Characterization, classification and authentication of fruit-based extracts by means of HPLC-UV chromatographic fingerprints, polyphenolic profiles and chemometric methods.

Naiara Pardo[a], Alex Vera[a], Sergio Barbosa[a], Míriam Hidalgo-Serrano[a], Oscar Núñez[a,b,c,*], Javier Saurina[a,b], Santiago Hernández-Cassou[a,b] and Lluis Puignou[a,b]

[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, Recinte Torribera, Av. Prat de la Riba 171, Edifici de Recerca (Gaudí), E-08921 Santa Coloma de Gramenet, Barcelona, Spain.

[c] Serra Húnter 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

HPLC-UV was applied to the analysis and characterization of fruit-based and fruit-processed products. A Kinetex C18 reversed-phase column was proposed under gradient elution for the determination of 17 polyphenols. Acceptable sensitivity (LODs below 0.16 mg/L), and good linearity (r^2 higher then 0.995), precision (RSD below 6.8%), and method trueness (relative errors below 11%) were obtained. Data corresponding to polyphenolic peak areas and HPLC-UV chromatographic fingerprints were then analyzed by exploratory principal component analysis (PCA) to extract information of the most significant variables contributing to characterization and classification of analyzed samples regarding the fruit of origin. HPLC-UV chromatographic data was further treated by partial least square (PLS) regression to determine the percentages of adulteration in cranberry-fruit extracts. It was found that even mixture samples containing low percentages of adulterants could be distinguished from genuine cranberry extracts. Highly satisfactory results were obtained, with overall errors in the quantification of adulterations below 4.3%.

Keywords: high performance liquid chromatography; UV-detection; polyphenols; principal component analysis; partial least square regression; food authentication

1. Introduction

The consumption of berry fruits associated with their contribution to improve human health because of their content on polyphenols, especially anthocyanins, is a subject of considerable interest (Basu, Rhone & Lyons, 2010; Seeram, 2008; Seeram, 2012). They contain several dietary constituents essential for human health such as fiber and vitamins (C and E), as well as bioactive phytochemicals (plant compounds that provide health benefits beyond basic nutrition) such as polyphenols and phenolic acids (Basu, Rhone & Lyons, 2010).

Lately, food products and nutraceuticals prepared with American red cranberries (*Vaccinium macrocarpon*) are gaining importance in our society due to some healthy effects on humans, including antioxidant activity, antimicrobial activity against bacteria involved in a wide range of diseases, antiinflammatory activity in periodontal disease, and antiproliferative activity on human oral, colon, and prostate cancer cell lines, among others (Sanchez-Patan, Bartolome, Martin-Alvarez, Anderson, Howell & Monagas, 2012). These healthy effects are attributed to their high content on specific polyphenols, although their most noticeable bioactivity deals with their capacity to inhibit the adhesion of pathogenic bacteria to uroepithelial cells of the urinary tract, thus preventing urinary tract infections (Feliciano, Krueger & Reed, 2015; Feliciano, Meudt, Shanmuganayagam, Krueger & Reed, 2014; Howell, Reed, Krueger, Winterbottom, Cunningham & Leahy, 2005; Nicolosi, Tempera, Genovese & Furneri, 2014; Patel, Scarano, Kondo, Hurta & Neto, 2011). The most common polyphenols found in cranberries are hydroxycinnamic and hydroxybenzoic acids, and flavonoids such as anthocyanins, flavonols, and flavan-3-ols (Borges, Degeneve, Mullen & Crozier, 2010; Diaz-Garcia, Obon, Castellar, Collado & Alacid, 2013; Howell, Reed, Krueger, Winterbottom, Cunningham & Leahy, 2005). In particular, flavan-3-ols (catechins and epicatechins) occur in cranberry in both monomeric and polymeric forms (i.e., proanthocyanidins, PACs). PACs are often classified according to the interflavan linkage as A-type and B-type molecules. B-type PACs are those in which monomeric units are linked through the C4 position of the upper unit and the C6 or C8 positions of the lower unit. In contrast, A-type PACs contain an additional ether-type bond between the C2 position of the upper unit and the hydroxyl group at C5 or C7 positions of the lower unit (C2–O–C5 or C2–O–C7). In general, 60% of PACs in cranberry are A-type ones (Gu, Kelm, Hammerstone, Beecher, Holden, Haytowitz, Gebhardt & Prior, 2004), while B-type PACS are predominantly found in other food products like tea, chocolate, blueberry or grapes. The most important difference between the two families of PCAs is that only the A-type PACs are capable of inhibiting the adhesion of bacteria to urinary tract issues

(Feliciano, Krueger & Reed, 2015; Feliciano, Meudt, Shanmuganayagam, Krueger & Reed, 2014; Krueger, Reed, Feliciano & Howell, 2013).

Nowadays, several concerns have arisen on some of the products sold in the market labeled as derived from American red cranberry extracts that may contain other more economic fruit extracts which do not provide the desired bioactivity to promote health beneficial effects (Krueger, 2015). Therefore, the prevention of this kind of frauds becomes an issue of great importance in our society, and the development of simple and reliable analytical methodologies able to classify and characterize natural extracts to achieve the correct authentication regarding the fruit of origin is necessary.

Several analytical methodologies have been proposed for the determination of polyphenols and phenolic acids in fruit products and pharmaceutical preparations. In general, a rough estimation of overall contents can be assessed by simple colorimetric methods. For example, a sensitive colorimetric assay able to tackle the total content on PACs is based on the reaction of these compounds with 4-dimethylaminociannamaldehyude (DMAC) (Feliciano, Shea, Shanmuganayagam, Krueger, Howell & Reed, 2012; Prior, Fan, Ji, Howell, Nio, Payne & Reed, 2010). However, this method is not capable of differentiating between A- and B-type PACs (Krueger, 2015). With this knowledge, the unscrupulous sellers can “spike” products with the lowest cost PAC source, and still provide specification (PAC levels) that buyers find acceptable.

Taking into account that polyphenol and phenolic acid content seems to be related to food features such as geographical areas, variety and manufacturing practices, etc., the contents of other less expensive polyphenols and phenolic acids compared to PACs can also be exploited as a source of analytical data to establish classification and characterization of fruit products (Saurina & Sentellas, 2015). Liquid chromatography (LC) with UV detection or coupled to mass spectrometry (LC-MS) are the most common techniques described for the determination of polyphenols and the characterization of a great variety of plants and fruit-based products (Alonso-Salces, Ndjoko, Queiroz, Ioset, Hostettmann, Berrueta, Gallo & Vicente, 2004; Engstrom, Palijarvi, Fryganas, Grabber, Mueller-Harvey & Salminen, 2014; Furuuchi, Yokoyama, Watanabe & Hirayama, 2011; Hamed, Al Ayed, Moldoch, Piacente, Oleszek & Stochmal, 2014; Navarro, Núñez, Saurina, Hernández-Cassou & Puignou, 2014; Parets, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2016; Puigventós, Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015; Rzeppa, Von Bargen, Bittner & Humpf, 2011; Wallace & Giusti, 2010).

The aim of the present work was to develop a simple, less expensive, and reliable high performance liquid chromatography method with UV-detection (HPLC-UV) for the determination of polyphenolic profiles in the analysis of fruit-based products. For that purpose, a total of 17

polyphenolic compounds belonging to different families (stilbenes, phenolic acids, flavonoids) were selected. A simple and cheap sample treatment, consisting of an extraction by sonication with acetone:water:hydrochloric acid (70:29.9:0.1 *v/v/v*) and centrifugation, was applied to the analysis of different kinds of cranberry-, grape-, blueberry-, and raspberry-based samples, including fruits, fruit juices, and raisins. Specific sample purification steps focused on the isolation of proanthocyaninds by employing sephadex sorbent (Navarro, Núñez, Saurina, Hernández-Cassou & Puignou, 2014) were prevented in order to reduce the cost of the proposed method and make it more applicable to any laboratory. Data corresponding to the polyphenolic composition as well as the HPLC-UV chromatographic fingerprints were considered as a source of potential descriptors to be exploited for the classification and characterization of fruit-based products by exploratory principal component analysis (PCA). Finally, cranberry-fruit extracts were adulterated with different amounts (2% to 50%) of grape, blueberry, or raspberry fruit extracts, and the polyphenolic profile and chromatographic fingerprinting data was evaluated for authentication purposes as well as the quantification of adulteration content by means of partial least squares (PLS) regression.

2. Materials and Methods

2.1. Chemicals

Unless specified, analytical grade reagents were always used. The polyphenols and phenolic acids studied (gallic acid, homogentistic acid, protocatechuic acid, protocatechualdehyde, (+)-catechin hydrate, gentisic acid, *p*-salicilic acid, chlorogenic acid, vanillic acid, (-)-epicatechin, syringic acid, syringaldehyde, ethyl gallate, *p*-coumaric acid, ferulic acid, resveratrol and quercitrin hydrate), whose structures and CAS numbers are shown in Table 1S (supplementary material), were all of them obtained from Sigma-Aldrich (Steinheim, Germany). Stock standard solutions of all polyphenols and phenolic acids (ca. 1,000 mg/L) were prepared in methanol in amber glass vials. Intermediate working solutions were prepared weekly from these stock standard solutions by appropriate dilution with Milli-Q water. All stock solutions were stored at 4 ºC for no more than 1 month. Methanol (Chromosolv® for HPLC, ≥99.9%), acetone and formic acid (≥98%) were also obtained from Sigma-Aldrich: Hydrochloric acid (25 o 33%) was provided from Merck (Seelze, Germany).

Water was purified using an Elix 3 coupled to a Milli-Q system (Millipore, Bedford, MA, USA) and filtered through a 0.22 µm nylon membrane integrated into the Milli-Q system.

144 *2.2. Instrumentation and methods*

145 The analysis of polyphenols and phenolic acids was performed on a Varian HPLC system
146 (California, USA) equipped with a ProStar 240 SDM ternary pump, a ProStar 430 Autosampler and
147 a ProStar 334 photodiode array (PDA) detector. Instrument control and data processing were
148 carried out with the System Control 6.3 software. Separation was performed in reversed-phase
149 mode by using a Kinetex C_{18} (100×4.6 mm i.d., 2.6 µm particle size) column from Phenomenex
150 (California, USA) at room temperature following a previously described method (Puigventós,
151 Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015). Gradient separation
152 using 0.1% formic acid in water (v/v) (solvent A) and methanol (solvent B) as mobile phases was
153 as follows: 0–3 min, linear gradient from 5 to 25 % B; 3–6 min, at 25 % B; 6–9 min, from 25 to 37
154 % B; 9–13 min, at 37 % B; 13–18 min, from 37 to 54 % B; 18–22 min, at 54%B; 22–26 min, from
155 54 to 95%B; 26–29 min, at 95 % B; 29–29.15 min, back to initial conditions at 5 % B; and from
156 29.15 to 36 min, at 5 % B. The mobile phase flow rate was 1 mL/min and the injection volume was
157 10 µL. PDA acquisition from 190 to 550 nm was performed to register UV-spectra and to
158 guarantee peak purity. For quantitation purposes on the 17 targeted polyphenols and phenolic acids,
159 direct UV absorption detection was employed at 280 nm (gallic acid, homogentistic acid,
160 protocatechualdehyde, (+)-catechin hydrate, (-)-epicatechin, syringic acid and ethyl gallate), 257
161 nm (protocatechuic acid, *p*-salicylic acid, vanillic acid and quercitrin hydrate) and 316 nm (gentistic
162 acid, chlorogenic acid, syringaldehyde, *p*-coumaric acid, ferulic acid and resveratrol).

163 HPLC-UV chromatographic fingerprints were obtained with an Agilent 1100 Series HPLC
164 instrument equipped with a G1311A quaternary pump, a G1379A degasser, a G1392A
165 autosampler, a G1315B diode-array detector and a PC with the Agilent Chemstation software (Rev.
166 A 10.02), all of them from Agilent Technologies (Waldbronn, Germany). Separation column,
167 chromatographic conditions were the same as previously described with the Varian HPLC system.
168

169 *2.3. Samples and sample treatment*

170 A total of 86 fruit-based samples, purchased from Barcelona markets, were analyzed. The
171 samples included 29 cranberry-based products (4 fruit samples, 10 raisin samples and 15 juice
172 samples), 27 grape-based products (4 fruit samples, 8 raisin samples and 15 juice samples), 18
173 blueberry-based products (6 fruit samples and 12 juice samples), and 12 raspberry-based fruit
174 samples.

175 All fruits and raisins were grinded using an Ike Ultra-Turrax machine (Staufen, Germany)
176 with different applicators. Water was added to raisins to improve the crushing. Then, all analyzed
177 samples were freeze-dried to achieve fully lyophilized products. To this end, 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.

Sample treatment was then carried out following a previously described method with some modifications (Navarro, Núñez, Saurina, Hernández-Cassou & Puignou, 2014; Puigventós, Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015). Briefly, 0.1 g of 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 minutes. 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 to evaluate the repeatability of the method and the robustness of the chemometric results.

For authentication studies by PLS regression, three cases were studied 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. This series of extracts was 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 other fruits were 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.

2.4. Data analysis

SOLO from Eigenvector Research was used for calculations with principal component analysis (PCA) and partial least square (PLS) regression (http://www.eigenvector.com/software/solo.htm). A detailed description of the theoretical background of this method is given elsewhere (Massart, 1997).

Data matrices to be treated by PCA consisted on (i) the peak area values of the 17 studied polyphenols and phenolic acids detected in the different samples under study and (ii) the HPLC-UV chromatographic profiles obtained at different acquisition wavelengths (257, 280, 316, 420 and 500 nm). In the first case, the dimension of the matrix was 86 samples×17 analytes. Normalization pretreatment with respect to the overall polyphenolic concentration was applied to provide similar weighs to all the samples. In the second case, HPLC-UV chromatograms were pretreated to improve the data quality while minimizing solvent and matrix interferences, peak shifting and baseline drifts. For additional details see (Pérez-Rafols & Saurina, 2015). Scatter plots of scores

and loadings of the principal components (PCs) were used to investigate the structure of maps of samples and variables, respectively.

Peak identification in the analyzed samples was performed by comparison of retention times and UV spectra with those of a polyphenolic standard solution. Peak purity was confirmed by comparison of UV spectra through the entire peak signal.

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. 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. For both training and test steps, X-data matrices consisted of the HPLC-UV chromatographic fingerprints of the corresponding matrices and the Y-data matrices contained the adulteration fruit-extract percentages.

3. Results and discussion

3.1. HPLC conditions

In previous works, LC-MS/MS methods for the determination of polyphenols in cranberry-based pharmaceuticals and several fruits or juice samples were established by using ESI and APPI as ionization sources and a triple quadrupole mass analyzer (Parets, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2016; Puigventós, Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015). Although a successful characterizations and classifications of the analyzed samples were achieved with the proposed methods, MS is a relatively expensive technique not available in all the laboratories focusing in food authentication problems. Moreover, in those preliminary studies the number of samples was more limited only several cranberry-based and grape-based products analyzed. For this reason, one of the main objectives of the present work was the development of an HPLC-UV method for the classification, characterization and authentication of fruits and fruit processed products, which will be a less expensive method in comparison to LC-MS/MS, and more accessible for any food control laboratory. Moreover, the number of samples was increased to include other fruits and fruit-processed products such as blueberry- and raspberry-based extracts that can also be used in the adulteration of cranberry products.

For that purpose, a total of 17 polyphenols and phenolic acids belonging to different families were selected (Table 1S, supplementary material) as target analytes, and their chromatographic separation was evaluated using the previously established separation (Puigventós, Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015). Fig. 1S (supplementary material) shows the HPLC-UV chromatogram obtained under gradient conditions (see experimental section) for a standard mixture of all the analyzed compounds at a concentration of 30 mg/L. As can be seen, an acceptable separation was obtained in less than 18 min. Only a small coelution between syringaldehyde and ethyl gallate (peaks 12 and 13) was observed, although it was considered acceptable for the intended purpose of the present work.

3.2. Instrumental quality parameters and method performance

The performance of the proposed HPLC-UV method was evaluated by determining instrumental quality parameters for the 17 polyphenols and phenolic acids analyzed and the figures of merit are given in Table 2S and Table 3S (supplementary material). Limits of detection (LODs), based on a signal-to-noise ratio of 3:1, were calculated using standard solutions at low concentration levels, and values between 0.16 mg/L (*p*-coumaric acid) and 2.90 mg/L ((-)-epicatechin) were achieved. Limits of quantitation (LOQs), based on a signal-to-noise ratio of 10:1, between 0.54 and 9.57 mg/L were obtained. Although these values are relatively higher in comparison to those achieved by LC-MS techniques (Parets, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2016; Puigventós, Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015), as expected, they were compatible with polyphenols and phenolic acids concentrations in natural fruit-based extracts are expected to be, in general, at the relatively low to high mg/L level. External calibration curves based on peak area at concentrations above LOQ to 100 mg/L were established and good linearities, with correlation coefficients (r^2) higher than 0.995 were achieved for all compounds.

Run-to-run and day-to-day precisions for migration time and compound quantification at two concentration levels, low level (LOQ) and medium level (21.9-35.5 mg/L), were calculated and the results are depicted in Table 2S (supplementary material). In order to obtain the run-to-run precision, five replicate determinations for each concentration level were carried out. Day-to-day precision was estimated from 15 replicate determinations at each concentration level on three nonconsecutive days (five replicates each day). For run-to-run precision, relative standard deviations (%RSD) in the range 0.9–3.9% were obtained at LOQ concentration levels. Lower RSD values (0.5–1.6%) were achieved at the medium concentration level, as expected. In terms of retention time, good run-to-run precisions were also obtained, with RSD values lower than 0.8% in

all cases. Very good day-to-day precision values were also obtained, although the values worsened a little in comparison to run-to-run precision, as expected, with RSDs in the ranges 2.8–6.8% and 2.5–6.1% for low and medium concentration levels, respectively. It should be mention that in terms of precision, similar results to those previously reported by employing LC-MS techniques were observed (Parets, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2016; Puigventós, Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015)

Due to the lack of any reference material containing the 17 polyphenols and phenolic acids studied, intra-day (within the same day) and inter-day (in different days) method trueness was evaluated at the two concentration levels by comparing spiked concentrations with the calculated concentrations using external calibration, and the results, expressed as the relative errors (%), are shown in Table 3S (supplementary material). As can be seen, the proposed HPLC-UV method showed, in general, a very good performance with lower relative error values for the medium concentration levels in comparison to the LOQ level, as can be expected. Regarding inter-day and intra-day trueness, very similar values were observed being intra-day slightly better, but none of the values exceeds an error of 11%, which is very acceptable for HPLC-UV methodologies.

The results obtained showed that the proposed HPLC-UV method was acceptable in terms of sensitivity, and very satisfactory in terms of precision and trueness for the determination of polyphenols and phenolic acids.

3.3. Exploratory studies by principal component analysis

Principal component analysis was used as exploratory method to study the classification of samples regarding the fruit of origin. PCA provided plots of scores and loadings, showing the distribution of the samples and variables on the principal components (PCs), respectively. The study of the plot of scores revealed patterns that may be correlated to sample characteristics, such as the type of fruit used on the extracts. The study of the distribution of variables from the plot of loadings provided information dealing with their correlations as well as dependencies of polyphenols and phenolic acids on vegetable oil properties. Both, peak area of polyphenols and phenolic acids and HPLC-UV chromatographic fingerprints were used for exploratory PCA studies.

Phenolic peak areas. First, fruit sample characterization was attempted using the peak area of the seventeen polyphenols and phenolic acids found in the analyzed samples. For that purpose, samples were processed as indicated in sample treatment section, and the final extracts were randomly analyzed with the proposed HPLC-UV method. Peak identification was achieved by comparison with retention time of standards and UV-spectra. Peak areas were used to build a data matrix with a dimension of 86 samples x 17 compounds to be subjected to PCA. Fig. 1a shows the

scatter plots of scores of PC1 *vs* PC2. As can be seen, QCs appeared in a compact group in the center area of the plot, demonstrating the good repeatability and robustness of the proposed HPLC-UV and chemometric methods. A preliminary classification of fruit samples showed that the most conflictive zone was in the center of the graph, where grape-based samples appeared mixed with some cranberry-based samples and close to the other two groups of samples (blueberry and raspberry ones). The two first principal components (PC1 and PC2) explained a 27.9% and 17.32% of the variability between samples, respectively. To corroborate the tendencies observed in PC1 *vs* PC2 plot (Fig. 1a), PC3, which retained a 15.13% of the variability between samples, was also considered, and the plot depicting PC2 *vs* PC3 is given in Fig. 1b. As can be seen, in general, the only difference is the distribution of the samples in the plot area. There were also three major zones, in which the raspberry- and blueberry-based samples were well separated, and the center area with the grape- and some cranberry-based samples. However, by considering both Fig. 1a and 2b, cranberry samples tended to display negative scores on PC2. Taking into consideration only the group of grape and cranberry samples, the PCA classification achieved up to this point is slightly worse than the one previously reported by employing the specific purification step for proanthocyanidins with sephadex sorbent (Navarro, Núñez, Saurina, Hernández-Cassou & Puignou, 2014). However, in the present work a higher number of grape samples, together with other fruit-based samples (blueberries and raspberries) were employed, and a less expensive methods was achieved. The plot of loadings (see Fig. S2 in the supplementary material) provided information on the analyzed polyphenols and phenolic acids. These figures manifest that there were several characteristic polyphenols in each group of samples while others were not discriminant at all. For example, compounds 8, 9 and 11 (chlorogenic, vanillic and syringic acids) seemed to be the most characteristic (and discriminant) for blueberry-based samples. The most relevant compounds in raspberry-based samples were signals 3, 4 and 13 (protocatechuic acid, protocatechualdehyde and ethyl gallate, respectively), and finally, for cranberry-based samples, the most significant compounds were signals 7, 14 and 15 (*p*-salicylic, *p*-coumaric and ferulic acids, respectively). Because grape-based samples appeared grouped close to the less discriminant area it is difficult to assign characteristic and/or discriminant polyphenols.

HPLC-UV chromatographic fingerprints. In a second approach, exploratory PCA characterization of the analyzed fruit-based samples was attempted by using raw chromatographic profiles (i.e., absorbance over time) as the analytical data. HPLC-UV chromatographic fingerprints were evaluated at several wavelengths: 257, 280, 316, 420 and 500 nm. Only HPLC-UV chromatographic fingerprints registered at 280 nm allowed achieving a certain distribution and

classification among analyzed samples. Fig. 2a shows the corresponding scatter plot of scores of PC1 *vs* PC2. As can be seen, certain discrimination among samples was achieved, being raspberry-based samples perfectly grouped at the top area of the plot and separated from the other groups, blueberry-based samples distributed at the bottom-right area of the plot, while no clear differentiation was obtained among cranberry- and grape-based samples, being grouped in the center-left area of the plot.

HPLC-UV chromatographic fingerprints were simplified by considering specific time segments that may contain richer information in reference to each fruit class. In a first approach, chromatographic profiles from 3 to 23 min were considered as the data (by removing the retention times corresponding to dead volume elution and gradient re-equilibration step). The scatter plot of PC1 *vs* PC2 obtained after PCA is depicted in Fig. 2b. A slightly improved sample classification in comparison to the previous one (Fig. 2a) was achieved. Sample distribution in the plot is more or less the same but they appeared more grouped among their specific fruit type. However, again, no clear discrimination among cranberry- and grape-based samples was obtained. It should be mention that the four cranberry fruit samples (CF1, CF2, CF3, and CF4) appeared completely separated from the other cranberry-based samples (raisins and juices), as in the previous experiment. This is due to the great differences in polyphenolic content among cranberry-based samples as can be seen in Fig. 3S (supplementary material) showing the segmented HPLC-UV chromatogram (from 3 to 23 min) of a cranberry fruit, raisin and juice sample. In a second approach, only the chromatographic retention time segments that were more different among the analyzed samples were considered. Thus, HPLC-UV chromatographic fingerprints by combining time segments from 4.7–6.5 min + 8–14 min + 15–17 min + 29–30 min were submitted to PCA, and the obtained results (score plot of PC1 *vs* PC2) are shown in Fig. 2b. This data simplification improved sample classification in comparison to the two previous experiments, although again a complete discrimination among cranberry- and grape-based samples was not possible Another model was built without including juices and raisins so only fruit samples were considered. Data treated by PCA corresponded to HPLC-UV chromatographic profiles segmented from 3 to 23 min. QCs considering only the fruit samples analyzed were also employed. The obtained results (score plot of PC1 *vs* PC2) are given in Fig. 2d. As can be seen, QCs appeared grouped in the center are of the plot showing the good repeatability and robustness of the HPLC-UV and chemometric methods employed. Regarding fruit samples, a very good distribution was observed, being raspberry fruits grouped in the center-top area of the plot, while the other samples appeared at the bottom of the plot, grape to the left, blueberry in the center and cranberry to the right area. In contrast to the results observed when employing phenolic peak area, the present developed method employing

HPLC-UV fingerprinting improved the PCA classification in comparison to the ones previously reported using sephadex purification of proanthocyanidins (Navarro, Núñez, Saurina, Hernández-Cassou & Puignou, 2014) and even LC-MS/MS methods (Parets, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2016; Puigventós, Navarro, Alechaga, Núñez, Saurina, Hernández-Cassou & Puignou, 2015). Taking into account these results, fruit samples were employed for the adulteration studies carried out.

3.3. Adulteration studies by partial least squares regression

PLS was employed to quantify the grape, blueberry or raspberry percentage of adulteration in the cranberry fruit extracts under study. All results discussed here corresponded to HPLC-UV chromatographic fingerprints recorded at 280 nm. It should be mention that models using chromatograms recorded at the other wavelengths were also investigated but, again, the most satisfactory PLS results were obtained at 280 nm. The PLS model was first established on the data set of calibration as indicated in the experimental section. On a first approach, the time window from 3 to 23 min was selected for PLS. The number of latent variables (LV) to be used for the assessment of the model was estimated by venetian blinds cross validation method, considering 6 data splits. PLS results obtained for the study of cranberry-fruit extracts adulterated with raspberry-fruit extracts are shown in Fig. 3. As can be seen (Fig. 3a), the lowest prediction error was attained with 4 LV in this particular example although, in general, the optimal number of latent variables ranged from 4 to 6 depending on the case. Fig. 3b depicts the scatter plot of scores on LV1 and LV2 showing the distribution of analyzed samples in agreement with the raspberry adulterated contents. The performance of predictions of raspberry percentages in both calibration and prediction steps was evaluated under the selected model conditions. Training and test results are depicted in Figs. 3c and 3d, respectively, and the obtained errors for both calibration and prediction steps are summarized in Table 1. The agreement between actual and predicted values was highly satisfactory. In the case of the test set, a prediction error of 4.65% was in the case of study. Better results were even achieved when blueberry-, and grape-fruit extracts were used as adulterants (see PLS results in Figs. S4 and S5 on the supplementary material for the adulteration with blueberry and grape, respectively). Overall prediction errors below 2.5% were obtained (Table 1).

In a second approach, in order to see if results can be improved at low adulteration concentrations, a PLS model employing only low adulteration levels was also considered. For that purpose, 100% cranberry-fruit extract samples, 50% adulterant extract samples, and 100% adulterant-fruit samples were removed from the calibration set, and the segmented HPLC-UV chromatographic fingerprints from 3 to 23 min obtained for the other samples were subjected to

PLS. The results are shown in Fig. 4 and in Figs. S6 and S7 (supplementary material) for raspberry-, blueberry-, and grape-fruit extracts used as adulterants, and the prediction errors obtained are also summarized in Table 1. Although calibration errors worsened slightly (but being lower than 1.71%), prediction errors improved when raspberry- and grape-fruits were used as adulterants extracts. In contrast, prediction errors worsened for the case of adulteration with blueberry. Anyway, overall prediction errors were always very satisfactory with values below 4.26%.

Finally, as the HPLC-UV chromatographic fingerprints of raspberry-fruit samples are quite characteristic in comparison to the ones observed for cranberry-fruit samples (see, as example, the HPLC-UV chromatograms from 3 to 23 min for each one of the analyzed fruits in Fig S8 of the supplementary material), a PLS model considering only several more specific HPLC-UV chromatographic time segments was evaluated when raspberry was used as adulterant fruit. Thus, HPLC-UV chromatographic fingerprints combining time segments from 5.4–6.3 min + 9.1–13.2 min + 16.2-16.4 min of the cranberry-fruit samples adulterated with raspberry extracts were submitted to PLS, and the results are shown in Fig. 5. Prediction errors are also summarized in Table 1. An important improvement on adulteration quantitation was observed, with a reduction on prediction errors below 2.03%. These results show that for some specific adulterants the simplification of HPLC-UV chromatographic fingerprints employing discriminant time segments may improve the identification of frauds.

4. Conclusions

A simple and feasible HPLC-UV method was developed for the determination of seventeen polyphenols and phenolic acids, showing acceptable LOD and LOQ values, good linearity, run-to-run and day-to-day precisions, and inter-day and intra-day method trueness.

Both peak areas of targeted compounds and chromatographic fingerprints recorded at various wavelengths were used as the analytical data to be further treated chemometrically. Exploratory PCA on phenolic peak areas provided a reasonable sample classification regarding the kind of fruit involved. The discrimination among samples improved when HPLC-UV chromatographic fingerprints were employed as this data resulted in richer source of discriminant features. The best characterization and classification of samples was observed when combining HPLC-UV chromatographic fingerprints at different time segments (4.7–6.5 min + 8–14 min + 15–17 min + 29–30 min), although still cranberry- and grape-based samples appeared grouped quite close. When the data set under study was reduced to fruit-based samples were considered for exploratory PCA, a very good characterization and classification of samples regarding the fruit of

origin was observed when employing HPLC-chromatographic fingerprints segmented from 3 to 23 min. Taking into account these results, fruit samples were considered to carry out further authentication studies focused on the quantitation of frauds.

The adulteration of cranberry fruit extracts with raspberry, blueberry or grape fruit extracts was here studied. The percentage of raspberry, blueberry or grape added as adulterant to the cranberry extracts was determined by multivariate calibration using PLS. Overall prediction errors in the quantitation of fruit adulterant percentage even at very low amounts (2%) were below 4.3%, showing that the proposed HPLC-UV method in combination with multivariate calibration was a simple and suitable strategy for the identification of frauds and to guarantee authentication of cranberry-based extracts employed for the production of cranberry-based pharmaceuticals and nutraceuticals.

Acknowledgements

The authors gratefully acknowledge the financial support received from Spanish Ministry of Economy and Competitiveness under the projects CTQ2014-65324 and CTQ2015-63968-C2-1-P, and from the Agency for Administration of University and Research Grants (Generalitat de Catalunya, Spain) under the projects 2014 SGR-377 and 2014 SGR-539.

Conflict of interest

The authors declare that they have no competing interests.

References

Alonso-Salces, R.M., Ndjoko, K., Queiroz, E.F., Ioset, J.R., Hostettmann, K., Berrueta, L.A., Gallo, B., & Vicente, F. (2004). On-line characterisation of apple polyphenols by liquid chromatography coupled with mass spectrometry and ultraviolet absorbance detection. *Journal of Chromatography A, 1046,* 89-100.

Basu, A., Rhone, M., & Lyons, T.J. (2010). Berries: emerging impact on cardiovascular health. *Nutrition Reviews, 68,* 168-177.

Borges, G., Degeneve, A., Mullen, W., & Crozier, A. (2010). Identification of Flavonoid and Phenolic Antioxidants in Black Currants, Blueberries, Raspberries, Red Currants, and Cranberries. *Journal of Agricultural and Food Chemistry, 58,* 3901-3909.

Diaz-Garcia, M.C., Obon, J.M., Castellar, M.R., Collado, J., & Alacid, M. (2013). Quantification by UHPLC of total individual polyphenols in fruit juices. *Food Chemistry, 138,* 938-949.

Engstrom, M.T., Palijarvi, M., Fryganas, C., Grabber, J.H., Mueller-Harvey, I., & Salminen, J.P. (2014). Rapid Qualitative and Quantitative Analyses of Proanthocyanidin Oligomers and Polymers by UPLC-MS/MS. *Journal of Agricultural and Food Chemistry, 62*, 3390-3399.

Feliciano, R.P., Krueger, C.G., & Reed, J.D. (2015). Methods to determine effects of cranberry proanthocyanidins on extraintestinal infections: Relevance for urinary tract health. *Molecular Nutrition and Food Research, 59*, 1292-1306.

Feliciano, R.P., Meudt, J.J., Shanmuganayagam, D., Krueger, C.G., & Reed, J.D. (2014). Ratio of "A-type" to "B-type" Proanthocyanidin Interflavan Bonds Affects Extra-intestinal Pathogenic Escherichia coli Invasion of Gut Epithelial Cells. *Journal of Agricultural and Food Chemistry, 62*, 3919-3925.

Feliciano, R.P., Shea, M.P., Shanmuganayagam, D., Krueger, C.G., Howell, A.B., & Reed, J.D. (2012). Comparison of Isolated Cranberry (Vaccinium macrocarpon Ait.) Proanthocyanidins to Catechin and Procyanidins A2 and B2 for Use as Standards in the 4-(Dimethylamino)cinnamaldehyde Assay. *Journal of Agricultural and Food Chemistry, 60*, 4578-4585.

Furuuchi, R., Yokoyama, T., Watanabe, Y., & Hirayama, M. (2011). Identification and Quantification of Short Oligomeric Proanthocyanidins and Other Polyphenols in Boysenberry Seeds and Juice. *Journal of Agricultural and Food Chemistry, 59*, 3738-3746.

Gu, L., Kelm, M.A., Hammerstone, J.F., Beecher, G., Holden, J., Haytowitz, D., Gebhardt, S., & Prior, R.L. (2004). Concentrations of proanthocyanidins in common foods and estimations of normal consumption. *Journal of Nutrition, 134*, 613-617.

Hamed, A.I., Al Ayed, A.S., Moldoch, J., Piacente, S., Oleszek, W., & Stochmal, A. (2014). Profiles analysis of proanthocyanidins in the argun nut (Medemia argun-an ancient Egyptian palm) by LC-ESI-MS/MS. *Journal of Mass Spectrometry, 49*, 306-315.

Howell, A.B., Reed, J.D., Krueger, C.G., Winterbottom, R., Cunningham, D.G., & Leahy, M. (2005). A-type cranberry proanthocyanidins and uropathogenic bacterial anti-adhesion activity. *Phytochemistry, 66*, 2281-2291.

Krueger, C.G., (2015) Authentication of Cranberry Ingredients, Natural Produtcs Insider. Available at: http://www.naturalproductsinsider.com/articles/2015/03/authentication-of-cranberry-ingredients.aspx . Accessed 15.05.2016.

Krueger, C.G., Reed, J.D., Feliciano, R.P., & Howell, A.B. (2013). Quantifying and characterizing proanthocyanidins in cranberries in relation to urinary tract health. *Analytical and Bioanalytical Chemistry, 405*, 4385-4395.

Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., de Jong, S., Lewi, P.J., & Smeyers-Verbeke, J. (1997) *Handbook of Chemometrics and Qualimetrics;* Elsevier: Amsterdam.

Navarro, M., Núñez, O., Saurina, J., Hernández-Cassou, S., & Puignou, L. (2014). Characterization of fruit products by capillary zone electrophoresis and liquid chromatography using the compositional profiles of polyphenols: Application to authentication of natural extracts. *J. Agric. Food Cheistry, 62*, 1038-1046.

Nicolosi, D., Tempera, G., Genovese, C., & Furneri, P.M. (2014). Anti-adhesion activity of A2-type proanthocyanidins (a cranberry major component) on uropathogenic E. coli and P. mirabilis strains. *Antibiotics*, *3*, 143-154, 12.

Parets, L., Alechaga, E., Núñez, O., Saurina, J., Hernández-Cassou, S. & Puignou, L., (2016). Ultrahigh pressure liquid chromatography-atmospheric pressure photoionization-tandem mass spectrometry for the determination of polyphenolic profiles in the characterization and classification of cranberry-based pharmaceutical preparations and natural extracts. *Anal. Methods*, *8*, 4363-4378.

Patel, K.D., Scarano, F.J., Kondo, M., Hurta, R.A.R., & Neto, C.C. (2011). Proanthocyanidin-rich extracts from cranberry fruit (Vaccinium macrocarpon Ait.) selectively inhibit the growth of human pathogenic fungi Candida spp. and Cryptococcus neoformans. *Journal of Agricultural and Food Chemistry, 59*, 12864-12873.

Pérez-Rafols, C., & Saurina, J. (2015). Liquid chromatographic fingerprints and profiles of polyphenolic compounds applied to the chemometric characterization and classification of beers. *Analytical Methods*, *7*, 8733-8739.

Prior, R.L., Fan, E., Ji, H., Howell, A., Nio, C., Payne, M.J., & Reed, J. (2010). Multi-laboratory validation of a standard method for quantifying proanthocyanidins in cranberry powders. *Journal of Science and Food Agriculture, 90*, 1473-1478.

Puigventós, L., Navarro, M., Alechaga, E., Núñez, O., Saurina, J., Hernández-Cassou, S., & Puignou, L. (2015). Determination of polyphenolic profiles by liquid chromatography-electrospray-tandem mass spectrometry for the authentication of fruit extracts. *Analytical and Bioanalytical Chemistry, 407*,597-608.

Rzeppa, S., Von Bargen, C., Bittner, K., & Humpf, H. (2011). Analysis of Flavan-3-ols and procyanidins in food samples by reversed phase high-performance liquid chromatography coupled to electrospray ionization tandem mass spectrometry (RP-HPLC-ESI-MS/MS). *Journal of Agricultural and Food Chemistry*, *59*, 10594-10603.

Sanchez-Patan, F., Bartolome, B., Martin-Alvarez, P.J., Anderson, M., Howell, A., & Monagas, M. (2012). Comprehensive assessment of the quality of commercial cranberry products. Phenolic characterization and in vitro bioactivity. *Journal of Agricultural and Food Chemistry, 60*, 3396-3408.

Saurina, J., & Sentellas, S. (2015). Determination of Phenolic Compounds in Food Matrices: Application to Characterization and Authentication. In Núñez, O., Gallart-Ayala, H., Martins, C.P.B., & Lucci, P. (Eds.), *Fast Liquid Chromatography-Mass Spectrometry Methods in Food and Environmental Analysis*, (pp. 517-547). Imperial College Press, London, UK.

Seeram, N.P. (2008). Berry Fruits: Compositional Elements, Biochemical Activities, and the Impact of Their Intake on Human Health, Performance, and Disease. *Journal of Agricultural and Food Chemistry*, *56*, 627-629.

Seeram, N.P. (2012). Emerging research supporting the positive effects of berries on human health and disease prevention. *Journal of Agricultural and Food Chemistry, 60*, 5685-5686.

562 Wallace, T.C., & Giusti, M.M. (2010). Extraction and normal-phase HPLC-fluorescence-
563 electrospray MS characterization and quantification of procyanidins in cranberry extracts. *Journal*
564 *of Food Science, 75*, C690-C696.

565
566
567
568
569

Figure captions

Fig. 1. Score plots of (a) PC1 *vs* PC2 and (b) PC2 *vs* PC3 obtained when using as analytical data for PCA the phenolic peak area information.

Fig. 2. Score plots of PC1 *vs* PC2 obtained with all analyzed samples when using as analytical data for PCA: (a) the full HPLC-UV chromatographic fingerprints; (b) the HPLC-UV chromatographic fingerprints segmented from 3 to 23 min; (c) the HPLC-UV chromatographic fingerprints segmented from 4.7–6.5 min + 8–14 min + 15–17 min + 29–30 min; (d) the HPLC-UV chromatographic fingerprints segmented from 3 to 23 min only with replicates of fruit samples.

Fig. 3. PLS applied to the quantification of the raspberry percentage on cranberry-fruit extracts adulterated when using HPLC-UV chromatographic fingerprints segmented from 3 to 23 min as data. (a) Root mean square error in cross validation (RMSECV) for the estimation of the optimum number of latent variables to be used for the assessment of the calibration model. (b) Plot of scores of latent variable 1 versus latent variable 2. (c) Scatter plot of actual *vs* calculated raspberry percentages in the validation of the calibration model. (d) Scatter plot of actual *vs* calculated raspberry percentages in the validation of predictions.

Fig. 4. PLS applied to the quantification of the raspberry percentage on cranberry-fruit extracts adulterated when using HPLC-UV chromatographic fingerprints segmented from 3 to 23 min as data, and considering only low adulteration levels. (a) Root mean square error in cross validation (RMSECV) for the estimation of the optimum number of latent variables to be used for the assessment of the calibration model. (b) Plot of scores of latent variable 1 versus latent variable 2. (c) Scatter plot of actual *vs* calculated raspberry percentages in the validation of the calibration model. (d) Scatter plot of actual *vs* calculated raspberry percentages in the validation of predictions.

Fig. 5. PLS applied to the quantification of the raspberry percentage on cranberry-fruit extracts adulterated when using the combination of HPLC-UV chromatographic fingerprints at different time segments (5.4–6.3 min + 9.1–13.2 min + 16.2-16.4 min) as data. (a) Root mean square error in cross validation (RMSECV) for the estimation of the optimum number of latent variables to be used for the assessment of the calibration model. (b) Plot of scores of latent variable 1 versus latent variable 2. (c) Scatter plot of actual *vs* calculated raspberry percentages in the validation of the

calibration model. (d) Scatter plot of actual *vs* calculated raspberry percentages in the validation of predictions.

Table 1. Prediction errors by PLS regression in the quantification of cranberry-fruit extracts adulterated with raspberry-, blueberry-, and grape-fruit extracts.

Data for PLS	Calibration error (%)			Prediction error (%)		
	Raspberry	Blueberry	Grape	Raspberry	Blueberry	Grape
HPLC-UV chromatographic fingerprints segment 3-23 min	0.15	0.06	0.17	4.65	2.32	2.53
HPLC-UV chromatographic fingerprints segment 3-23 min (only with low adulteration levels)	0.37	0.96	1.71	2.90	4.26	2.01
HPLC-UV chromatographic fingerprints segments 5.4–6.3 min + 9.1–13.2 min + 16.2-16.4 min	0.10	-	-	2.03	-	-

Figure 1

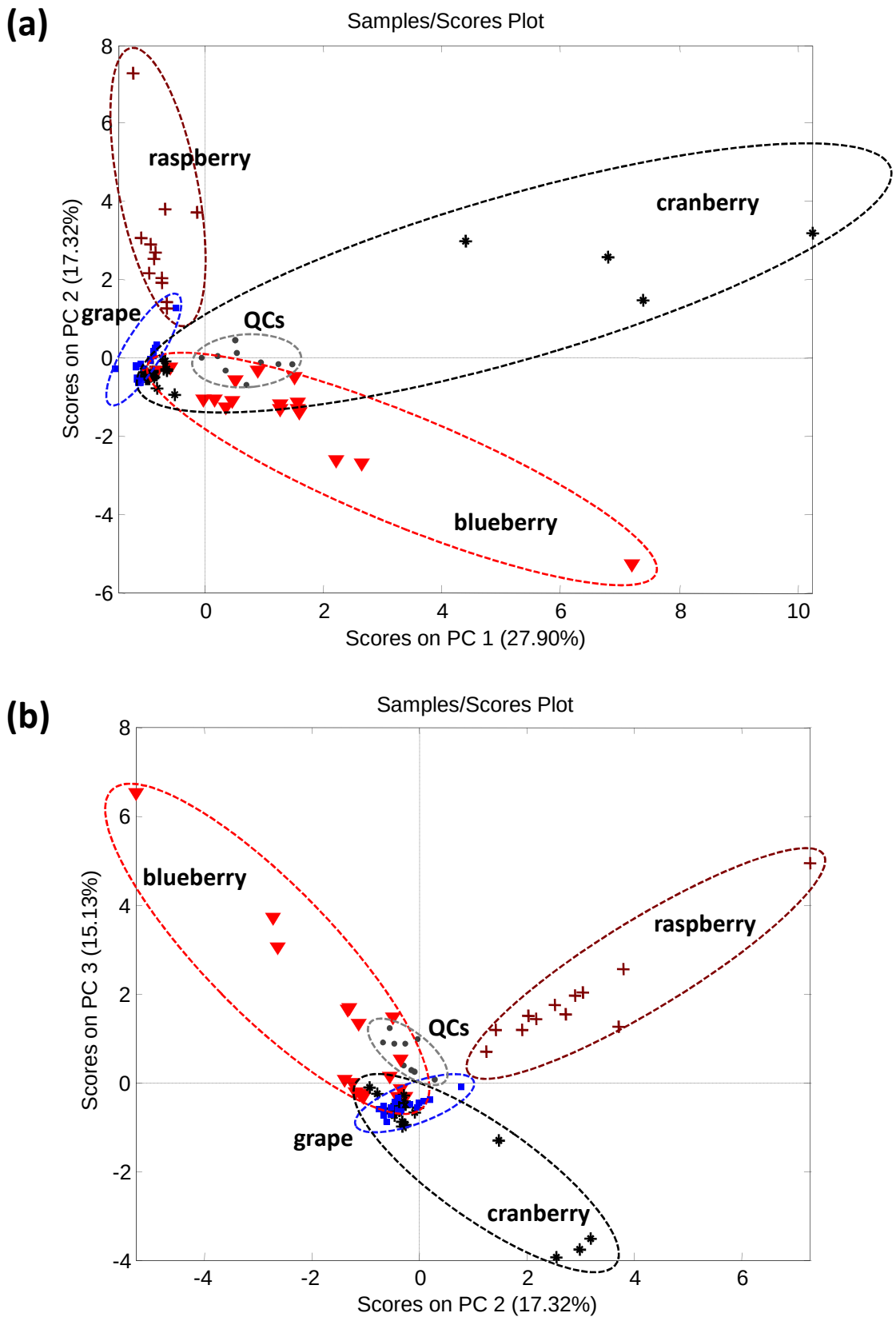

(a)
Samples/Scores Plot
raspberry
cranberry
grape
QCs
blueberry
Scores on PC 2 (17.32%)
Scores on PC 1 (27.90%)
(b)
Samples/Scores Plot
blueberry
raspberry
QCs
grape
cranberry
Scores on PC 3 (15.13%)
Scores on PC 2 (17.32%)

Figure 2

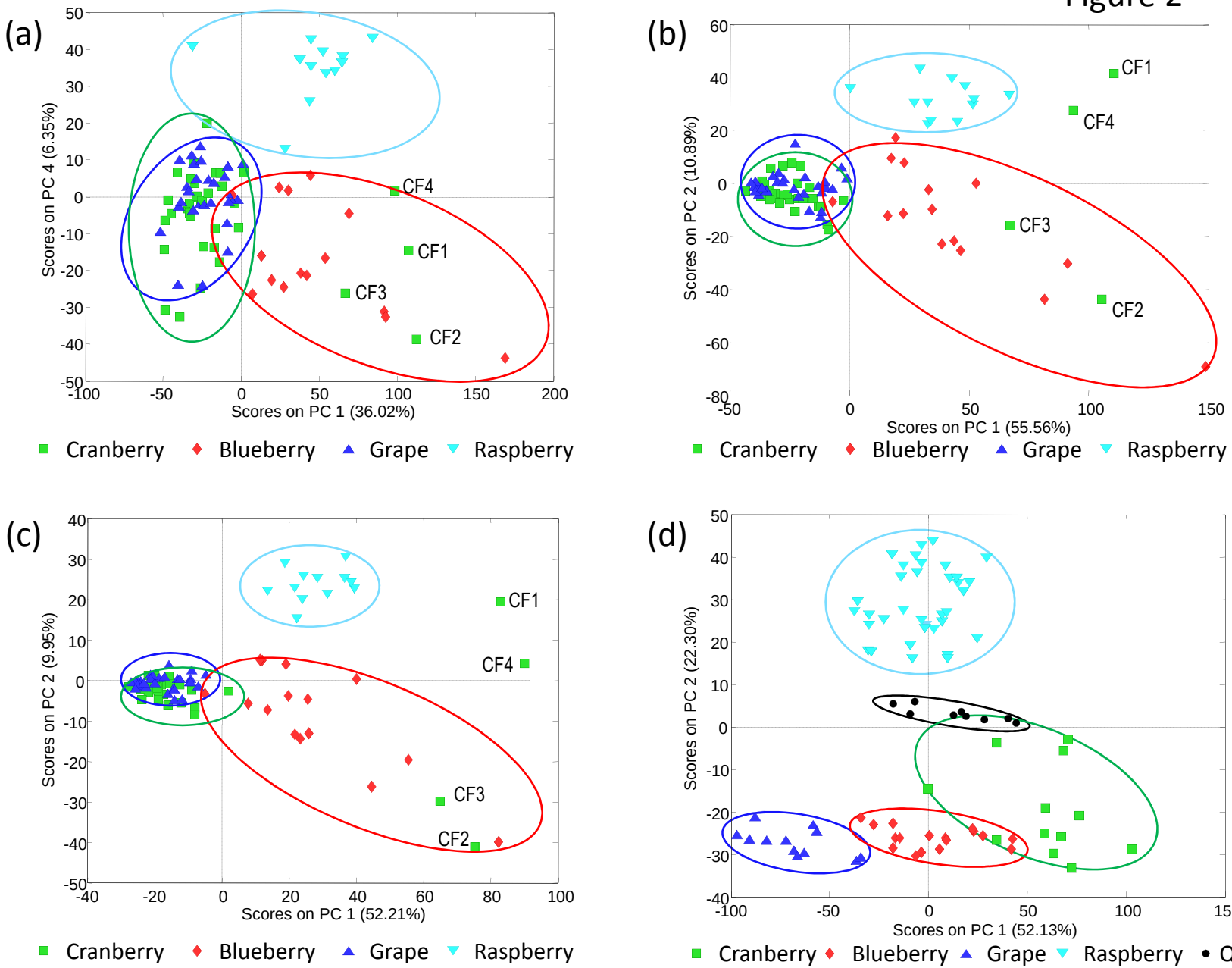

(a)
Scores on PC 4 (6.35%)
Scores on PC 1 (36.02%)
CF4
CF1
CF3
CF2
Cranberry
Blueberry
Grape
Raspberry
(b)
Scores on PC 2 (10.89%)
Scores on PC 1 (55.56%)
CF1
CF4
CF3
CF2
Cranberry
Blueberry
Grape
Raspberry
(c)
Scores on PC 2 (9.95%)
Scores on PC 1 (52.21%)
CF1
CF4
CF3
CF2
Cranberry
Blueberry
Grape
Raspberry
(d)
Scores on PC 2 (22.30%)
Scores on PC 1 (52.13%)
Cranberry
Blueberry
Grape
Raspberry
QC

Figure 3

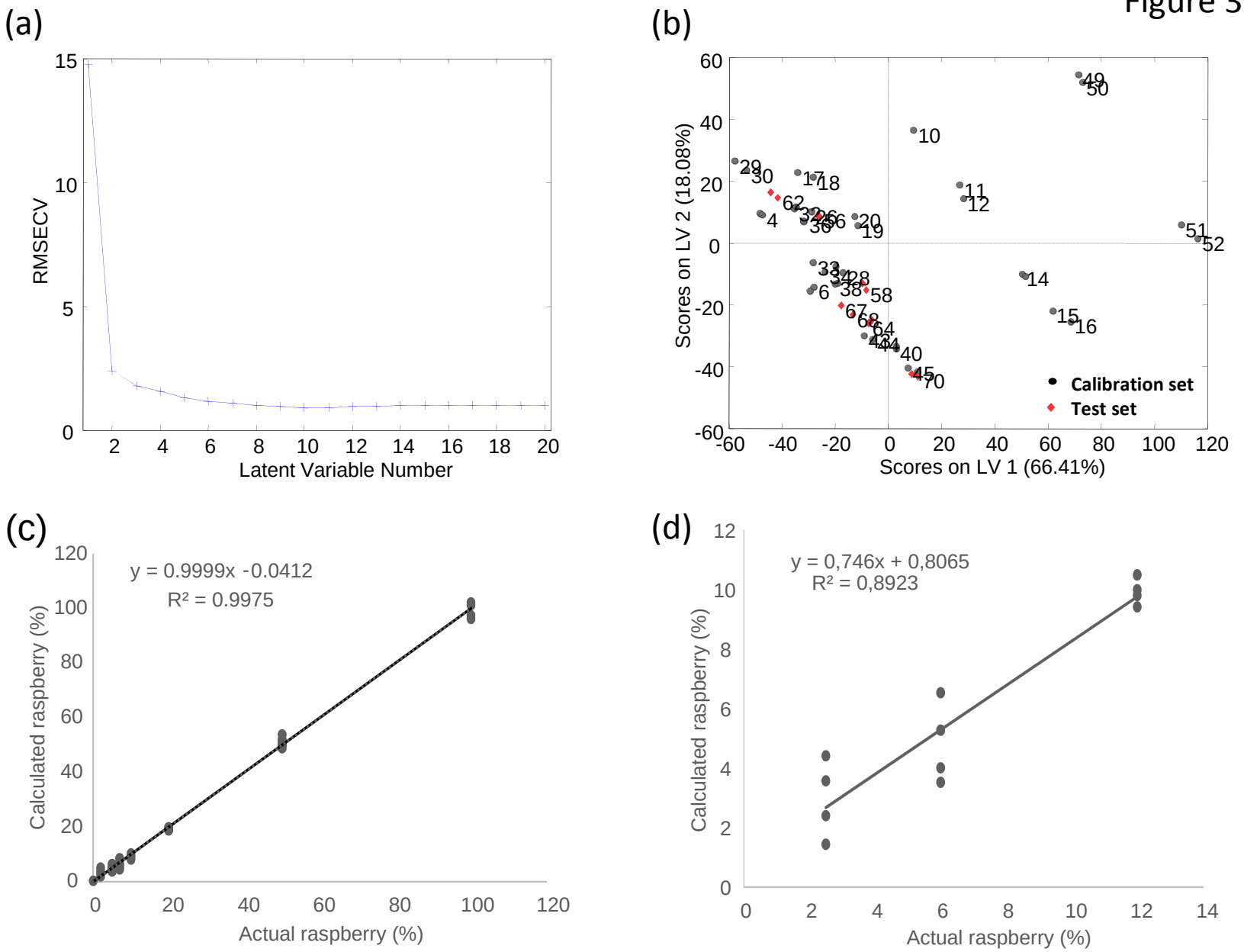
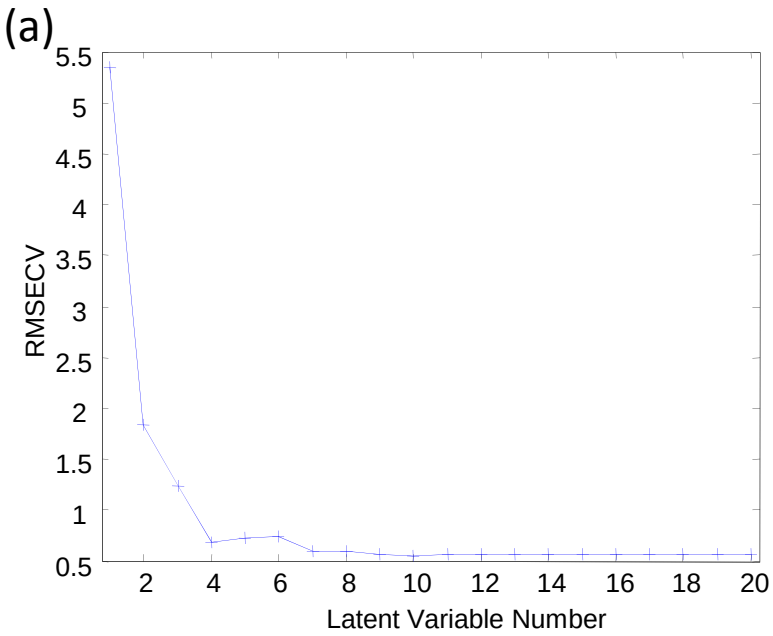
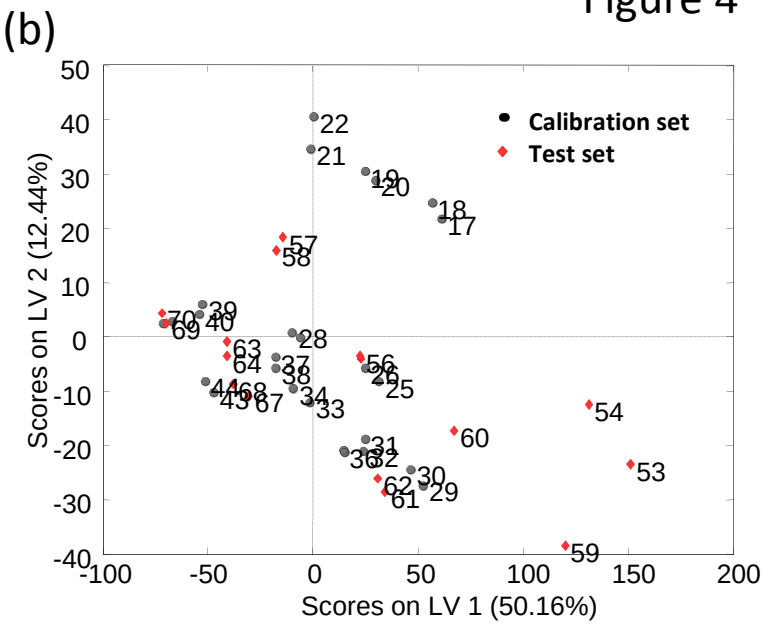
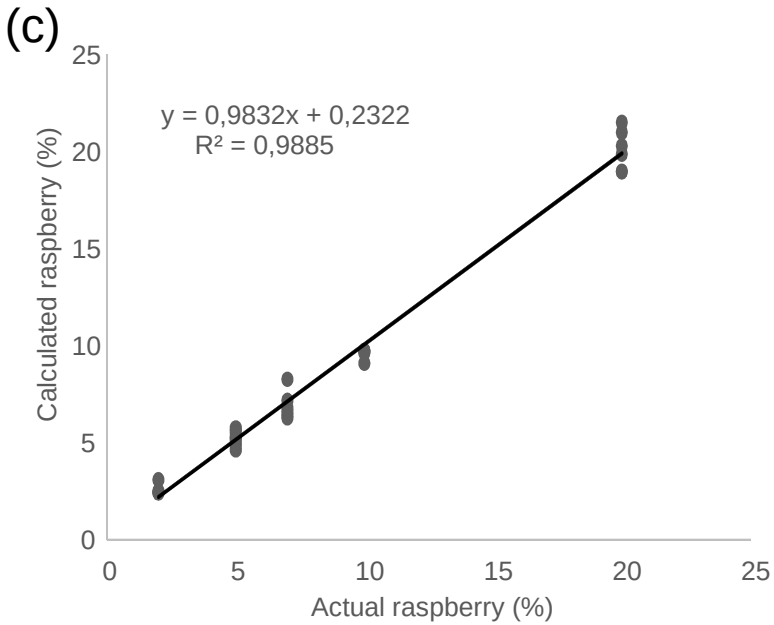

(a)
RMSECV
Latent Variable Number
(b)
Scores on LV 2 (18.08%)
Scores on LV 1 (66.41%)
Calibration set
Test set
(c)
y = 0.9999x -0.0412
R² = 0.9975
Calculated raspberry (%)
Actual raspberry (%)
(d)
y = 0,746x + 0,8065
R² = 0,8923
Calculated raspberry (%)
Actual raspberry (%)

Figure 4


(a)
RMSECV
Latent Variable Number


(b)
Calibration set
Test set
Scores on LV 2 (12.44%)
Scores on LV 1 (50.16%)


(c)
y = 0,9832x + 0,2322
R² = 0,9885
Calculated raspberry (%)
Actual raspberry (%)

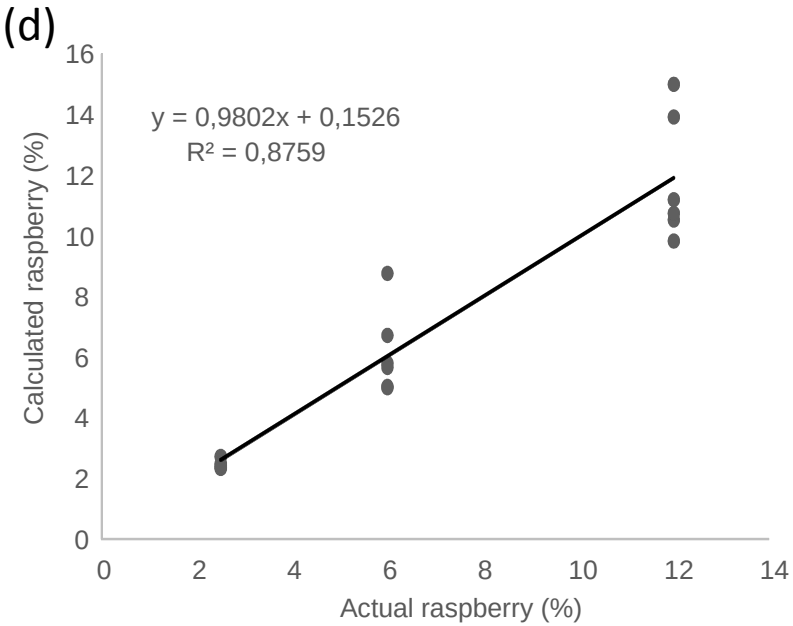

(d)
y = 0,9802x + 0,1526
R² = 0,8759
Calculated raspberry (%)
Actual raspberry (%)

Figure 5

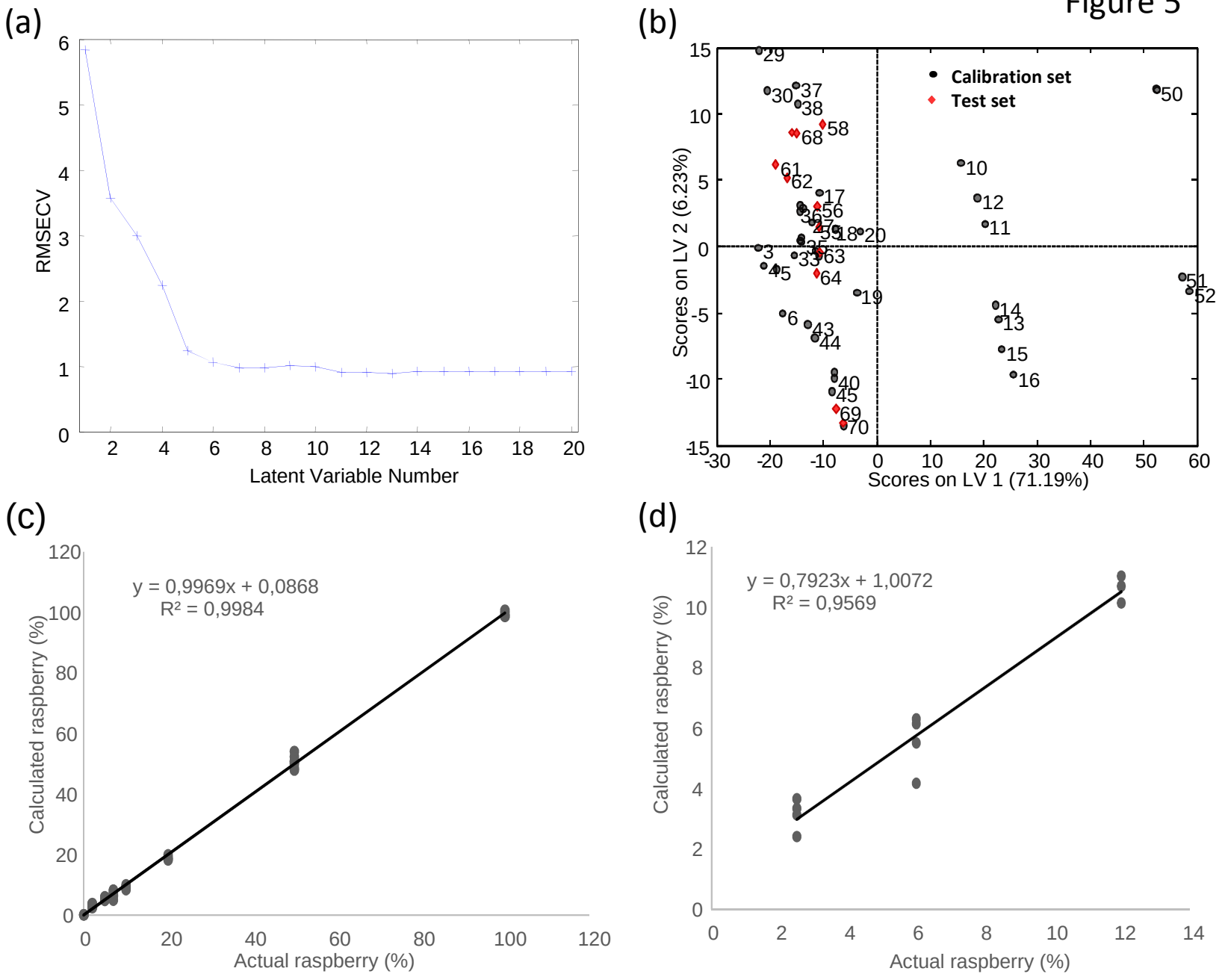

(a)
RMSECV
Latent Variable Number
(b)
Calibration set
Test set
Scores on LV 2 (6.23%)
Scores on LV 1 (71.19%)
(c)
y = 0,9969x + 0,0868
R² = 0,9984
Calculated raspberry (%)
Actual raspberry (%)
(d)
y = 0,7923x + 1,0072
R² = 0,9569
Calculated raspberry (%)
Actual raspberry (%)

Table S1. Chemical structures and classification of the studied polyphenols and phenolic acids.

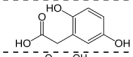
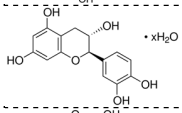
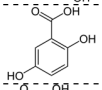
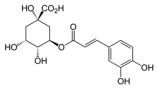
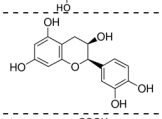
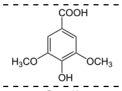
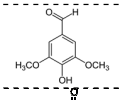
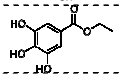
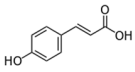
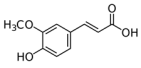
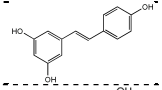
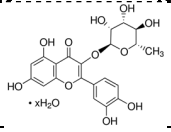
Peak	Phenolic compound	Family	Structure	CAS number
1	Gallic acid	Phenolic acid		149-91-7
2	Homogentístic acid	Phenolic acid		451-13-8
3	Protocatechuic acid	Phenolic acid		99-50-3
4	Protocatechualdehyde	Phenolic aldehyde		139-85-5
5	(+)-Catechin hydrate	Flavanol		225937-10-0
6	Gentisic acid	Phenolic acid		490-79-9
7	*p*-Salicylic acid	Phenolic acid		99-96-7
8	Chlorogenic acid	Phenolic acid		327-97-9
9	Vanillic acid	Phenolic acid		121-34-6
10	(-)-Epicatechin	Flavanol		490-46-0
11	Syringic acid	Phenolic acid		530-57-4
12	Syringaldehyde	Phenolic aldehyde		134-96-3
13	Ethyl gallate	Phenolic acid		831-61-8
14	*p*-Coumaric acid	Phenolic acid		501-98-4
15	Ferulic acid	Phenolic acid		537-98-4
16	Resveratrol	Stilbene		501-36-0
17	Quercitrin hydrate	Flavone		522-12-3

Table S2. Instrumental quality parameters of the proposed HPLC-UV method.

Peak	Compound	LOD (mg/L)	LOQ (mg/L)	Linearity (r^2)	run-to-run precision (%RSD, n=5)				day-to-day precision (%RSD, n=5×3)			
					Migration time		Concentration		Migration time		Concentration	
					Low level[a]	Medium level[b]	Low level[a]	Medium level[b]	Low level[a]	Medium level[b]	Low level[a]	Medium level[b]
1	Gallic acid	0.27	0.88	0.998	0.8	0.7	3.0	1.2	2.2	0.9	6.5	2.7
2	Homogentistic acid	0.89	2.93	0.998	0.3	0.4	3.1	1.4	1.6	0.6	4.4	2.5
3	Protocatechuic acid	0.26	0.87	0.997	0.3	0.1	1.9	1.4	0.5	0.6	3.7	3.9
4	Protocatechualdehyde	0.24	0.80	0.997	0.5	0.5	3.6	1.1	1.3	0.7	4.4	3.5
5	(+)-Catechin hydrate	0.81	2.67	0.997	0.8	0.6	2.4	1.2	1.7	1.0	4.5	3.2
6	Gentistic acid	0.76	2.52	0.996	0.3	0.5	2.3	1.4	0.6	0.8	4.2	3.8
7	*p*-Salicylic acid	0.25	0.83	0.997	0.3	0.5	1.9	1.5	0.6	1.1	6.8	4.2
8	Chlorogenic acid	2.90	9.57	0.995	0.4	0.6	1.3	1.4	0.8	1.3	3.1	4.3
9	Vanillic acid	0.31	1.03	0.986	0.4	0.6	2.5	1.3	0.6	1.0	3.0	6.1
10	(-)-Epicatechin	2.39	7.88	0.997	0.8	0.7	1.1	1.2	2.4	1.3	4.7	3.2
11	Syringic acid	0.87	2.87	0.998	0.4	0.6	3.9	0.4	1.7	1.0	4.2	3.2
12	Syringaldehyde	1.03	3.41	0.999	0.3	0.5	2.3	1.6	0.6	0.8	3.3	4.6
13	Ethyl gallate	0.74	2.46	0.996	0.4	0.3	3.5	0.5	1.7	0.8	4.6	3.3
14	*p*-Coumaric acid	0.16	0.54	0.998	0.1	0.3	2.1	1.2	0.4	0.6	2.8	4.5
15	Ferulic acid	0.25	0.81	0.996	0.1	0.2	3.2	1.0	0.3	0.5	4.5	4.8
16	Resveratrol	0.85	2.81	0.995	0.1	0.4	0.9	0.9	0.6	0.7	2.9	3.9
17	Quercitrin hydrate	0.76	2.52	0.996	0.1	0.3	3.4	1.3	0.4	0.5	4.5	4.3

[a] LOQ

[b] 21.88-35.48 mg/L (depending on the compound)

Table S3. Intra-day and inter-day trueness values at low and medium concentration levels.

Peak	Compound	Trueness									
		Low concentration level					Medium concentration level				
		Concentration value (mg/L)	Intra-day		Inter-day		Concentration value (mg/L)	Intra-day		Inter-day	
			Calculated value (mg/L)	Relative error (%)	Calculated value (mg/L)	Relative error (%)		Calculated value (mg/L)	Relative error (%)	Calculated value (mg/L)	Relative error (%)
1	Gallic acid	0.88	0.89	1.63	0.87	0.08	21.88	22.21	1.52	22.59	3.25
2	Homogentistic acid	2.93	2.94	0.33	3.01	2.87	24.38	24.12	1.07	24.68	1.21
3	Protocatechuic acid	0.87	0.96	10.04	0.97	10.90	27.23	25.01	8.14	28.09	3.15
4	Protocatechualdehyde	0.80	0.80	0.34	0.78	3.07	25.03	25.01	0.05	25.52	1.97
5	(+)-Catechin hydrate	2.67	2.69	0.97	2.83	6.26	27.77	27.76	0.01	28.36	2.13
6	Gentistic acid	2.52	2.41	4.48	2.32	7.82	26.26	26.38	0.45	26.39	0.51
7	*p*-Salicylic acid	0.83	0.84	0.96	0.82	1.18	26.08	25.41	2.60	26.76	2.57
8	Chlorogenic acid	9.57	9.62	0.51	9.65	0.80	29.92	30.64	2.42	30.38	1.56
9	Vanillic acid	1.03	1.01	2.25	0.99	3.88	32.14	32.77	1.97	32.41	0.85
10	(-)-Epicatechin	7.88	8.22	4.29	8.02	1.78	24.63	25.92	5.24	24.94	1.26
11	Syringic acid	2.87	2.97	3.57	3.06	6.59	29.91	30.00	0.29	30.86	3.18
12	Syringaldehyde	3.41	3.33	2.14	3.44	0.87	35.48	34.55	2.63	36.41	2.63
13	Ethyl gallate	2.46	2.44	0.85	2.53	3.03	25.59	25.80	0.82	26.03	1.73
14	*p*-Coumaric acid	0.54	0.54	1.01	0.55	1.03	33.80	32.85	2.83	34.65	2.51
15	Ferulic acid	0.81	0.80	1.13	0.81	0.09	25.39	24.62	3.06	26.08	2.69
16	Resveratrol	2.81	2.65	5.64	2.60	7.33	29.23	29.22	0.04	29.30	0.24
17	Quercitrin hydrate	2.52	2.52	0.28	2.46	2.35	26.22	26.07	0.57	26.26	0.15

Fig. S1. HPLC-UV chromatogram (254 nm) of a standard mixture of polyphenols and phenolic acids at 30 mg/L. Peak identification: (1) gallic acid, (2) homogentistic acid, (3) protocatechuic acid, (4) protocatechualdehyde, (5) (+)-catechin hydrate, (6) gentisic acid, (7) p-salicylic acid, (8) chlorogenic acid, (9) vanillic acid, (10) (-)-epicatechin, (11) syringic acid, (12) syringaldehyde, (13) ethyl gallate, (14) p-coumaric acid, (15) ferulic acid (16) resveratrol and (18) quercitrin hydrate.

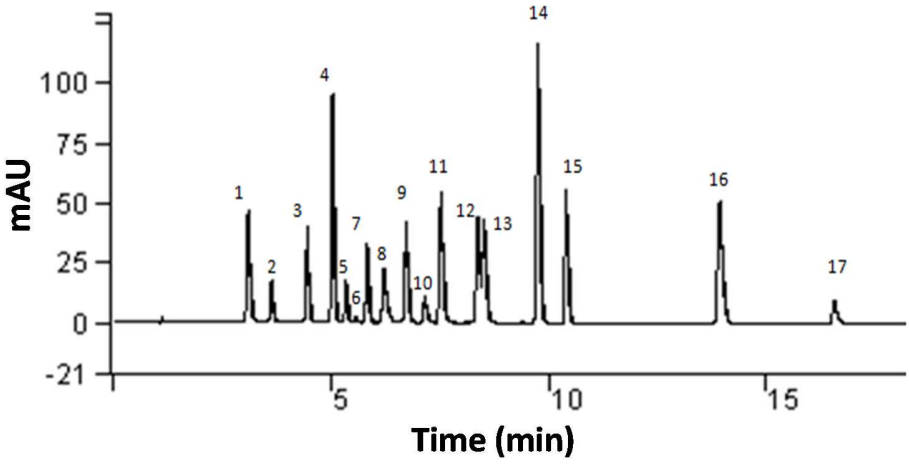


Fig. S2. Loading plots of (a) PC1 *vs* PC2 and (b) PC2 *vs* PC3 obtained when using as analytical data for PCA the phenolics peak area information.

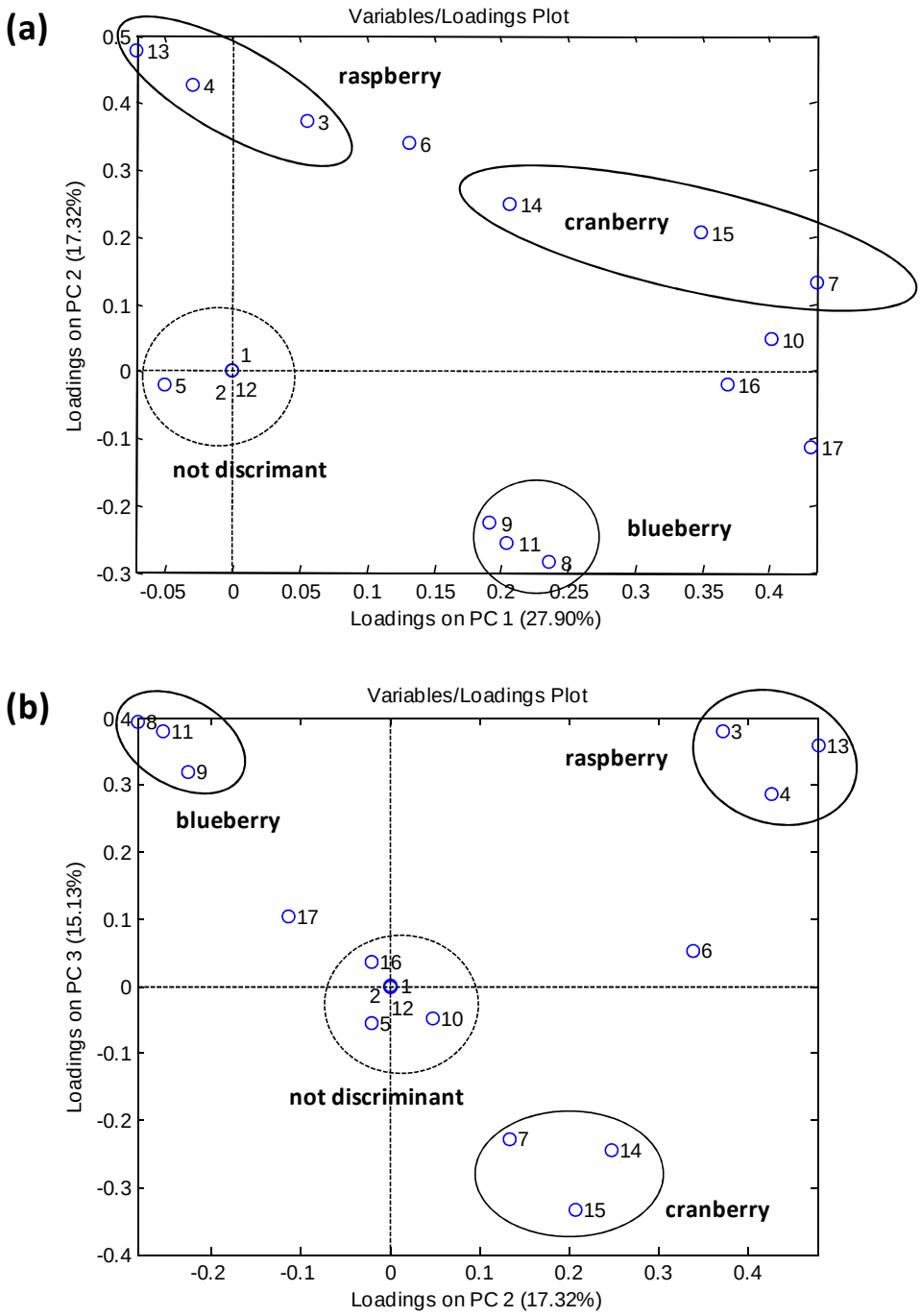


Fig. S3. Segmented HPLC-UV chromatographic fingerprints (from 3 to 23 min) of three cranberry samples (fruit, raisin and juice) acquired at 280 nm.

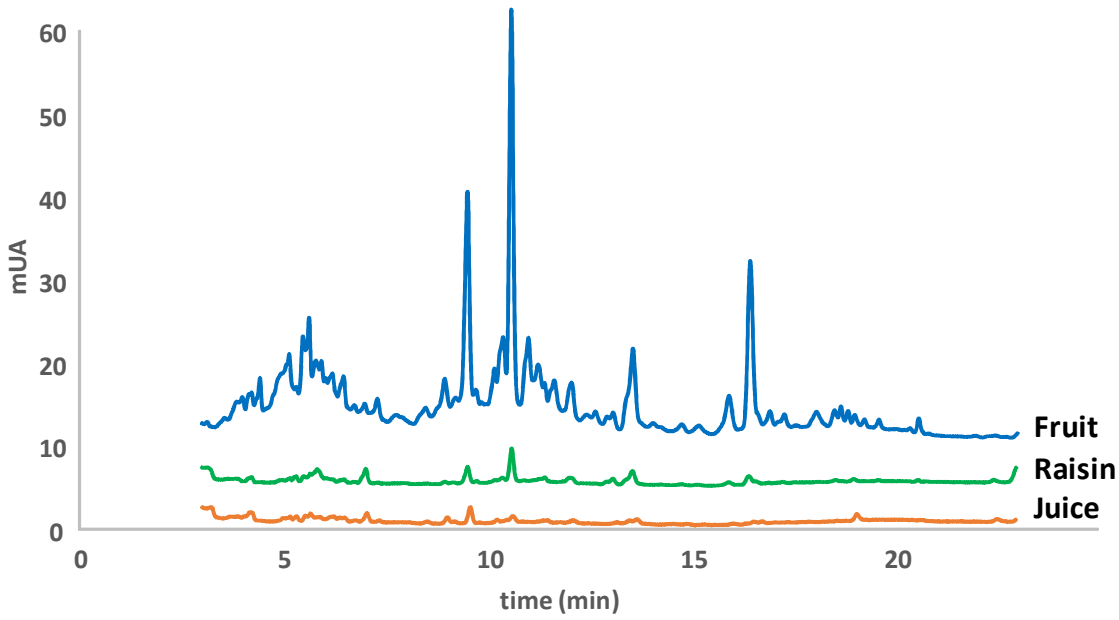


Fig. S4. Partial least squared regression applied to the quantification of the blueberry percentage on cranberry-fruit extracts adulterated when using HPLC-UV chromatographic fingerprints segmented from 3 to 23 min as data. (a) Root mean square error in cross validation (RMSECV) for the estimation of the optimum number of latent variables to be used for the assessment of the calibration model. (b) Plot of scores of latent variable 1 versus latent variable 2. (c) Scatter plot of actual vs calculated blueberry percentages in the validation of the calibration model. (d) Scatter plot of actual vs calculated blueberry percentages in the validation of predictions.

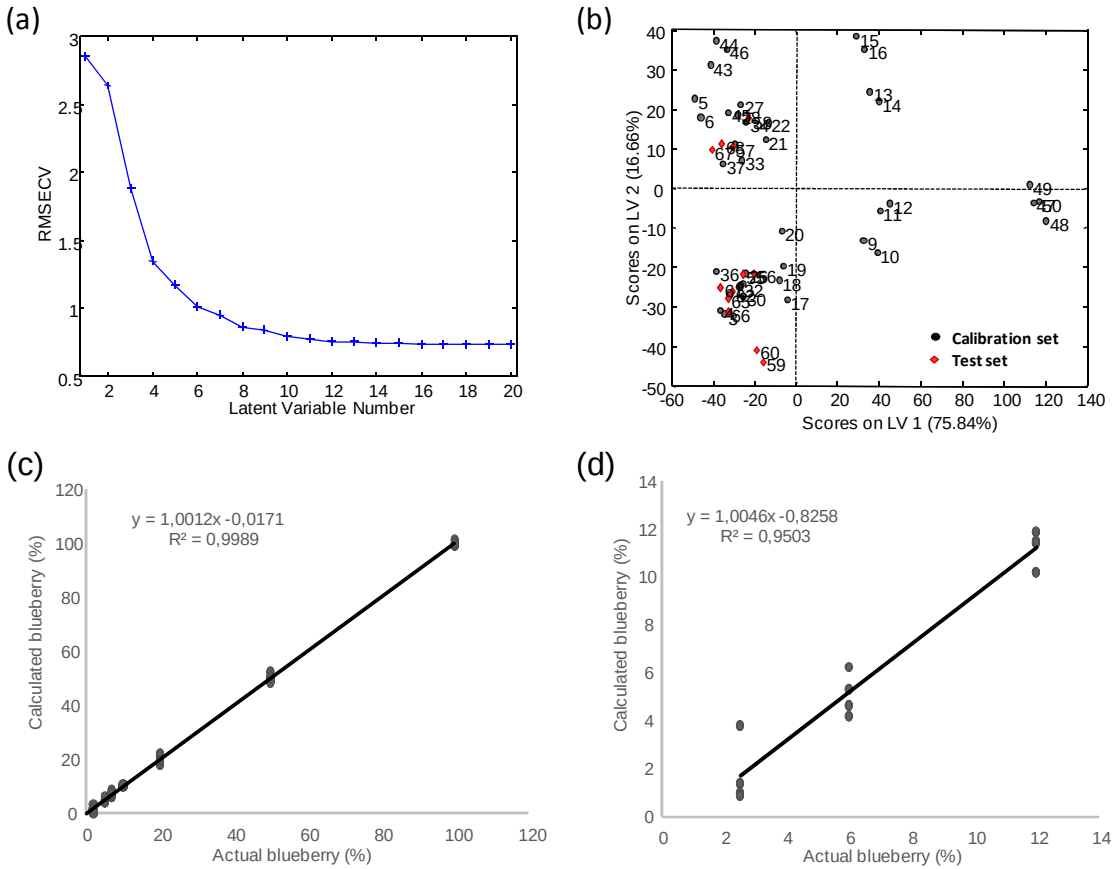


Fig. S5. Partial least squared regression applied to the quantification of the grape percentage on cranberry-fruit extracts adulterated when using HPLC-UV chromatographic fingerprints segmented from 3 to 23 min as data. (a) Root mean square error in cross validation (RMSECV) for the estimation of the optimum number of latent variables to be used for the assessment of the calibration model. (b) Plot of scores of latent variable 1 versus latent variable 2. (c) Scatter plot of actual vs calculated grape percentages in the validation of the calibration model. (d) Scatter plot of actual vs calculated grape percentages in the validation of predictions.

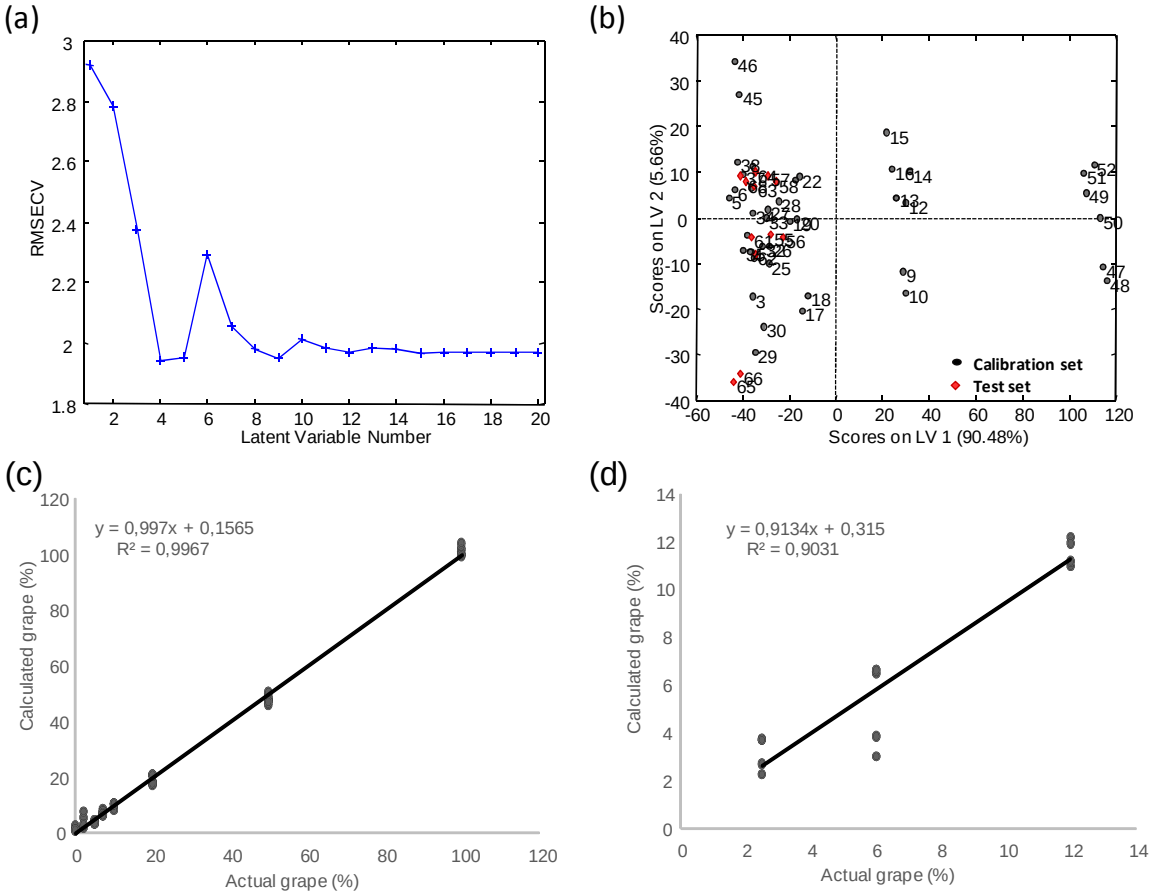


Fig. S6. Partial least squared regression applied to the quantification of the blueberry percentage on cranberry-fruit extracts adulterated when using HPLC-UV chromatographic fingerprints segmented from 3 to 23 min as data, and considering only low adulteration levels. (a) Root mean square error in cross validation (RMSECV) for the estimation of the optimum number of latent variables to be used for the assessment of the calibration model. (b) Plot of scores of latent variable 1 versus latent variable 2. (c) Scatter plot of actual vs calculated blueberry percentages in the validation of the calibration model. (d) Scatter plot of actual vs calculated blueberry percentages in the validation of predictions.

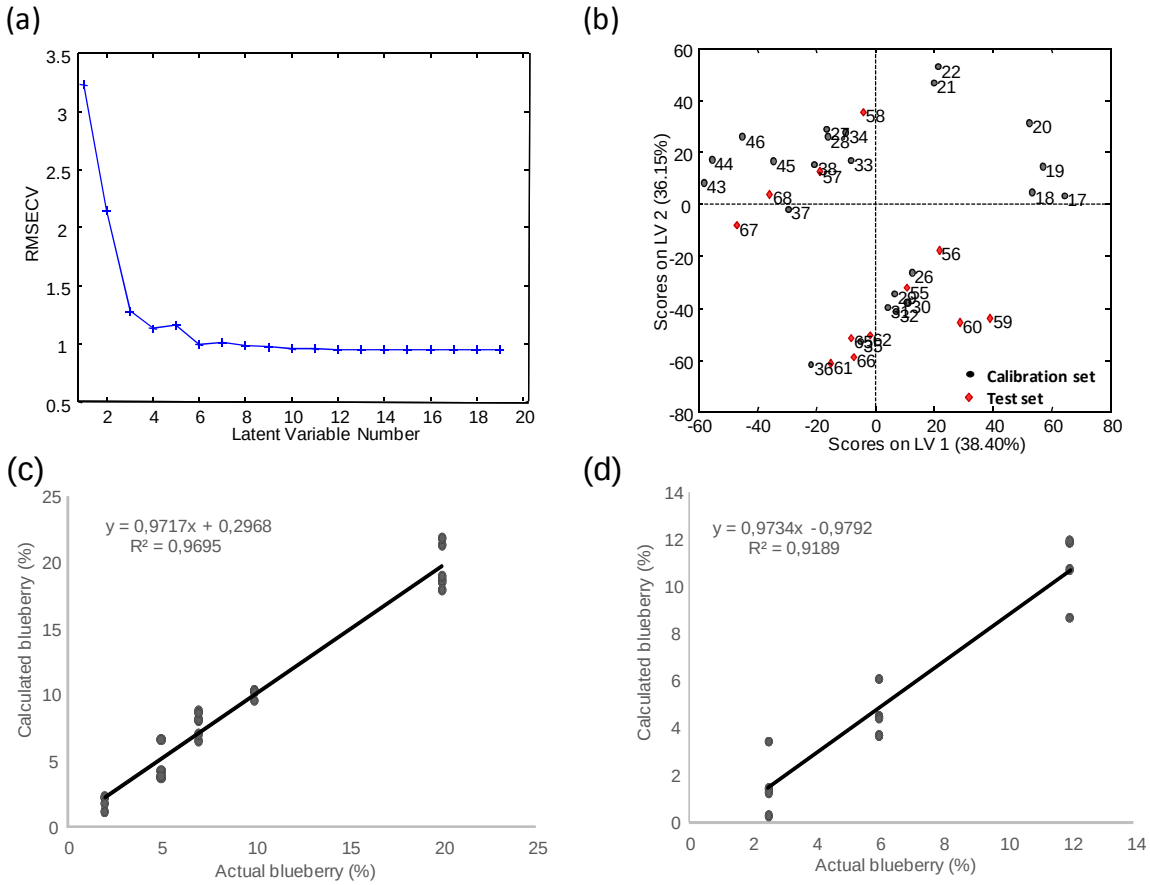


Fig. S7. Partial least squared regression applied to the quantification of the grape percentage on cranberry-fruit extracts adulterated when using HPLC-UV chromatographic fingerprints segmented from 3 to 23 min as data, and considering only low adulteration levels. (a) Root mean square error in cross validation (RMSECV) for the estimation of the optimum number of latent variables to be used for the assessment of the calibration model. (b) Plot of scores of latent variable 1 versus latent variable 2. (c) Scatter plot of actual vs calculated grape percentages in the validation of the calibration model. (d) Scatter plot of actual vs calculated grape percentages in the validation of predictions.

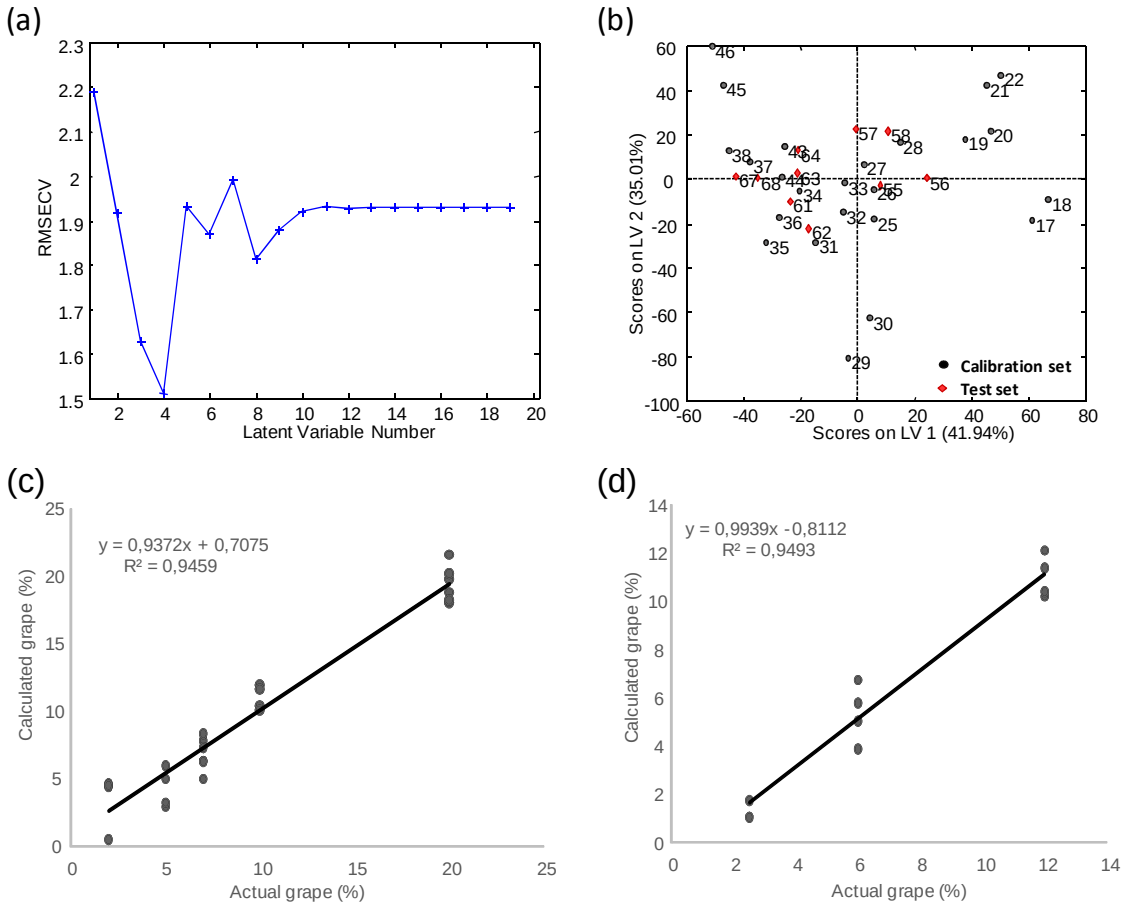


Fig. S8. Segmented HPLC-UV chromatographic fingerprints (from 3 to 23 min) of a raspberry, cranberry, blueberry and grape fruit sample.