A Computational Tool for Accelerated Analysis of Oligomeric Proanthocyanidins in Plants

Mengliang Zhang[†], Jianghao Sun[†], and Pei Chen^{*}

Food Composition and Methods Development Lab, Beltsville Human Nutrition Research Center,
Agricultural Research Service, United States Department of Agriculture, Beltsville, Maryland
20705-2350, USA

[†]Contributed equally to this manuscript.

^{*}Corresponding author: Tel.: +1 301 504 8144; fax: +1 301 504 8314.
E-mail address: pei.chen@ars.usda.gov

Abbreviations: DAD: diode array detection; DP: degrees of polymerization; EA: epiafzelechin/afzelechin; EC: epicatechin/catechin catechin/epicatechin; EG: epigallocatechin/gallocatechin; EF: epifisetinidol/fisetinidol; ER: epirobinetinidol/robinetinidol; ESI: Electrospray ionization; HRAM-MS: high resolution accurate mass-mass spectrometry; MRRF: molar relative response factors; PA: Proanthocyanidin; UHPLC: ultra high-performance liquid chromatography

ABSTRACT

A computational tool was developed to facilitate proanthocyanidin analysis using data collected by ultra-high-performance liquid chromatography-diode array detection-highresolution accuratemass-mass spectrometry (UHPLC-DAD-HRAM-MS). Both identification and semi-quantitation of proanthocyanidins can be achieved by the developed computational tool. It can extract proanthocyanidin chromatographic peaks, deconvolute the isotopic patterns of A-type, B-type, and multi-charged proanthocyanidins ions, and predict proanthocyanidin structures. Proanthocyanidins were quantified by an external calibration curve of catechin and molar relative response factors (MRRFs) of proanthocyanidins. Quantitation results including concentrations of total proanthocyanidins, individual proanthocyanidins, and proanthocyanidins with different degrees of polymerization and different types of linkage were calculated by the program and exported into an Excel spreadsheet automatically. The program was applied to the analysis of seven plant materials including apple, cranberry, dark chocolate, grape seed extract, jujube, litchi, and mangosteen. The identification results were compared with the results obtained by manual processing. The program can greatly save the time needed for the data analysis of proanthocyanidins.

Keywords: Proanthocyanidins; Data processing; Deconvolution; Identification; Quantitation; LC-DAD-MS; Flavonoids; Food analysis; Food composition

1. Introduction

Proanthocyanidins are a group of polyhydroxyflavan-3-ol oligomers and polymers with the monomeric flavan-3-ols as core units, which are linked by carbon–carbon bonds (B-type proanthocyanidins) or additional ether bond O₇–C₂ (A-type proanthocyanidins). Proanthocyanidins have a high structural diversity that is affected by the two chiral centers in each flavan-3-ol unit, existence of substitutions (e.g., galloyl), degrees of polymerization (DP), orders and types of linkage between the flavan-3-ol units, and different monomer units, e.g., epicatechin/catechin catechin/epicatechin (EC), epiafzelechin/afzelechin (EA), epigallocatechin/gallocatechin (EG), epifisetinidol/fisetinidol (EF), and epirobinetinidol/robinetinidol (ER) (Figure S1) (Hümmer and Schreier, 2008; Lin et al., 2014). These structural differences are often associated with the various nutritional outcomes of proanthocyanidins. For example, B-type proanthocyanidins are more common in plants (Chai et al., 2012), but A-type PAs are reported to have stronger bioactivities (De Bruyne et al., 1999; Foo et al., 2000b; Howell, 2002; Kylli et al., 2011). The adsorption of the proanthocyanidins in the gastrointestinal tract is affected by DP; absorption is more difficult as DP increases (Hellstrom et al., 2009; Kallio et al., 2014). Therefore, it is important to quantify proanthocyanidins by different types of linkage and different degrees of polymerization.

Analytical methods for the identification and quantification of PA oligomers in food to identify the role of PAs in health and nutrition is very important (Feliciano et al., 2012; Lin et al., 2014), but PA analysis presents a great challenge. Both normal and reverse-phase high-performance liquid chromatography (HPLC) has been applied to the analysis of PAs, but the selection of a chromatographic method mainly depended on the PAs expected to be present in

the matrix and the research focus (Hümmer and Schreier, 2008). In general, normal-phase HPLC is used for the separation of high DP PAs, since PA's hydrophilicity increases as its DP increases, and reverse-phase HPLC is usually used for the separation of oligomeric PAs and isomers with low DP (Kallio et al., 2014; Kelm et al., 2006). The identification of the types and DP values of PAs is also difficult, even after the HPLC separation.

HPLC coupled with electrospray ionization (ESI) mass spectrometric (MS) detection has become the most important technique for the identification of proanthocyanidins (Hümmer and Schreier, 2008). However, as the DP increases, the increased number of oligomers often causes proanthocyanidins retention time overlap in LC chromatograms, even when ultra-high-performance liquid chromatography (UHPLC) is used. Moreover, the multiply charged ions under ESI of proanthocyanidins with high degree polymerization usually stack on the single charged ions of monomers in the MS spectra. Thus, identification and quantitation of proanthocyanidins manually are very difficult, due to the presence of multiply charged ions, co-eluting peaks, and isomers (Appeldoorn et al., 2009b; Hokkanen et al., 2009; Irungu et al., 2008; Lin et al., 2014; Nunez et al., 2006; Sarnoski et al., 2012).

In this study, a computer program was developed to process data acquired by ultra-high-performance liquid chromatography-diode array detection-electrospray ionization high-resolution mass spectrometry (UHPLC-DAD-ESI/HRMS) for identification and quantification of proanthocyanidins. The problems of the co-eluting peaks and overlap of singly charged ions with multiply charged proanthocyanidin oligomers ions were addressed automatically by the computer program for the first time. Proanthocyanidins were quantified by an external calibration curve of catechin and molar relative response factors (MRRFs) of proanthocyanidins. Seven plant materials were analyzed by the proposed program and method.

2. Materials and methods

2.1. Materials and reagents

Formic acid, HPLC-grade methanol and acetonitrile were purchased from Fisher Scientific (Pittsburgh, PA). HPLC-grade water was prepared from distilled water using a Milli-Q system (Millipore Laboratory, Bedford, MA). Catechin was obtained from Sigma-Aldrich (St. Louis, MO).

Apple, cranberry, dark chocolate, jujube, litchi, and mangosteen were purchased from local grocery stores, and lyophilized immediately upon arrival. Grape seed extract was kindly supplied by Triarco Industries, Inc. (Wayne, NJ). The catechin standard solutions at concentration of 4.12, 8.24, 16.48, 32.96, 41.20 $\mu\text{g/mL}$ were prepared by dilution of 412 $\mu\text{g/mL}$ stock solution with methanol/water (70:30, v/v), and 2 μL of the standard solution were injected into the system.

Each powdered sample (250 mg) was extracted with 5.00 mL of methanol/water (60:40, v/v) and then sonicated for 60 min at room temperature. The slurry mixture was centrifuged at 2500 rpm for 15 min (IEC Clinical Centrifuge; Damon/IEC Division, Needham, MA). The supernatant was filtered through a 17-mm (0.45 μm) PVDF syringe filter (VWR Scientific, Seattle, WA), and 2 μL of the extract were used for each injection.

2.2. Instruments

The UHPLC-PDA HRAM MS system consisted of an LTQ Orbitrap XL mass spectrometer with an Accela 1250 binary pump, a PAL HTC Accela TMO autosampler, a photodiode array (PDA) detector (Thermo Fisher Scientific, San Jose, CA), and a G1316A column compartment (Agilent, Santa Clara, CA) was used. The chromatographic separation was achieved using a UHPLC column (200 mm $\times$ 2.1 mm i.d., 1.9 μm , Hypersil Gold AQ RP-C₁₈; Thermo Fisher

Scientific, Inc., Waltham, MA) with an HPLC/UHPLC pre-column filter (UltraShield Analytical Scientific Instruments, Richmond, CA) at a flow rate of 0.3 mL/min. The mobile phases consisted of **A** (0.1% formic acid in water, v/v) and **B** (0.1% formic acid in acetonitrile, v/v). The linear gradient was from 4 to 20% **B** (v/v) at 40 min, to 35% **B** at 60 min, and to 100% **B** at 61 min and held at 100% **B** to 65 min. The UV-vis spectra were recorded from 190 to 600 nm. For mass spectrometer, negative ionization mode was used, and the conditions were set as follows: sheath gas at 70 (arbitrary units), auxiliary and sweep gas at 15 (arbitrary units), spray voltage at 4.8 kV, capillary temperature at 300 °C, capillary voltage at 15 V, and tube lens at 70 V. The mass range was from m/z 100 to 1500 with a resolution of 30000, FTMS AGC target at 2×10^5 , FT-MS/MS AGC target at 1×10^5 , isolation width of 1.5 amu, and maximum ion injection time of 500 ms. The most intense ion was selected for the data-dependent scan to offer its MS^2 , MS^3 , and MS^4 product ions with a normalization collision energy at 35%.

2.3. Data formats

RAW files were acquired initially. The DAD data were extracted from RAW files by Xcalibur plug-in tool, MSGet (Kazusa DNA Research Institute, 2008) and then were converted to text files. ProteoWizard (Kessner et al., 2008) was used to convert MS data in RAW files to mzXML, which can be then read into MATLAB (MATLAB R2012b, MathWorks Inc., Natick, MA; https://www.mathworks.com/products/new_products/release2012b.html) by the built-in 'mzxmlread' function in MATLAB bioinformatics toolbox.

2.4. Structure information calculation from HRMS spectra

In HRMS spectra, the high-resolution deprotonated molecular ions can be used to calculate the formula of PAs. An algorithm was implemented to calculate the exact molecular formula, DP,

the flavan-3-ols (EC, EA, EF, EG, and ER) in the oligomer, the number of galloyls, and the number of A-type bonds from high resolution deprotonated molecular ions in HRMS data. The algorithm was based on the following equations which have been reported by Lin et al. (2014) (Supporting information equation S1–S4).

3. Results and discussion

3.1. Instrument settings

Chromatographic separation is a crucial step in our study. The purity of chromatographic peak is important because the UV-vis spectrum of each individual peak will be used to compare with that of a reference peak; the co-eluting compounds may distort the UV-vis spectrum if the peaks are overlapping, which may further affect the result of UV-vis spectra comparison. A 65-min UHPLC gradient program was used to separate compounds in the sample extracts. For MS setting, negative ionisation mode was adopted because it was proved to be more sensitive and selective than positive mode for the analysis of proanthocyanidins in LC/ESI-MS methods (Foo et al., 2000a; Gu et al., 2003).

The program took advantage of both UV-vis spectra and high resolution MS spectra for proanthocyanidin analysis. Full UV-vis spectra in range of 190–600 nm are recommended to be recorded. Resolution of the mass spectrometer should be 15,000 or higher. The program can read data in ‘mzXML’ format which is a popular open format and data from major MS instrument vendors can convert to this format (Deutsch, 2012).

3.2. Program workflow

Overall the program executes the data processing using the following steps: UV chromatogram peak detection, target peak extraction, peak identification using MS spectra, and

quantitation. The computer program used for data processing in this work was developed based on FlavonQ (Zhang et al., 2015).

3.3. PA peak detection and extraction

PAs are known to have characteristic UV absorbance at 274–280 nm (Lin et al., 2012). The UV chromatogram at 278 nm was extracted from UV data, and the chromatographic peaks were detected by the program automatically. Figure 1A shows an example of a UV chromatogram at 278 nm from Fuji apple with 64 peaks detected.

In the sample extract, PAs co-exist with other flavonoids and other compounds with absorbance at 278 nm. Thus, differentiation of PAs from other compounds is a crucial step. In our study, the PA peaks are identified using a UV similarity analysis, where the full UV-vis spectrum (190–600 nm) of each chromatographic peak is compared with that of a reference PA peak or a known PA peak, either endogenous, spiked, or external. Each UV-vis spectrum was normalized to unit vector length before the comparison (Zhang and Harrington, 2013, 2015). The comparison is expressed by the Pearson product-moment correlation coefficient (Zhang et al., 2015). The non-PA peaks are eliminated by a user-adjustable threshold for the similarity analysis. Figure 2 shows full UV-Vis spectra of all 64 chromatographic peaks detected in Figure 1A from Fuji apple. By comparing with the full UV-vis spectrum of catechin (PA reference) and setting the similarity analysis threshold at 90%, 11 of the 64 peaks were detected as PA peaks by the program (blue in Figure 2), and they are labeled in Figure 1B. Non-PA peaks are not discussed in detail because they are not the focus of this study. The UV chromatograms at 278 nm for the other six plant materials are shown in supporting information Figure S2–S7.

3.4. PA identification and deconvolution

After the PA peaks are extracted from the UV chromatogram, each peak is identified using its HRMS spectra. The five most intense ions in the full MS scan spectrum of each peak are selected as potential candidate ions for PAs and their structure information is calculated using equations S1–S4 in supporting information. The masses close to the most intense ion (± 3 amu) are also selected from the full MS scan spectrum to examine the isotopic peaks and look for possible overlapping PA peaks.

Compared with a B-type interflavan bond, an A-type interflavan bond loses two hydrogen atoms to form a C–C linkage which causes a molecular weight difference of 2.0156 (2×1.0078 amu). Their isotope patterns can overlap in the MS spectrum. In ESI-MS data, the presence of multiple charged ions complicates the interpretation of the MS spectrum further. Figure 3A, 3B, and 3C represent the mass spectra at m/z 862 to 869 for deprotonated molecules of trimer B-type PA ($[M-H]^-$), trimer A-type PA ($[M-H]^-$), and hexamer B-type PA ($[M-2H]^{2-}$) from grape seed extract. The mass spectra with overlapping isotope patterns are shown in Figure 3D.

Feliciano et al. (2012) proposed three methods to deconvolute a mass spectrum with only single charged deprotonated molecular ions ($[M-H]^-$) overlapping isotope patterns from MALDI-TOF-MS data. All are based on the natural abundance of C isotopes within PA oligomers. In this study, the deconvolution method is developed for ESI-MS data, and accounts for the existence of double and triple charged deprotonated molecules ($[M-2H]^{2-}$ and $[M-3H]^{3-}$) in the overlapping isotope patterns. $[M-2H]^{2-}$ and $[M-3H]^{3-}$ ions for PA oligomers typically only occur for PA oligomers with higher DP (at least pentamers and molecular weights larger than 1350 Da) in ESI-MS spectra (Lin et al., 2014). Discovery of $[M-2H]^{2-}$ and $[M-3H]^{3-}$ ions is achieved by examining the distance between ^{12}C and ^{13}C isotope ions of PA, for example, 0.50

amu distance indicates the presence of $[M-2H]^{2-}$ ions, and 0.33 amu distance indicates the presence of $[M-3H]^{3-}$ ions.

Deconvoluting the overlapping isotope patterns which differentiates the ^{12}C and ^{13}C isotope ions of PA in the MS peak cluster can be quite challenging. The following strategy is used in our program:

- 1) Examine the ion with the smallest m/z in the cluster (e.g., m/z 863.1824 in Figure 3D), check charge states (single, double, or triple charges), and calculate the formula using equation S1, S2 or S3 in supporting information;
- 2) If no PA formula can be obtained from step 1, skip it and examine the next ion on the right and repeat step 1;
- 3) If the ion passes step 1, calculate the relative intensity of its ^{13}C isotope ions by the equation below:

$$I_{^{13}C} = \frac{N!}{k!(N-k)!} \times \left(\frac{1.0816}{100}\right)^k \quad (1)$$

where $I_{^{13}C}$ represents the relative intensity of the ^{13}C isotope ion compared with the ^{12}C ion, N represents the number of C atom in PA formula, k represents the number of ^{13}C atoms, and $\frac{1.0816}{100}$ is the natural abundance ratio between the ^{13}C and ^{12}C isotopes of carbon. Subtract intensity of ^{13}C isotope ions from MS spectrum (e.g., subtract yellow portions in m/z 864.1902, 865.1980, and 866.2058), examine the next ion on the right and repeat steps 1 to 4 until you reach the last ion of the MS spectrum.

The strategy above can also be used to estimate the relative intensity ratio of each PA in the MS peak cluster. However, it is easier and more efficient to put the data in a matrix form and use

the ‘pinv’ function in Matlab to calculate the ratio of PAs in the MS peak clusters. Using Figure 3E as an example, the experimental MS spectrum (blue) can be deconvoluted by the equation below:

$$\begin{bmatrix} R_{red} \\ R_{yellow} \\ R_{green} \end{bmatrix} = \begin{bmatrix} R_{red} \frac{m}{z} 863.18 & R_{yellow} \frac{m}{z} 863.18 & R_{green} \frac{m}{z} 863.18 \\ R_{red} \frac{m}{z} 864.19 & R_{yellow} \frac{m}{z} 864.19 & R_{green} \frac{m}{z} 864.19 \\ R_{red} \frac{m}{z} 864.69 & R_{yellow} \frac{m}{z} 864.69 & R_{green} \frac{m}{z} 864.69 \\ R_{red} \frac{m}{z} 865.20 & R_{yellow} \frac{m}{z} 865.20 & R_{green} \frac{m}{z} 865.20 \\ R_{red} \frac{m}{z} 865.70 & R_{yellow} \frac{m}{z} 865.70 & R_{green} \frac{m}{z} 865.70 \\ R_{red} \frac{m}{z} 866.21 & R_{yellow} \frac{m}{z} 866.21 & R_{green} \frac{m}{z} 866.21 \\ R_{red} \frac{m}{z} 867.21 & R_{yellow} \frac{m}{z} 867.21 & R_{green} \frac{m}{z} 867.21 \\ R_{red} \frac{m}{z} 868.22 & R_{yellow} \frac{m}{z} 868.22 & R_{green} \frac{m}{z} 868.22 \end{bmatrix}^{-1} \times \begin{bmatrix} I_{\frac{m}{z} 863.18} \\ I_{\frac{m}{z} 864.19} \\ I_{\frac{m}{z} 864.69} \\ I_{\frac{m}{z} 865.20} \\ I_{\frac{m}{z} 865.70} \\ I_{\frac{m}{z} 866.21} \\ I_{\frac{m}{z} 867.21} \\ I_{\frac{m}{z} 868.22} \end{bmatrix} \quad (2)$$

$\vdots$
X

$\vdots$
A

$\vdots$
Y

where matrix X represents the intensity ratios of trimer B-type PA (red), trimer A-type PA (yellow) and hexamer B-type PA (green); matrix A represents the theoretical isotope ion distributions of trimer B-type PA, trimer A-type PA and hexamer B-type PA (corresponding to each column, respectively), which can be obtained from equation 1; and matrix Y is the observed intensities from m/z 863 to 869 in the MS spectrum. After normalizing the mass spectrum to the unit vector (Faver et al., 2011, 2012; Wang et al., 2016), the estimated ratios of B-type PA trimer (deprotonated molecular ions m/z 865.1980), A-type PA trimer (deprotonated molecular ions m/z 863.1823), and double charged B-type PA hexamer (deprotonated molecular ions m/z 864.1902 with double charges) can be displayed on the figure and also in the spreadsheet. The deconvolution result can be displayed as a reflected image (Figure 3E), and the similarity

between the original MS spectrum (Figure 3E top) and the deconvoluted MS spectrum (Figure 3E bottom) is evaluated by using a spectral-contrast-angle method, and expressed as a percentage (Equation 3) (Wan et al., 2002). An example of acquired MS spectrum is shown in Figure S8.

$$\cos \theta = \frac{\sum_i Ie_i Id_i}{\sqrt{\sum_i Ie_i^2 \sum_i Id_i^2}} \times 100\% \quad (3)$$

where Ie is the relative intensity of experimental MS peaks, and Id is the relative intensity of deconvoluted MS peaks. The deconvoluted MS spectra with low value of $\cos \theta$ ($<90\%$) were questionable and labeled for further manual examination and interpretation.

The putative identification result of PAs in seven samples using the proposed approach is shown in Table 1. The results from manual identification of seven samples were used as references to compare. Manual identification was conducted by the study of retention time, UV-vis spectrum, $[M-H]^-$, and MS^{2-5} product ions. For certain ions of interest, selected ion monitoring (SIM) mode was used for further investigation (Hammerstone et al., 1999). The details about manual identification of proanthocyanidins in seven samples can be found in our lab's previous publication (Lin et al., 2014). Forty-three PAs with different molecular masses (isomers are not differentiated) were discovered in the seven tested samples using the computer program, while only 23 of them were identified manually. For the 23 PAs identified by both the computer program and manual method, 13 of them were found to be present in more samples identified by the computer program; for example, procyanidin B2 expressed as EC-EC in Table 1 was found in all the samples by the computer program but was only found in apple, dark chocolate, grape seed extract and mangosteen samples by manual identification. The PAs identified exclusively by the computer program were present as coexisting ions in small quantities which do not have MS^n spectra available; thus they were not confirmed as

proanthocyanidins with the manual method and not reported in Lin's study (Lin et al., 2014). It is easy to filter these peaks out. However, it is also beneficial to leave them in for possible future studies. The user can quickly eliminate them as there are no MSⁿ spectra for these ions.

In rare cases, manual identification discovered PAs in the sample when the computer program did not. For example, PA tetramer EC-A-EC-A-EC-EC was found in cranberry extract and litchi by manual identification, but was only found in litchi by the computer program. Two possible reasons are: 1) the PA peak was too small to be detected by the peak detection step, and 2) the PA co-eluted with other compounds which caused distortion of the UV-Vis spectrum and was excluded in the peak extraction step. In either case, the undiscovered PAs were not major PAs in the samples, based on their quantity (next section).

It is worth noting that in Lin's study, 247 PAs in 90 isomeric groups were predicted, but the computer program cannot differentiate the isomers of PAs in the current stage. For example, epicatechin-(4 β →8)-catechin (procyanidin B1, Figure S9-1), (-)-epicatechin-(4 β →8)-(-)-epicatechin (procyanidin B2, Figure S9-2), catechin-(4 α →8)-catechin (procyanidin B3, Figure S9-3), and catechin-(4 α →8)-epicatechin (procyanidin B4, Figure S9-4) are all identified as EC-EC by the computer program. Expertise in the field of proanthocyanidin research is needed to identify the isomers by interpreting the MS fragmentation patterns in MSⁿ spectra and comparing them with reference standards.

3.5. PA identification accuracy

As discussed above, PA peak extraction by UV-vis spectrum similarity analysis can filter most non-PA peaks, and mass spectrum analysis can further exclude non-PA peaks by examining the major ions in full scan. However, incorrect identification of PAs may occur, especially in

case of overlapping between non-PA and PA chromatographic peaks. For example, Peak ID 24 in Figure 1 is a non-PA peak in Fuji apple sample, which is overlapped with Peak ID 25 (a flavan-3-ol). UV-vis spectrum of Peak ID 24 was distorted by flavan-3-ol in Peak ID 25 and it was misclassified as a PA candidate peak after the UV-vis spectrum similarity analysis. In the MS spectrum of Peak ID 24, m/z 289.0712 ($[M-H]^-$ for flavan-3-ol) was also observed as one of the major ions (See Figure S10). Therefore Peak ID 24 was misidentified as a PA peak (false positive peak). Similar examples can be found in different samples in Figure S2–S7 (false positive peaks were labeled in red color). As emphasized before, the chromatographic separation is critical to the correct identification of PA peaks, and a better separation could reduce the false positive identifications. Table 2 summarized chromatographic peak analysis results by the program. PAs dominated in grape seed extract sample and no false positive peaks were found. For other samples, various polyphenols were found and a couple of false positive peaks were observed, but they are all small peaks and are susceptible to interfering. These peaks could be further excluded by manual verification.

3.6. Quantitation of PAs

The MRRF values represent the ratio of the molar UV absorptivity of the compounds to that of catechin (as a reference compound). An external standard calibration curve of catechin was constructed in this study using weighted least squares linear regression with a weighting factor of $1/x^2$ (x represents concentration) (Zhang and Harrington, 2014; Zhang and Harrington, 2015). It was found that the UV absorbance response of the monomers in molar units was additive for PAs and was not affected by the types of linkages between the flavan-3-ol units (A- or B- type). The MRRF values for EC, EG and gallic acid were determined to be 1.00, 0.31 and 2.8 (Lin and Harnly, 2012). The MRRF values for EA and EF were assigned to be 1.00 (Lin et al., 2014).

Since the structure information of PAs was obtained from the PA identification step, the respective MRRF values were easily calculated automatically.

Some UV chromatographic peaks may contain multiple co-eluted PAs. The deconvolution results of MS spectra in the PA identification step were used to calculate the concentration of individual PAs in a PA-mixture peak (Equation 4). The relative ratio of PAs in a PA-mixture peak was based on the MS ion counts which assumes that different PAs responded the same in the MS instrument.

$$C_{PA_1} = \frac{R_{PA_1} \times C_{cat} \times MW_{PA}}{MW_{cat} \times MRRF_{PA_1}} \quad (4)$$

where R_{PA_1} represents the relative ratio of a specific PA in the chromatographic peak which can be obtained from Equation 2; C_{cat} represents the concentration of the chromatographic peak equivalent to catechin; C_{PA_1} represents the calculated concentration for a specific PA; MW_{cat} ; and MW_{PA} are the molecular weights of catechin and PA, respectively.

After we examined all the chromatographic peaks in which multiple PAs were co-eluted, we found that most peaks contained PAs with the same units but different DPs or different types of linkages such as the examples in Figure 3E (EC-EC-EC vs EC-EC-A-EC vs EC-EC-EC-EC-EC). A novel quantitative method was proposed for the first time to calculate masses of co-eluting PAs in a single chromatographic peak (equation 5). From equation 5, we found that the amount of PAs by mass in a chromatographic peak can be simply represented by the monomer PA in the peak. The amount of PAs calculated by this method is considered to be more accurate, because the assumption of constant MS response for all PAs in equation 4 is not always true and is not a necessity in the calculation of equation 5. The PA concentrations in percent dry weight can be simply calculated using equation 6. It is worth noting that this method can measure the

total PA mass of each chromatographic peak accurately, but to quantify co-eluted PAs respectively in a peak, their responses in MS spectra still need to be used.

Assume a chromatographic peak contains monomer and oligomers, and R represents the relative ratios of them, so $R_{oligomer} + R_{monomer} = 100\%$

$$\text{Given } DP \approx \frac{MW_{oligomer}}{MW_{monomer}} \text{ and } DP \approx \frac{MRRF_{oligomer}}{MRRF_{monomer}},$$

$$\text{so } \frac{MW_{monomer}}{MRRF_{monomer}} \approx \frac{MW_{oligomer}}{MRRF_{oligomer}}$$

$$\begin{aligned} M &= \frac{R_{oligomer} \times C_{cat} \times MW_{oligomer} \times V_s}{MW_{cat} \times MRRF_{oligomer}} + \frac{R_{monomer} \times C_{cat} \times MW_{monomer} \times V_s}{MW_{cat} \times MRRF_{monomer}} \\ &\approx \frac{R_{oligomer} \times C_{cat} \times MW_{monomer} \times V_s}{MW_{cat} \times MRRF_{monomer}} + \frac{R_{monomer} \times C_{cat} \times MW_{monomer} \times V_s}{MW_{cat} \times MRRF_{monomer}} \\ &= \frac{(R_{oligomer} + R_{monomer}) \times C_{cat} \times MW_{monomer} \times V_s}{MW_{cat} \times MRRF_{monomer}} = \frac{C_{cat} \times MW_{monomer} \times V_s}{MW_{cat} \times MRRF_{monomer}} \quad (5) \end{aligned}$$

$$C(\%, w/w) = \frac{M}{W_s} \times 100\% \quad (6)$$

where MW and R are the molecular weight and relative ratio of PAs; C_{cat} and MW_{cat} are the concentration of the chromatographic peak equivalent to catechin and the molecular weight of catechin; V_s and W_s are the volume of the extract and the sample weight; and M is the mass of PAs in a chromatographic peak.

The concentrations of PAs as percent dry weight were exported and summarized in an Excel spreadsheet which includes the individual PA concentrations in each chromatographic peak, concentrations of the monomers through the hexamers, concentrations of A-type PAs and B-type

PAs, and concentration of total PAs. An example of the exported results for Fuji apple sample is shown in Table 2. The concentrations of PAs in all seven samples are shown in Table 3. Over 87% of the proanthocyanidins in cranberry were A-type proanthocyanidins which are the main bioactive components associated with the prevention of urinary tract infections (Krueger et al., 2013). In our study, monomers, A-type dimers and trimers were found to be the terminal units of proanthocyanidins in cranberries which agreed with a previous study, but a lower proportion of A-type dimers and trimers (46.1%) were reported in that study (Gu et al., 2002). About 60% proanthocyanidins in litchi fruit pericarp were A-type proanthocyanidins, similar to those of other authors (Le Roux et al., 1998). Our study also found that proanthocyanidins in mangosteen fruit pericarp were dominated by B-type oligomers (49.8%) and monomers (43.8%) with a small amount of A-type oligomers (6.3%), the same as indicated in the literature (Fu et al., 2007). The chocolate, Fuji apple, grape seed extract, and jujube fruit samples barely contained any A-type proanthocyanidins (<3%) (Appeldoorn et al., 2009a; Lazarus et al., 1999; Lin et al., 2014; Miller et al., 2006; Ohnishi-Kameyama et al., 1997; Wojdyło et al., 2016).

3.7. Computation speed

The computation speed of the program for the PA data analysis was about 1 min per sample which was affected by the number of chromatographic peaks, number of PA mass spectra peaks, and the performance of the computer. The proposed computational tool is designed to facilitate the data analysis of PAs, not to replace human experts. Although the program could not provide positive confirmation of the structures for PAs isomers, it significantly reduced the time and expertise needed for data analysis by automatically eliminating non-PA peaks, deconvoluting overlapping MS peaks of PAs, quantifying PAs with various output results, and facilitating the structure elucidation of PAs, since all the information about the substitutions, degrees of

polymerization, types of linkage, and the monomer units are provided by the program. Compared with manual data processing, it also avoided the false negatives arising from human error, i.e., small PA peaks which were ignored by manual data processing.

4. Conclusion

In this study a computer program was developed as a UHPLC-HRMS data processing tool for putative identification and quantitation of proanthocyanidins. The isotopic patterns of A-type, B-type, and multi-charged proanthocyanidins ions were deconvoluted and their structures predicted. Quantitative information for proanthocyanidins, such as concentrations for individual proanthocyanidins, proanthocyanidins with different degrees of polymerization and different types of linkage, and total proanthocyanidins was exported to an Excel spreadsheet. The program facilitated the proanthocyanidins data analysis by extracting proanthocyanidin peaks, providing putative proanthocyanidin identification, and producing quantitative results. It was determined that this program saves enormous amounts of time compared to manual data-mining. The program and strategy were successfully applied to proanthocyanidin analysis in seven plant materials.

Acknowledgement.

This research is supported by the Agricultural Research Service of the U.S. Department of Agriculture, an Interagency Agreement with the Office of Dietary Supplements at the National Institutes of Health. We thank Dr. Long-Ze Lin of the Food Composition and Methods Development Lab, Beltsville Human Nutrition Research Center, Agricultural Research Service, United States Department of Agriculture for his valuable suggestions to this study.

Funding

This research is supported by the Agricultural Research Service of the U.S. Department of Agriculture, an Interagency Agreement with the Office of Dietary Supplements at the National Institutes of Health (Y01 OD001298-01). The authors have declared no conflict of interest.

References

- Appeldoorn, M.M., Sanders, M., Vincken, J.P., Cheynier, V., Le Guerneve, C., Hollman, P.C.H., Gruppen, H., (2009a). Efficient isolation of major procyanidin A-type dimers from peanut skins and B-type dimers from grape seeds. *Food Chemistry*, 117(4), 713-720.
- Appeldoorn, M.M., Vincken, J.P., Sanders, M., Hollman, P.C.H., Gruppen, H., (2009b). Combined Normal-Phase and Reversed-Phase Liquid Chromatography/ESI-MS as a Tool To Determine the Molecular Diversity of A-type Procyanidins in Peanut Skins. *Journal of Agricultural and Food Chemistry*, 57(14), 6007-6013.
- Chai, W.M., Shi, Y., Feng, H.L., Qiu, L., Zhou, H.C., Deng, Z.W., Yan, C.L., Chen, Q.X., (2012). NMR, HPLC-ESI-MS, and MALDI-TOF MS Analysis of Condensed Tannins from *Delonix regia* (Bojer ex Hook.) Rat and Their Bioactivities. *Journal of Agricultural and Food Chemistry*, 60(19), 5013-5022.
- De Bruyne, T., Pieters, L., Witvrouw, M., De Clercq, E., Vanden Berghe, D., Vlietinck, A.J., (1999). Biological evaluation of proanthocyanidin dimers and related polyphenols. *Journal of Natural Products*, 62(7), 954-958.
- Deutsch, E.W., (2012). File Formats Commonly Used in Mass Spectrometry Proteomics. *Molecular & Cellular Proteomics*, 11(12), 1612-1621.
- Faver, J.C., Zheng, Z., Merz, K.M., (2011). Model for the fast estimation of basis set superposition error in biomolecular systems. *Journal of Chemical Physics*, 135(14).

398 Faver, J.C., Zheng, Z., Merz, K.M., (2012). Statistics-based model for basis set superposition
 399 error correction in large biomolecules. *Physical Chemistry Chemical Physics*, 14(21), 7795-
 400 7799.

401 Feliciano, R.P., Krueger, C.G., Shanmuganayagam, D., Vestling, M.M., Reed, J.D., (2012).
 402 Deconvolution of matrix-assisted laser desorption/ionization time-of-flight mass
 403 spectrometry isotope patterns to determine ratios of A-type to B-type interflavan bonds in
 404 cranberry proanthocyanidins. *Food Chemistry*, 135(3), 1485-1493.

405 Foo, L.Y., Lu, Y.R., Howell, A.B., Vorsa, N., (2000a). The structure of cranberry
 406 proanthocyanidins which inhibit adherence of uropathogenic P-fimbriated *Escherichia coli* in
 407 vitro. *Phytochemistry*, 54(2), 173-181.

408 Foo, L.Y., Lu, Y.R., Howell, A.B., Vorsa, N., (2000b). A-type proanthocyanidin trimers from
 409 cranberry that inhibit adherence of uropathogenic P-fimbriated *Escherichia coli*. *Journal of*
 410 *Natural Products*, 63(9), 1225-1228.

411 Fu, C., Loo, A.E.K., Chia, F.P.P., Huang, D., (2007). Oligomeric proanthocyanidins from
 412 mangosteen pericarps. *Journal of Agricultural and Food Chemistry*, 55(19), 7689-7694.

413 Gu, L., Kelm, M., Hammerstone, J.F., Beecher, G., Cunningham, D., Vannozzi, S., Prior, R.L.,
 414 (2002). Fractionation of polymeric procyanidins from lowbush blueberry and quantification
 415 of procyanidins in selected foods with an optimized normal-phase HPLC-MS fluorescent
 416 detection method. *Journal of Agricultural and Food Chemistry*, 50(17), 4852-4860.

417 Gu, L., Kelm, M.A., Hammerstone, J.F., Zhang, Z., Beecher, G., Holden, J., Haytowitz, D., Prior,
 418 R.L., (2003). Liquid chromatographic/electrospray ionization mass spectrometric studies of
 419 proanthocyanidins in foods. *Journal of Mass Spectrometry*, 38(12), 1272-1280.

420 Hammerstone, J.F., Lazarus, S.A., Mitchell, A.E., Rucker, R., Schmitz, H.H., (1999).
 421 Identification of procyanidins in cocoa (*Theobroma cacao*) and chocolate using high-
 422 performance liquid chromatography mass spectrometry. *Journal of Agricultural and Food*
 423 *Chemistry*, 47(2), 490-496.

424 Hellstrom, J.K., Torronen, A.R., Mattila, P.H., (2009). Proanthocyanidins in common food
 425 products of plant origin. *Journal of Agricultural and Food Chemistry*, 57(17), 7899-7906.

426 Hokkanen, J., Mattila, S., Jaakola, L., Pirttila, A.M., Tolonen, A., (2009). Identification of
 427 phenolic compounds from lingonberry (*Vaccinium vitis-idaea* L.), bilberry (*Vaccinium*
 428 *myrtillus* L.) and hybrid bilberry (*Vaccinium x intermedium* Ruthe L.) leaves. *Journal of*
 429 *Agricultural and Food Chemistry*, 57(20), 9437-9447.

430 Howell, A.B., (2002). Cranberry proanthocyanidins and the maintenance of urinary tract health.
 431 *Critical Reviews in Food Science and Nutrition*, 42(3 Suppl), 273-278.

432 Hümmer, W., Schreier, P., (2008). Analysis of proanthocyanidins. *Molecular Nutrition & Food*
 433 *Research*, 52(12), 1381-1398.

434 Irungu, J., Go, E.P., Zhang, Y., Dalpathado, D.S., Liao, H.X., Haynes, B.F., Desaire, H., (2008).
 435 Comparison of HPLC/ESI-FTICR MS versus MALDI-TOF/TOF MS for glycopeptide
 436 analysis of a highly glycosylated HIV envelope glycoprotein. *Journal of the American*
 437 *Society for Mass Spectrometry*, 19(8), 1209-1220.

438 Kallio, H., Yang, W., Liu, P., Yang, B., (2014). Proanthocyanidins in wild sea buckthorn
 439 (*Hippophae rhamnoides*) berries analyzed by reversed-phase, normal-phase, and hydrophilic
 440 interaction liquid chromatography with UV and MS detection. *Journal of Agricultural and*
 441 *Food Chemistry*, 62(31), 7721-7729.

442 Kazusa DNA Research Institute, C., Japan, (2008). MSGet Manual, Komics Wiki.

443 Kelm, M. A., Johnson, J. C., Robbins, R. J., Hammerstone, J. F., Schmitz, H. H., (2006). High-
 444 performance liquid chromatography separation and purification of cacao (*Theobroma cacao*
 445 L.) procyanidins according to degree of polymerization using a diol stationary phase. *Journal*
 446 *of Agricultural and Food Chemistry*, 54, 1571-1576.

447 Kessner, D., Chambers, M., Burke, R., Agus, D., Mallick, P., (2008). ProteoWizard: open source
 448 software for rapid proteomics tools development. *Bioinformatics*, 24(21), 2534-2536.

449 Krueger, C.G., Reed, J.D., Feliciano, R.P., Howell, A.B., (2013). Quantifying and characterizing
 450 proanthocyanidins in cranberries in relation to urinary tract health. *Analytical and*
 451 *Bioanalytical Chemistry*, 405(13), 4385-4395.

452 Kylli, P., Nohynek, L., Puupponen-Pimia, R., Westerlund-Wikstrom, B., Leppanen, T., Welling,
 453 J., Moilanen, E., Heinonen, M., (2011). Lingonberry (*Vaccinium vitis-idaea*) and European
 454 cranberry (*Vaccinium microcarpon*) proanthocyanidins: isolation, identification, and
 455 bioactivities. *Journal of Agricultural and Food Chemistry*, 59(7), 3373-3384.

456 Lazarus, S.A., Adamson, G.E., Hammerstone, J.F., Schmitz, H.H., (1999). High-performance
 457 liquid Chromatography/Mass spectrometry analysis of proanthocyanidins in foods and
 458 beverages. *Journal of Agricultural and Food Chemistry*, 47(9), 3693-3701.

459 Le Roux, E., Doco, T., Sarni-Manchado, P., Lozano, Y., Cheynier, V., (1998). A-type
 460 proanthocyanidins from pericarp of *Litchi chinensis*. *Phytochemistry*, 48(7), 1251-1258.

461 Lin, L.Z., Harnly, J., Zhang, R.W., Fan, X.E., Chen, H.J., (2012). Quantitation of the
 462 hydroxycinnamic acid derivatives and the glycosides of flavonols and flavones by UV
 463 absorbance after identification by LC-MS. *Journal of Agricultural and Food Chemistry*,
 464 60(2), 544-553.

465 Lin, L.Z., Harnly, J.M., (2012). Quantitation of Flavanols, Proanthocyanidins, Isoflavones,
 466 Flavanones, Dihydrochalcones, Stilbenes, Benzoic Acid Derivatives Using Ultraviolet
 467 Absorbance after Identification by Liquid Chromatography-Mass Spectrometry. *Journal of*
 468 *Agricultural and Food Chemistry*, 60(23), 5832-5840.

469 Lin, L.Z., Sun, J., Chen, P., Monagas, M.J., Harnly, J.M., (2014). UHPLC-PDA-ESI/HRMSn
 470 profiling method to identify and quantify oligomeric proanthocyanidins in plant products.
 471 *Journal of Agricultural and Food Chemistry*, 62(39), 9387-9400.

472 Miller, K.B., Stuart, D.A., Smith, N.L., Lee, C.Y., McHale, N.L., Flanagan, J.A., Ou, B., Hurst,
 473 W.J., (2006). Antioxidant activity and polyphenol and procyanidin contents of selected
 474 commercially available cocoa-containing and chocolate products in the United States.
 475 *Journal of Agricultural and Food Chemistry*, 54(11), 4062-4068.

476 Nunez, V., Gomez-Cordoves, C., Bartolome, B., Hong, Y.J., Mitchell, A.E., (2006). Non-
 477 galloylated and galloylated proanthocyanidin oligomers in grape seeds from *Vitis vinifera* L.
 478 cv. Graciano, Tempranillo and Cabernet Sauvignon. *Journal of the Science of Food and*
 479 *Agriculture*, 86(6), 915-921.

480 Ohnishi-Kameyama, M., Yanagida, A., Kanda, T., Nagata, T., (1997). Identification of catechin
 481 oligomers from apple (*Malus pumila* cv. Fuji) in matrix-assisted laser desorption/ionization
 482 time-of-flight mass spectrometry and fast-atom bombardment mass spectrometry. *Rapid*
 483 *Communications in Mass Spectrometry*, 11(1), 31-36.

484 Sarnoski, P.J., Johnson, J.V., Reed, K.A., Tanko, J.M., O'Keefe, S.F., (2012). Separation and
 485 characterisation of proanthocyanidins in Virginia type peanut skins by LC-MSn. *Food*
 486 *Chemistry*, 131(3), 927-939.

487 Wan, K.X., Vidavsky, I., Gross, M.L., (2002). Comparing similar spectra: from similarity index
 488 to spectral contrast angle. *Journal of the American Society for Mass Spectrometry*, 13(1), 85-
 489 88.

490 Wang, S., Farahani, M.R., Baig, A.Q., Sajjad, W., (2016). The Sadhana polynomial and the
 491 Sadhana index of polycyclic aromatic hydrocarbons PAHK. *Journal of Chemical and*
 492 *Pharmaceutical Research*, 8(6), 526-531.

493 Wojdyło, A., Carbonell-Barrachina, A.A., Legua, P., Hernandez, F., (2016). Phenolic
 494 composition, ascorbic acid content, and antioxidant capacity of Spanish jujube (*Ziziphus*
 495 *jujube* Mill.) fruits. *Food Chemistry*, 201, 307-314.

496 Zhang, M., Harrington, P.D., (2014). Simultaneous quantification of Aroclor mixtures in soil
 497 samples by gas chromatography/mass spectrometry with solid phase microextraction using
 498 partial least-squares regression. *Chemosphere*, 118C, 187-193.

499 Zhang, M., Harrington, P.d.B., (2013). Automated pipeline for classifying Aroclors in soil by gas
 500 chromatography/mass spectrometry using modulo compressed two-way data objects. *Talanta*,
 501 117, 483-491.

502 Zhang, M., Harrington, P.d.B., (2015). Application of chemometrics to resolve overlapping mass
 503 spectral peak clusters between trichloroethylene and its deuterated internal standard. *Rapid*
 504 *Communications in Mass Spectrometry*, 29(9), 789-794.

505 Zhang, M., Sun, J., Chen, P., (2015). FlavonQ: An Automated Data Processing Tool for Profiling
 506 Flavone and Flavonol Glycosides with Ultra-High-Performance Liquid Chromatography-
 507 Diode Array Detection-High Resolution Accurate Mass-Mass Spectrometry. *Analytical*
 508 *Chemistry*, 87(19), 9974-9981.

509

510 **Figure Captions**

511 Figure 1. UV chromatogram at 278 nm for Fuji apple sample: peaks are labeled by peak
512 detection process (A) and peaks are labeled by proanthocyanidin peak extraction process (B).
513 False positive peaks are labeled with red color.

514 Figure 2. Similarity analysis by comparing full UV-vis spectra of each chromatographic peak
515 with UV-vis spectrum of catechin (inset).

516 Figure 3. Mass spectra to illustrate isotopic distribution for (A) trimer B-type PA $[M-H]^-$ in red,
517 (B) trimer A-type PA $[M-H]^-$ in yellow, (C) hexamer B-type PA $[M-2H]^{2-}$ in green, (D) 3
518 mixtures at ratio of 1:0.55:0.1, and (E) comparison of deconvolution spectrum with experimental
519 spectrum in blue.

520

521

522

Figure 1. UV chromatogram at 278 nm for Fuji apple sample: peaks are labeled by peak detection process (A) and peaks are labeled by proanthocyanidin peak extraction process (B). False positive peaks are labeled with red color.

Figure 2. Similarity analysis by comparing full UV-vis spectra of each chromatographic peak with UV-vis spectrum of catechin (inset).

Figure 3. Mass spectra to illustrate isotopic distribution for (A) trimer B-type PA $[M-H]^-$ in red, (B) trimer A-type PA $[M-H]^-$ in yellow, (C) hexamer B-type PA $[M-2H]^{2-}$ in green, (D) 3 mixtures at ratio of 1:0.55:0.1, and (E) comparison of deconvolution spectrum with experimental spectrum in blue.

Table 1. Comparison of proanthocyanidins identification results in seven samples: computer program versus manual interpretation.

DP	proanthocyanidin	[M-H] ⁻ (<i>m/z</i>)	mol. formula	presence in foods ^a	
				program identification	manual identification
monomer	epiafzelechin	273.0762	C ₁₅ H ₁₃ O ₅	A, C, M	C, L
	catechin/epicatechin	289.0712	C ₁₅ H ₁₃ O ₆	ALL	ALL
	epiafzelechin gallate	425.0873	C ₂₂ H ₁₇ O ₉	A, D, G, J, L, M	NONE
	catechin/epicatechin-gallate	441.0821	C ₂₂ H ₁₇ O ₁₀	G, J, M	G, M
dimer	EA/EF—A—EC	559.1240	C ₃₀ H ₂₃ O ₁₁	G, M	NONE
	EA/EF—EC	561.1397	C ₃₀ H ₂₅ O ₁₁	G, M	G, M
	EC—A—EC	575.1189	C ₃₀ H ₂₃ O ₁₂	A, C, D, G, J, L, M	A, C, D, L, M
	EC—EC	577.1346	C ₃₀ H ₂₅ O ₁₂	A, C, D, G, J, L, M	A, D, G, M
	EC—A—EG	591.1139	C ₃₀ H ₂₃ O ₁₃	G, M	NONE
	EC—EG	593.1295	C ₃₀ H ₂₅ O ₁₃	G, M	G
	EG—A—EG	607.1088	C ₃₀ H ₂₃ O ₁₄	L	NONE
	(EA/EF—A—EA/EF)g	695.1401	C ₃₇ H ₂₇ O ₁₄	G, M	NONE
	(EA/EF—EA/EF)g	697.1557	C ₃₇ H ₂₉ O ₁₄	G, M	NONE
	(EA/EF—A—EC)g	711.1350	C ₃₇ H ₂₇ O ₁₅	C, D, G, J, L, M	NONE
	(EA/EF—EC)g	713.1506	C ₃₇ H ₂₉ O ₁₅	A, D, G, J, M	NONE
	(EC—A—EC)g	727.1299	C ₃₇ H ₂₇ O ₁₆	L, M	NONE
	(EC—EC)g	729.1455	C ₃₇ H ₂₉ O ₁₆	D, G, M	G
	(EC—EC)2g	881.1565	C ₄₄ H ₃₃ O ₂₀	G	G
trimer	EA—EA—EC	833.2082	C ₄₅ H ₃₇ O ₁₆	M	M
	EA—A—EC—EC	847.1874	C ₄₅ H ₃₅ O ₁₇	G, L, M	L, M
	EA/EF—EC—EC	849.2031	C ₄₅ H ₃₇ O ₁₇	G, M	G, M
	EC—A—EC—A—EC	861.1666	C ₄₅ H ₃₃ O ₁₈	L	NONE
	EC—EC—A—EC	863.1823	C ₄₅ H ₃₅ O ₁₈	A, C, D, G, J, L, M	C, L, M
	EC—EC—EC	865.1979	C ₄₅ H ₃₇ O ₁₈	A, C, D, G, J, L, M	A, C, D, G, J L, M
	EC—EC—A—EG	879.1772	C ₄₅ H ₃₅ O ₁₉	G	NONE
	(EA/EF—EC—A—EC)g	999.1984	C ₅₂ H ₃₉ O ₂₁	G	NONE
	(EA/EF—EC—EC)g	1001.2140	C ₅₂ H ₄₁ O ₂₁	G	NONE
	(EC—EC—A—EC)g	1015.1933	C ₅₂ H ₃₉ O ₂₂	G	NONE

	EC)g				
	(EC—EC—EC)g	1017.2089	C ₅₂ H ₄₁ O ₂₂	G	G
	(EC—EC—EC)2g	1169.2199	C ₅₉ H ₄₅ O ₂₆	G	G
tetramer	EA/EF—EA/EF— EC—EC	1121.2715	C ₆₀ H ₄₉ O ₂₂	M	NONE
	EA/EF—EC—A— EC—EC	1135.2508	C ₆₀ H ₄₇ O ₂₃	L, M	L
	EA/EF—EC—EC— EC	1137.2664	C ₆₀ H ₄₉ O ₂₃	L, M	M
	EC—A—EC—A— EC—EC	1149.2301	C ₆₀ H ₄₅ O ₂₄	L	C, L
	EC—EC—A— EC—EC	1151.2457	C ₆₀ H ₄₇ O ₂₄	A, C, D, G, J, L, M	C, L
	EC—EC—EC—EC	1153.2614	C ₆₀ H ₄₉ O ₂₄	A, C, D, G, J, L, M	A, D, G, J, M
pentamer	EC—EC—A— EC—EC—EC	1439.3091	C ₇₅ H ₅₉ O ₃₀	L	L
	EC—EC—EC— EC—EC	1441.3248	C ₇₅ H ₆₁ O ₃₀	A, J, M	M
hexamer [*]	EA/EF—EA/EF— EC—EC—EC—EC	1697.3983	C ₉₀ H ₇₃ O ₃₄	G	NONE
	EC—EC—EC— EC—EC—EC	1729.3881	C ₉₀ H ₇₃ O ₃₆	A, D, G, M	NONE
	EC—EC—EC— EC—EG—EG	1761.3780	C ₉₀ H ₇₃ O ₃₈	G	NONE
	(EC—EC—EC— EC—EC—EC)2g	2033.4101	C ₁₀₄ H ₈₁ O ₄₄	G	NONE
octamer [*]	EC—EC—EC— EC—EG—EG— EG—EG	2305.5149	C ₁₂₀ H ₉₇ O ₄₈	A, D, J, M	NONE

^aAbbreviations: A, apple; C, cranberry extract; D, dark chocolate; G, grape seed extract; J, jujube; L, litchi; M, mangosteen; ALL, all tested samples; NONE, none of tested samples; DP, degree of polymerization; g, galloyl; EC, EA, EG, and EF, epicatechin/catechin, epiafzelechin/afzelechin, epigallocatechin/gallocatechin, and epifisetinidol/fisetinidol, respectively. The signal unit1→unit2 or unit1→A→unit2 expresses the units bonded by B-type (4,8- or 4,6-) bond or A-type (plus additional C2—O—C7- or C2—O—C5-) bond, respectively.

*Ions are in double charged deprotonated molecules form ([M-2H]²⁻).

Table 2. Chromatographic peak analysis results by proposed program

sample name	number of peaks detected by program	number of proanthocyanidin candidate peaks identified by program	number of false positive peaks
cranberry	78	9	2
dark chocolate	52	28	3
Fuji apple	64	11	2
grape seed extract	71	67	0
jujube fruit	51	30	5
litchi fruit pericarp	72	48	2
mangosteen fruit pericarp	72	49	3

Table 3. Quantitation result for proanthocyanidins in Fuji apple sample

peak ID	RT (min)	initial peak ID	total concentration ¹	concentration of B type ¹	concentration of A type ¹	Concentration with different DPs ¹			
						mono	di	tri	tetra
1 [*]	7.61	9	0	0	0	0	0	0	0
2 [*]	9.32	13	0	0	0	0	0	0	0
3	13.09	21	25.8	25.8	0	15.6	10.2	0	0
4 [#]	13.89	23	0	0	0	0	0	0	0
5 [#]	14.06	24	0	0	0	0	0	0	0
6	14.30	25	34.2	34.2	0	34.2	0	0	0
7 [*]	17.78	32	0	0	0	0	0	0	0
8	18.30	33	14.8	13.5	1.3	1.0	4.4	9.4	0
9	19.75	34	16.8	15.5	1.3	3.2	1.4	1.1	11.1
10 [*]	21.79	36	0	0	0	0	0	0	0
11 [*]	25.14	42	0	0	0	0	0	0	0
total			91.6	89.0	2.6	54.0	16.1	10.5	11.1

¹ Concentration equivalent to catechin, mg/100g (dry weight, DW);^{*} Peaks with proanthocyanidins which are below the limit of quantitation (8.24 mg/100g).[#] False positive peaks

531

532

Table 4. Quantitation result summaries for proanthocyanidins in 7 samples

sample name	total concentration ¹	total concentration of B type ¹	total concentration of A type ¹	total concentration with different DPs ¹						
				mono	di	tri	tetra	penta	hexa	octa
cranberry	350	0	307	42.9	210.2	97.4	0	0	0	0
dark chocolate	106	16.9	1.3	87.4	8.9	6.0	3.3	0	0	0
Fuji apple	91.6	35.0	2.6	54.0	16.1	10.5	11.1	0	0	0
grape seed extract	22973	9429	435	13109	8894	947	20.9	0	3.3	0
jujube fruit	1611	875	46.5	690	433	284	197	6.6	0.4	0.4
litchi fruit pericarp	2455	81.0	1490	884	741	514	306	12.2	0	0
mangosteen fruit pericarp	6387	2800	405	3182	1908	864	406	27.2	0	0

¹ Concentration equivalent to catechin, mg/100g.

Graphic Abstract