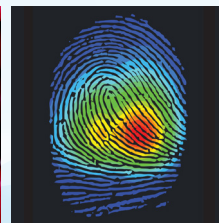


Identification and Quantification of Potential Adulterants in Cranberry Fruit Juice Dry Extracts using Absorbance-Transmittance Excitation-Emission (A-TEEM) Spectroscopy



Application Note
Food Sciences
FL-2021-05-05

Introduction

Dietary supplements derived from Cranberry (*Vaccinium macrocarpon*) promote urinary tract health and may reduce the risk of recurrent urinary tract infections (UTIs) in healthy women (1). The key beneficial compounds responsible are thought to be A-type proanthocyanidins (PACs), which are complex oligomers and polymers of flavan-3-ols.

Unscrupulous dietary supplement suppliers may adulterate cranberry with cheaper ingredients that mimic the analytical signature in conventional assays. Importantly, current chromatographic methods, which are expensive and time-consuming, may typically be limited to thresholds of 10% or more adulterant: Cranberry (w:w) due to problems resolving shared overlapping components (2). Further, the more commonly applied UV method for PAC detection, known as dimethylaminocinnamaldehyde (DMAC) adopted by the industry for the quantification of total PACs, is unable to effectively differentiate between cranberry and common adulterants (2). In 2019, The United States Pharmacopeial (USP) published in PF 45(6) a new monograph for proposal for Cranberry Fruit Juice Dry Extract, which defines the requirements for the identity, purity and composition for this dietary ingredient. In this application note we evaluate the ability to identify and quantify some key adulterants. These materials include extracts derived from Grape Seed (*Vitis vinifera* L), Pine Bark (*Pinus* sp.) and Peanut Skin (*Arachis hypogaea* L). It is important to keep in mind that adulterated cranberry products may not only be less effective (best case), but may cause a life-threatening allergic response (worst case), if the adulteration is due to peanut skin extract derived from Peanut Skin.

The Aqualog A-TEEM spectrometer was hypothesized to be a suitable method, owing to both its high sensitivity to anthocyanins, flavan-3-ols and PACs, as well as its fast acquisition rates (<45 s) and lack of any requirements for reagents, derivatization agents, columns and mobile phase solvents. The experiments were designed to determine both the quantification limits of A-TEEM spectroscopy and the discrimination accuracy for a selection of adulterant materials. Samples included USP reference materials for Maritime Pine Bark Dry Extract and Grape Seed Oligomeric Proanthocyanidin, as well as commercial dry extracts derived from Cranberry Fruit Juice, Grape Seed, Pine Bark

and Peanut Skin (Table 1). Data were evaluated using unsupervised multivariate decomposition analyses of each material, as well as concentration-supervised regression and discrimination analyses of potential adulterant spiking (grape seed, pine or peanut skin dry extracts). We report on the Aqualog A-TEEM spectrometer's ability to positively identify and sensitively quantify the spiked adulterant materials in the cranberry.

Materials and Methods

Samples and Preparation

Table 1 lists the samples used in this study with respect to the general category and the genus and species of origin. Highlighted in bold are the USP reference standards for Oligomeric Grape Seed Dry Extract and Maritime Pine Dry Extract, and the samples selected as authentic materials for Cranberry Fruit Dry Extract and Peanut Skin Dry Extract.

Sample ID	Sample type	Botanical name
GR-1	Grape Seed Extract	<i>Vitis vinifera</i>
GR-2 (ref)	Grape Seed Extract	<i>Vitis vinifera</i>
GR-3	Grape Seed Extract	<i>Vitis vinifera</i>
PB-1	Pine Bark Extract	<i>Pinus radiata</i>
PB-2	Pine Bark Extract	<i>Pinus radiata</i>
PB-3	Pine Bark Extract	<i>Pinus sylvestris</i>
PB-4	Pine Bark Extract	<i>Pinus massoniana</i>
PB-5	Pine Bark Extract	<i>Pinus pinaster</i>
PB-6	Pine Bark Extract	<i>Pinus pinaster</i>
PB-7 (ref)	Pine Bark Extract	<i>Pinus pinaster</i>
PS-1 (ref)	Peanut Skin Extract	<i>Arachis hypogaea</i>
PS-2	Peanut Skin Extract	<i>Arachis hypogaea</i>
CR-1	Cranberry Fruit Juice Extract	<i>Vaccinium macrocarpon</i>
CR-2	Cranberry Fruit Juice Extract	<i>Vaccinium macrocarpon</i>
CR-3	Cranberry Fruit Juice Extract	<i>Vaccinium macrocarpon</i>
CR-4 (ref)	Cranberry Fruit Juice Extract	<i>Vaccinium macrocarpon</i>
CR-5	Cranberry Fruit Juice Extract	<i>Vaccinium macrocarpon</i>
CR-6	Cranberry Fruit Juice Extract	<i>Vaccinium macrocarpon</i>
CR-7	Cranberry Fruit Juice Extract	<i>Vaccinium macrocarpon</i>

Table 1: Sample descriptions for cranberry extracts and potential adulterants used in this study.

All samples arrived as dry powders and were stored in darkness at room temperature. Samples were dissolved in 50% EtOH pH2 solution at a stock concentration of 1 mg/ml by vortexing for 30 s. Final solutions for A-TEEM analysis were diluted 200 fold. All samples were evaluated relative to a blank cell of 50% EtOH pH 2 solvent.

A-TEEM Measurements

All A-TEEM measurements were made with a HORIBA Aqualog UV-800C (Piscataway, NJ) using Aqualog v. 4.2 software. Samples were equilibrated to, and maintained at 20°C and stirred in the sample chamber. The excitation and absorbance range was 240-700 nm at 5 nm increments, and emission range was 250-800 nm at 4.66 nm, and interpolated to 5 nm increments. The CCD was binned by 8 pixels and the Gain was medium. All instrumental and sample bias corrections were applied, including dark offset removal, blank solvent subtraction, NIST-traceable excitation and emission spectral correction, inner filter effect correction and water Raman scattering unit (Figure 1.) CIE 1931 x/y color index representations of (A) the Cranberry (CR1-7) extracts and (B) the Pine Bark (PB1-7) extracts and (C) a summary of all extracts from Table 1, also including USP reference materials indicated by the colored arrows. All CIE values were corrected for the w:v dilution factor for normalization. Integration time was 0.15 s for a total of <45 s per A-TEEM acquisition. Samples were analyzed using a HORIBA Fast-01 autoinjector with 10 ml vials maintained at 20°C and with 4 repeat injections per sample.

Multivariate Analyses

All multivariate analyses were performed using Eigenvector Inc. Solo v8.8. A-TEEM data preprocessing for PARAFAC analysis included: Non-negative constraints for scores and loadings and masking of the first and second order

Rayleigh lines (16 and 32 nm, respectively). A-TEEM data block pre-processing for Principal Components Analysis and Extreme Gradient Boost regression and discrimination included: Unfolding the 3-way matrix into a 2-way array and class-dependent automatic clutter removal, centering and scaling. The concentration data block for the adulteration experiments was mean-centered.

Results

A-TEEM Characterization of Cranberry, Pine Bark, Grape Seed and Peanut Skin Dry Extracts

One of the most powerful features of the A-TEEM is the ability to use the Absorbance and Transmittance data to evaluate the color and CIE Chromaticity of solutions. In Figure 1 we first compare the CIE 1931 chromaticity x/y plots for each of the Cranberry (A) and Pine Bark (B) extract solutions prepared using the same w:v dilution, noting the CIE values are corrected for the dilution factor. Clearly all of the CR1-7 extracts exhibit RGB color locations in the orange to purple-red regions with y-values <0.35 and x-values >0.4. The CR-4 authentic material is shown at x/y (0.5177/0.31318). The Pine Bark extracts (PB1-7) (USP Reference Material) exhibit color indices in the yellow-orange-brown regions with greater y values than the CR samples. PB-7 is located at x/y (0.3448/0.35623). Figure 1 C summarizes the CIE x/y coordinates for all samples from Table 1 to emphasize that each of the potential adulterant materials exhibit much lower purple red content compared to the CR samples, which is attributed to the significant anthocyanin content of the CR samples compared to the adulterants.

Consistent with the unique CIE index variations for each sample shown in Figure 1, Figure 2 shows that each sample also exhibits a unique Principal Component Analysis (PCA) score cluster index with respect to PC1 and PC2.

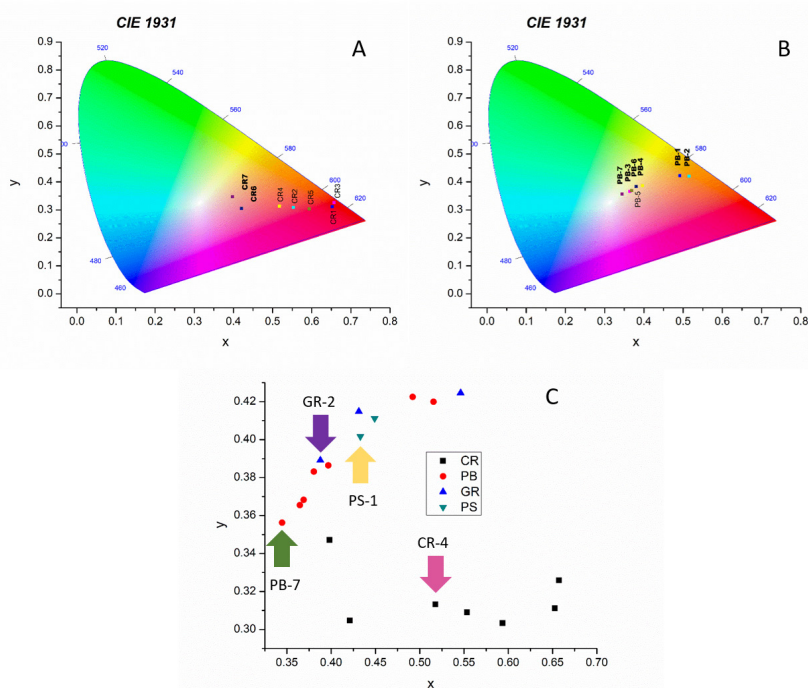


Figure 1: CIE 1931 x/y color index representations of (A) the Cranberry (CR1-7) extracts and (B) the Pine Bark (PB1-7) extracts and (C) a summary of all extracts from Table 1, also including USP reference materials indicated by the colored arrows. All CIE values were corrected for the w:v dilution factor.

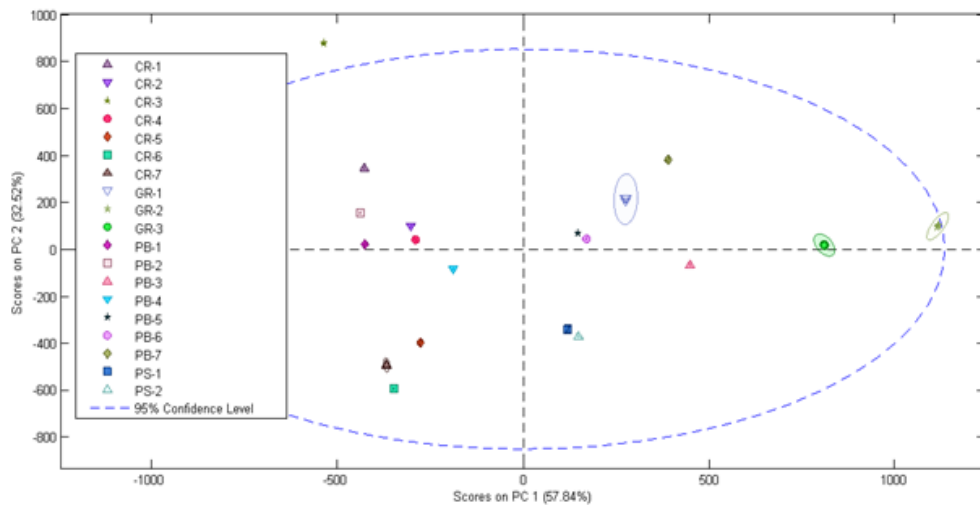


Figure 2: Principal Components Analysis of all 19 samples from Table 1. Each sample is represented in quadruplicate and defined by a 99% confidence Ellipse interval, noting in some cases the ellipse is smaller than the symbol size as scaled. All samples are represented to the right of the figure legend.

Each sample is identified with a unique symbol as indicated in the legend, and also defined by a 99% confidence ellipse. In many cases the ellipse may be smaller than the symbol size as scaled.

Figure 3 shows that each of the reference USP reference standard (GR-2 and PB-7) and authentic materials (CR-4 and PS-1) prepared using the same w:v dilution exhibit both some similar and other significantly varying features with respect to their A-TEEM contours in the UV excitation region (240-400 nm). Compared are samples (A) Peanut Skin (PS-1), (B) Pine Bark (PB-7), (C) Cranberry (CR-4) and (D) Grape Seed (GR-2). Each extract exhibits a relatively broad major bimodal contour feature with excitation/emission peaks around 240/315 nm and 280/315 nm. PS-1 (A) shows a very minor contour around 330/380nm. PB-7 (B) exhibits a second contour with similar coordinates but slightly more intense than in PS-1. Compared to PS-1 and PB-7, CR-4 shows the most intense complex minor contours extending to emission regions from 380 nm to beyond 450 nm. GR-2 shows the relatively weakest minor contours around 315/370 nm.

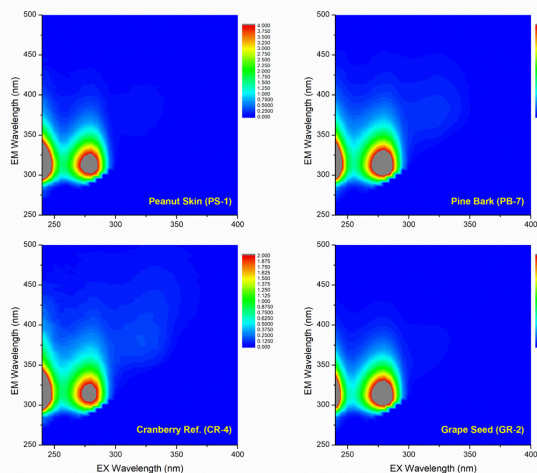


Figure 3. A-TEEM Contour Plots for the USP reference material extracts for (A) Peanut Skin (PS-1), (B) Pine Bark (PB-7), (C) Cranberry (CR-4) and (D) Grape Seeds (GR-2). Peak contours (gray) represent off scale values in order to emphasize contours in the longer excitation and emission wavelength regions.

Figure 4 further evaluates the reference USP reference standards and authentic materials (CR-4, PB-7, GR-2 and PS-1) by comparing their integrated excitation (thick lines) and emission (thin line profiles) extracted from the A-TEEM contours from Figure 2.

The clearest result is that there are strong differences in the fluorescence intensities, with the strongest being GR-2, followed by PB-7, PS-1 then CR-4 respectively. This is significant because in addition to spectral profile variation among samples seen in Figure 2, the A-TEEM analyses are also very sensitive to the fluorescence intensity variation associated with the unique quantum yield sample properties.

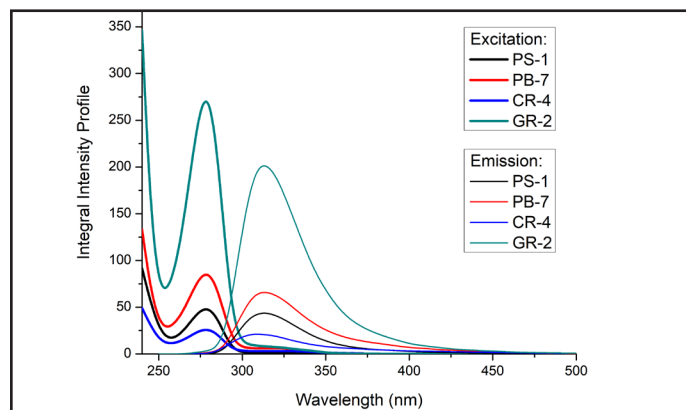


Figure 4: Integrated excitation (thick lines) and emission (thin lines) profiles extracted from the EEM contour plots of Figure 1 for the USP reference materials PS-1, PB-7, CR-4 and GR-2.

Unsupervised PARAFAC Score-based Regression Analysis of Cranberry Adulterant Spiking Experiment

Having established that each of the USP reference standards and authentic materials could be distinguished based on both CIE and unsupervised PCA decomposition, a spiking experiment was undertaken to evaluate the ability to determine adulterant detection limits. The CR-4 solution was diluted in a reciprocating series of sample preparations with either GR-2 or PB-7 each ranging from 1 to 100% contribution to the final solution. The data from both dilution series were analyzed by Parallel Factor Analysis (PARAFAC) models each comprising three components.

PARAFAC is a tri-linear decomposition method that yields unique components each defined by orthogonally related excitation and emission spectral loadings and their scores, which represent the total intensity contribution for a given sample. The component number assignments are made in descending order (Component 1, Component 2...) based on their relative score contribution for all samples in the model. Figure 5 shows the PARAFAC component loading contours representing the cross product of each component's excitation and emission spectral loadings. It is clear that the spectral contours are different, especially comparing components 1 and 2 for the Grape Adulteration to the Pine Bark Adulteration. This indicates both adulterants exhibit unique spectral intensity profiles. Component 3 being the minor more red shifted component appears similar for the two adulterants.

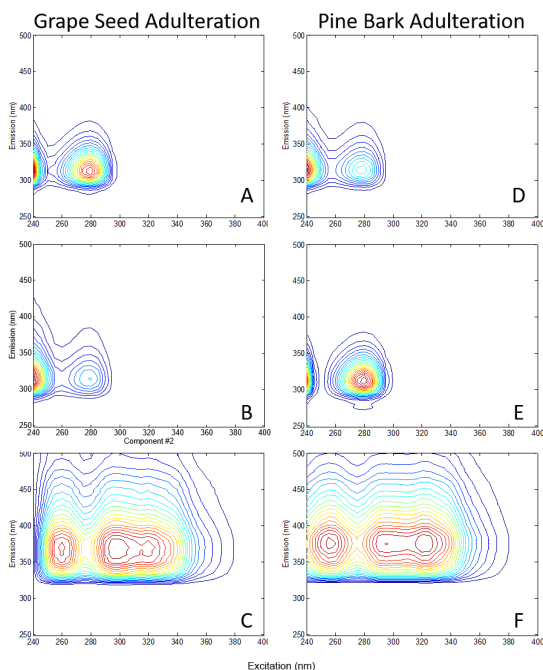


Figure 5: Contour plots from three components Parallel Factor Analysis models where Cranberry (CR-4) was adulterated on a v:v basis (from 1-100%) with either (A-C) GR-2 or (D-F) PB-7. Component 1 is represented in (A, D), Component 2 in (B, E) and Component 3 in (C, F), respectively. The contours represent the cross product of the excitation and emission loadings for each component. Components numbers are assigned based on their total intensity contribution to all samples in the model.

The PARAFAC scores for the two adulteration series were systematically evaluated and regressed against the % composition for GR-2 (A) and PB-7 (B) in Figure 6. The scores for Component 1 were averaged for the four replicates per sample and a linear regression line was determined relating the scores to the % observed composition. This equation was then inverted to determine the Predicted % values and the regression results shown for GR-2 and PB-7, respectively. In both models it was possible to observe the effects of $\geq 1\%$ adulteration based on the 0 and 1% data points. The GR-2 regression yielded significantly better linear fit statistics with respect to the adjusted coefficient of variance (adjusted R2) and Standard Errors for the Slope and Intercept parameters. Further, the 95% Prediction Band was narrower for GR-2 than PB-7.

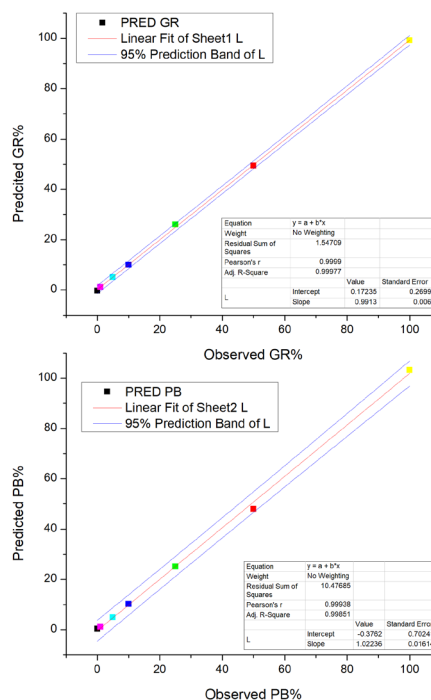


Figure 6: Linear regression analyses of the 'unsupervised' Parallel Factor Analysis scores for Cranberry reference sample (CR-4) spiked with (A) Grape Seed (GR-2) extract and (B) Pine Bark extract. The adulterants' solutions were evaluated from 1 to 100% of the total volume reciprocally with the CR-4 solution. The linear regression statistics are located in the respective Figure legends and are defined graphically with 95% prediction bands in Panels A and B.

Overall, these results were consistent with the observation in Figure 2 that GR-2 shows a higher fluorescence intensity and hence a higher signal to noise for the same w:v ratio.

Supervised Extreme Gradient Boost Analyses of Cranberry Adulterant Spiking Experiment

Given the ability to distinguish $\geq 1\%$ of the adulterants GR-2 and PB-7 in Cranberry Fruit Juice Dry Extract (CR-4) using an unsupervised method to relate the adulterant concentration and A-TEEM spectral data, we investigated whether the regression statistics and detection limits could be improved by using a concentration supervised regression method. The A-TEEM data from Figure 6 were analyzed using Extreme Gradient Boost (XGB) regression to directly predict the % adulteration of GR-2 and PB-7 in Cranberry Fruit Dry Extract (CR-4). In contrast to the PARAFAC analysis which included separate models for each adulterant the XGB regression combined the A-TEEM files for both adulterants in the same data block to determine if they could be distinguished simultaneously. Here a Calibration set comprising 75% of the samples was used to predict the adulterant % concentrations in a prediction set comprising 25% of the samples. The linear regressions for the prediction set were evaluated for GR-2 (A) and PB-7 (B). Compared to the unsupervised models in Figure 6 the regression statistics were strongly improved for the XGB prediction including the adjusted R2 and the Standard Errors for the Slope and Intercept parameters. Interestingly, as observed in the unsupervised PARAFAC models, the XGB GR-2 model exhibited better linear fit statistics than the PB-7 XGB model. Table 2 compares the limit of quantification (LOQ) for the PARAFAC and XGB models for the spiked experiments with GR-2 and PB-7

also including Calibration, Cross Validation and Prediction data LOQ for GR-2 was <3%, compared to around 7.48 % for PB-7. The LOQ values were reduced greatly in the XGB models to around 2.2E-04 and 5.3E-02% for GR-2 and PB-7, respectively.

PARAFAC	SE Intercept	Slope	LOD	LOQ
GR-2	2.6992E-01	9.9130E-01	8.9855E-01	2.9652E+00
PB-7	7.0241E-01	1.0224E+00	2.2673E+00	7.4819E+00
XGB				
GR2	2.0292E-05	9.9900E-01	6.7030E-05	2.2120E-04
PB-7	4.7800E-03	1.0000E+00	1.5774E-02	5.2054E-02

Table 2: Limits of Detection (LOD) and Quantification (LOQ) from unsupervised Parallel Factor Analysis (PARAFAC) and supervised Extreme Gradient Boost (XGB) regression analyses calculated from the adulterant spiking data from Figures 6 and 8, respectively.

Based on the success of the supervised XGB regression method, a concentration threshold-based XGB discrimination model was evaluated using the same A-TEEM data as in Figure 7.

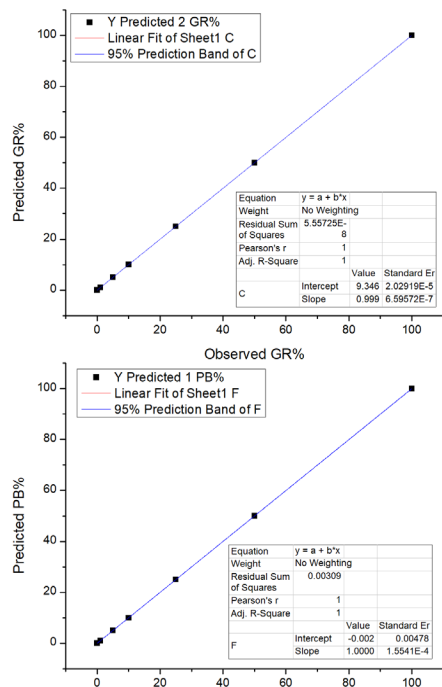


Figure 7: Concentration supervised Extreme Gradient Boost (XGB) regression analysis of the adulterant spike data from Figure 6. Panels A and B show the results of Grape Seed (GR%) and Pine Bark (PB%), respectively. The regression line, fit statistics and 95% Prediction Bands for each adulterant are shown in the respective Panels.

Three class groups were assigned based on the following concentration ranges: 100% Cranberry Extract (CR-4), 0-100% Grape Seed Extract (GR-2) and 0-100% Pine Bark Extract (PB-7). The confusion matrix in Figure 8 shows all True Positive and True Negative results were 1.000 and all False Positive and False Negative results were 0.000 for all

samples in the Calibration, Cross Validation and Prediction data. The results support the ability to both identify the unique adulterants, as well as simultaneously identify cranberry samples with ≥1% of either adulterant.

```

XGBDA Classification Using Rule: Pred Most Probable

MODEL RESULTS
Confusion Matrix:
Class: TPR FPR TNR FNR N Err P F1
100% Cranberry 1.00000 0.00000 1.00000 0.00000 6 0.00000 1.00000 1.00000
Grape Seed 1-100% 1.00000 0.00000 1.00000 0.00000 18 0.00000 1.00000 1.00000
Pine Bark 1-100% 1.00000 0.00000 1.00000 0.00000 18 0.00000 1.00000 1.00000

Confusion Table:
Actual Class
100% *1 Grape*2 Pine *3
Predicted as 100% Cranberry 6 0 0
Predicted as Grape Seed 1-100% 0 18 0
Predicted as Pine Bark 1-100% 0 0 18
Predicted as Unassigned 0 0 0

Key:
100% *1 = 100% Cranberry
Grape*2 = Grape Seed 1-100%
Pine *3 = Pine Bark 1-100%

CV RESULTS
Confusion Matrix (CV):
Class: TPR FPR TNR FNR N Err P F1
100% Cranberry 1.00000 0.00000 1.00000 0.00000 6 0.00000 1.00000 1.00000
Grape Seed 1-100% 1.00000 0.00000 1.00000 0.00000 18 0.00000 1.00000 1.00000
Pine Bark 1-100% 1.00000 0.00000 1.00000 0.00000 18 0.00000 1.00000 1.00000

Confusion Table (CV):
Actual Class
100% *1 Grape*2 Pine *3
Predicted as 100% Cranberry 6 0 0
Predicted as Grape Seed 1-100% 0 18 0
Predicted as Pine Bark 1-100% 0 0 18
Predicted as Unassigned 0 0 0

Key:
100% *1 = 100% Cranberry
Grape*2 = Grape Seed 1-100%
Pine *3 = Pine Bark 1-100%

PREDICTION RESULTS
Confusion Matrix:
Class: TPR FPR TNR FNR N Err P F1
100% Cranberry 1.00000 0.00000 1.00000 0.00000 2 0.00000 1.00000 1.00000
Grape Seed 1-100% 1.00000 0.00000 1.00000 0.00000 6 0.00000 1.00000 1.00000
Pine Bark 1-100% 1.00000 0.00000 1.00000 0.00000 6 0.00000 1.00000 1.00000

Confusion Table:
Actual Class
100% *1 Grape*2 Pine *3
Predicted as 100% Cranberry 2 0 0
Predicted as Grape Seed 1-100% 0 6 0
Predicted as Pine Bark 1-100% 0 0 6
Predicted as Unassigned 0 0 0

Key:
100% *1 = 100% Cranberry
Grape*2 = Grape Seed 1-100%
Pine *3 = Pine Bark 1-100%
    
```

Figure 8: Confusion matrix results for a class-supervised Extreme Gradient Boost Discrimination Analysis (XGBDA) model of the adulterant spike experiment data shown in Figure 7. Three class groups were defined including 100% Cranberry (CR-4) samples, samples containing (GR-2) Grape Seed Extract (1-100% extract, and samples containing (PB-7) Pine Bark Extract (1-100%). The results refer to the "Most Probable" prediction rule noting the results were equivalent for the "Strict" prediction rule. Compared are the probabilities for True Positive results (TPR), False Positive results (FPR), True Negative results (TNR) and False Negative results (FNR) for the Calibration, Cross Validation and Prediction data.

Conclusions

The Aqualog A-TEEM spectrometer has the capacity to specifically identify and sensitively quantify key adulterants that are most commonly detected in Cranberry dietary supplements, including Pine Bark, Grape Seed and Peanut Skin extracts. Compared to other more expensive and time-consuming analytical methods, the A-TEEM technique exhibits up to an order of magnitude or more higher sensitivity (≤1%) and reduction in analysis time (less than 1 minute). The analyses and reporting can also be automated for use in laboratory operations in near real-time. Thus, A-TEEM spectroscopy should be considered a valuable technique for developing product specific methods for identity and composition required for cGMP as well, and for adulterant detection in the food and dietary supplement industry.

References

1. T. Brendler and S. Gafner (2017) Adulteration of Cranberry (Vaccinium macrocarpon). Botanical Adulterants Bulletin. www.botanicaladulterants.org
2. J.H. Cardellina II and S. Gafner (2018) Cranberry Products Laboratory Guidance. Document. www.botanicaladulterants.org

info.sci@horiba.com

www.horiba.com/scientific

USA: +1 732 494 8660
UK: +44 (0)1604 542 500
China: +86 (0)21 6289 6060
Taiwan: +886 3 5600606

France: +33 (0)1 69 74 72 00
Italy: +39 06 51 59 22 1
India: +91 80 41273637
Brazil: +55 (0)11 2923 5400

Germany: +49 (0) 6251 8475 0
Japan: +81(75)313-8121
Singapore: +65 (0)6 745 8300
Other: +33 (0)1 69 74 72 00

HORIBA
Scientific