 0,8283
Calculated grape (%)
Actual grape (%)

**(b)**


Scatter plot of actual vs calculated blueberry percentages in the validation of the calibration model
y = 0,9984x + 0,0441
R² = 0,9984
Calculated blueberry (%)
Actual blueberry (%)
Scatter plot of actual vs calculated blueberry percentages in the validation of prediction
y = 0,7885x + 1,9154
R² = 0,8096
Calculated blueberry (%)
Actual blueberry (%)

**(c)**


Scatter plot of actual vs calculated raspberry percentages in the validation of the calibration model
y = 0,9992x + 0,0222
R² = 0,9992
Calculated raspberry (%)
Actual raspberry (%)
Scatter plot of actual vs calculated raspberry percentages in the validation of prediction
y = 0,9431x + 1,8684
R² = 0,8426
Calculated raspberry (%)
Actual raspberry (%)

## Electronic Supplementary Information

**Fig. S1.** Structure of a trimeric proanthocyanidin with A-type and B-type linkages.

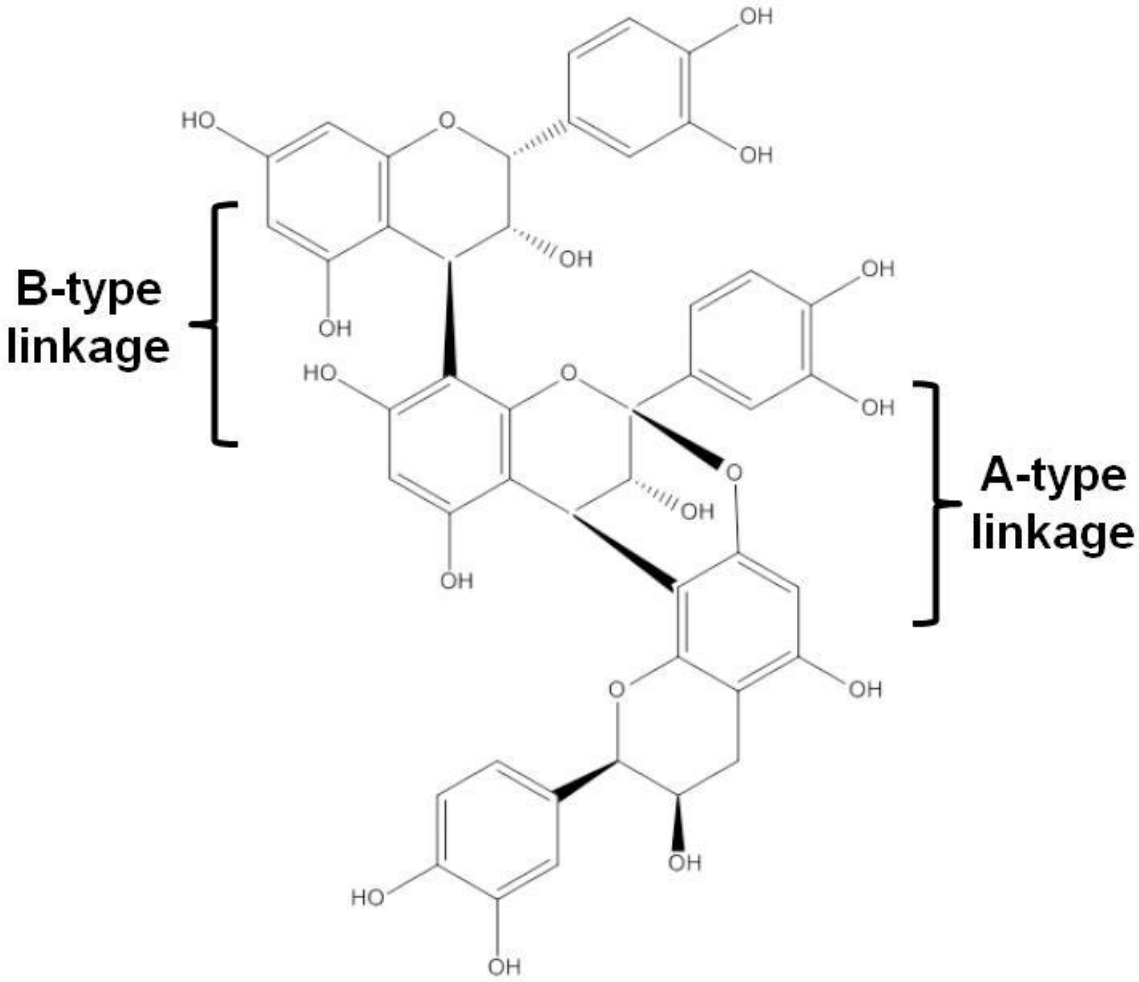


**Fig. S2.** Latent variable number *vs* CV classification error average plots for the built PLS-DA models of: (a) cranberry *vs* raspberry, (b) cranberry *vs* blueberry, and (c) cranberry *vs* grape.