Identifying Cranberry Juice Consumers with Predictive OPLS-DA Models of Plasma Metabolome and Validation of Cranberry Juice Intake Biomarkers in a Double-Blinded, Randomized, Placebo-controlled, Cross-over Study

Shaomin Zhao [1], Haiyan Liu [2], Zhihua Su[3], Christina Khoo [2], Liwei Gu[1, *]

1 Food Science and Human Nutrition Department, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, Florida 32611, United States

2 Ocean Spray Cranberries, Inc., Lakeville-Middleboro, Massachusetts 02349, United States

3 Department of Statistics, College of Liberal Arts and Sciences, University of Florida, Gainesville, Florida 32611, United States

This study was registered with ClinicalTrials.gov ID: NCT02860468

(URL: https://clinicaltrials.gov/ct2/show/NCT02860468)

This manuscript contains supplemental data.

Funding source: this research was funded in part by Ocean Spray Cranberries, In.

* Corresponding author:

Tel: (352) 294-3730; Fax: (352) 392-9467; Email: Lgu@ufl.edu

Received: 05/12/2019; Revised: 02/04/2020; Accepted: 07/04/2020

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/mnfr.201901242.

List of Abbreviations

HRMS	High resolution mass spectrometry
OPLS-DA	Orthogonal partial least squares-discriminant analysis
PCA	Principle component analysis
QC	Quality control
VIP	Variable influence on projection

KEYWORDS: Cranberry; Metabolomics; UHPLC-Q-Orbitrap-HRMS; Orthogonal partial least squares-discriminant analysis; Procyanidins

ABSTRACT

Background: Cranberry products were tested in many clinical trials on their efficacy to prevent urinary tract infections, however, low compliance and high dropout have plagued these trials due to cranberries' bitterness, astringency, and lack of method to verify consumption. It is possible to use predictive multivariate models built upon validated biomarkers to verify human consumption of a food using a nutrimetabolomics approach.

Objectives: 1) to validate previously identified cranberry juice intake biomarkers, 2) to build predictive OPLS-DA models to classify cranberry juice consumers, and 3) discover additional discriminant metabolites.

Methods: A 21-day double blinded, randomized, placebo-controlled, cross-over study was conducted among healthy young women aging 18-29. Plasma were collected at baseline and after 3-day and 21-day consumption of cranberry or placebo juice. Plasma metabolome were analyzed using UHPLC-Q-Orbitrap-HRMS.

Results: 18 discriminant metabolites in positive mode and 18 discriminant metabolites in negative mode from a previous 3-day open-label study were re-discovered in the present blinded study. Predictive OPLS-DA models was built to identify cranberry juice consumers and non-consumers. The models were able to identify cranberry juice consumers over placebo juice group with 96.9% correction rates after 3-day consumption at both positive and negative mode. This present study revealed 84 and 109 additional discriminant metabolites in positive and negative mode respectively. Twelve of them were tentatively identified.

Conclusion: Cranberry juice consumers were classified with high correction rates using predictive OPLS-DA models built upon validated plasma biomarkers. Additional biomarkers were tentatively identified. These OPLS-DA models and biomarkers provided an objective approach to verify participant compliance in future clinical trials.

INTRODUCTION

Cranberries have many health benefits but are best known for their ability to prevent the urinary tract infections which affect 50-60% of women in their lifetime [1]. However, clinical trials on the effectiveness of cranberry products against the occurrence and recurrence of urinary tract infection yield conflicting results. Some trials showed positive outcome while results from others were negative. Results from the same research group sometime were contradictive [2]. The 2012 Cochrane review concluded that low compliance and high dropout were common problems in clinical trials of cranberry products due to their bitterness and astringency [3]. There is no objective method to evaluate cranberry consumption. Currently, self-reporting by participant is used to verify the compliance of cranberry consumptions in all clinical trials and human intervention studies.

Nutrimetabolomics is a rapidly growing field which applies integrative metabolomic analysis in nutritional studies [4]. There is a trend to replace traditional dietary assessment method with metabolomic techniques using specific food intake biomarkers. Metabolomics studies have been done to identify plasma or urinary biomarkers after consumption of pea, almond, or certain type of diet [5-7]. Data mining methods like principle component analysis (PCA), and partial least squares discriminant analysis (PLS-DA) and it's variant orthogonal PLS-DA (OPLS-DA) are commonly used to build multivariate models for biomarker identification. However, any statistical differences should be validated with independent studies, which were rarely done partially due to high cost of metabolomic research, leaving question marks on the repeatability of identified biomarkers and model performance.

A previous untargeted metabolomic study from our lab reported that consumption of cranberry juice for three days significantly altered plasma metabolome in 18 women in an open-label cross-over intervention. This study revealed 39 and 42 discriminant metabolic features in plasma in negative mode and positive mode, respectively, as potential biomarkers of cranberry juice consumption [8]. The current independent study was conducted to test the reproductivity of previously found metabolic features and the performance of predictive OPLS-DA models to classify cranberry juice consumers and non-consumers. Both biomarker validity and the predictability of the predictive OPLS-DA models were evaluated in 16 women using placebo-controlled double-blinded crossover design.

MATERIALS AND METHODS

Chemicals and materials

Acetonitrile, methylene chloride, methanol, acetic acid, formic acid and acetone were purchased from Fischer Scientific Co. (Pittsburgh, PA, USA). Creatine-D_3, L-leucine-D_{10}, L-tryptophan-2, 3, 3-D_3, caffeine-D_3 were from CDN Isotopes (Pointe-Claire, Quebec, Canada).

Dosage information

Cranberry juice cocktail (double strength, 54% juice) and placebo juices in 250-mL bottles were provided by Ocean Spray Cranberries (Lakeville-Middleboro, MA, USA). Both types of juice were coded in 3-digit random numbers by Ocean Spray. Placebo juice was 100% apple juice with colorant. One cannot tell the difference between placebo and cranberry juice by appearance. Procyanidin, anthocyanin, and sugar content of cranberry juice and placebo juice are shown in supplemental Table S1. Participants were asked to drink two bottles of cranberry juice or placebo per day for 21 days then after 14-day wash-out, switched to

alternative regimen and repeated the protocol. This dosage was matched previous open-labeled study and could be achieved by consuming single strength cranberry juice cocktail (27%) sold in the market [8].

Subjects and study design

The human study was approved by Institutional Review Boards at the University of Florida. The experimental design is depicted in **Figure 1** and **Figure S1**. Twenty-two healthy female college students between 21–29 years old with BMI of 18.5–25.0 were recruited in Gainesville, Florida. The sample size matched a previous open-label study [8]. Each subject was provided with a list of foods that contained significant amount of procyanidins, including cranberries, apples, grapes, blueberries, chocolate and plums. They were advised to avoid these foods during the 1st –10th day and the rest of the study. The study was randomized, double-blinded, placebo-controlled. On the morning of the 10th day, a baseline blood was drawn from all human subjects after an overnight fast. Participants were then randomly assigned into two groups (n=11 for each group) and asked to consume either cranberry juice or placebo. Complete randomization was applied to randomly assign participant with cranberry juice or placebo [9]. Both participants and researchers were blinded. Each participant had equal chance to receive cranberry juice or placebo. Six bottles (250 ml/bottle) of juice were given to participants to drink in the morning and evening of the 10th, 11th and 12th days. On the morning of 13th day, all subjects returned to the clinical unit and a blood sample was collected from participants after fasting overnight. Another supply of either cranberry juice or placebo were given to participants to drink in the morning and evening from the 13th to 30th day. Appointments were scheduled weekly from the 14th -29th day for juice distribution and feedback collection. On the 30th day morning, a blood was

drawn from all participants after an overnight fast. After a two-week wash out, participants switched to the alternative regimen and repeated the protocol. Blood samples were centrifuged at 2000 g for 10 min at 4 °C to obtain plasma. All plasma samples were aliquoted and kept in a −80 °C freezer until analysis. One participant was removed from the study due to missed appointments. Seventeen participants completed all six blood collections and plasma from these 17 participants were analyzed.

Multiple strategies were applied to ensure high compliance in the present study. All participants were undergraduate and graduate students from University of Florida. Each participant was well educated about the scientific merit of this study and the importance of compliance before the treatment. During the study, reminders of drinking juice were sent at least twice per week to help with the compliance. Questionnaires collecting the information of how many bottles missed during each session were completed by participants. According to the questionnaire, more than 95% compliance were reached for both 3-day and 21-day consumption. Therefore, after removing the outlier sample points from one participant due to low compliance, we assumed 100% compliance of the rest 16 participants, which enabled us to test the model predictability with this dataset.

UHPLC-Q-Orbitrap-HRMS analyses

Plasma sample preparation and UHPLC-Q-Orbitrap-HRMS analyses followed published method [8]. Briefly, plasma sample was mixed with acetonitrile: acetone: methanol (8:1:1, v: v: v) and centrifuged to precipitate protein. Supernatant was dried by nitrogen stream and reconstituted with 0.1% formic acid water. L-leucine-D10, L-tryptophan-2, 3, 3-D3, creatine-D3 and caffeine-D3 were added as internal standards. All 80 samples were prepared in the same manner. Quality control (QC) samples, which were pooled plasma from American Red

Cross, were prepared and analyzed with experimental samples to monitor the stability and validity of instrumental acquisition.

Data pre-processing

UHPLC-HRMS data were converted to mzXML using MSConvert from ProteoWizard [10] and then processed using MZmine 2.12 [11]. Peaks in each sample were extracted, deconvoluted, and deisotoped. Alignment using join aligner algorithm was conducted with a 10-ppm tolerance for m/z values and 0.2 min tolerance for retention time. Gap filling using peak finder algorithm was performed to fill in missing peaks. The resultant data set was imported into SIMCA (Version 14.0, Umetrics, Umea, Sweden) for multivariate statistical analysis.

Statistical analysis

Data mining strategy is depicted in supplemental Figure S2. Data acquired using positive and negative ionization were pareto scaled/mean centered and log-transformed before multivariate statistical analyses. VIP was used to summarize the importance of X-variables on models and top 1500 VIP (VIP >1) variables of PLS-DA model were selected to build OPLS-DA model [12]. Validation of the model was tested using sevenfold internal cross-validation and permutation tests for 200 times using SIMCA.

Identification of discriminant metabolites

Discriminant metabolic features were identified based on their accurate masses and/or product ion spectra in negative and positive mode. HMDB was searched to assist metabolite identification [13]. Confidence level of identification was based on definitions by Schrimpe-Rutledge et al [14].

RESULTS AND DISCUSSION

A schematic study flow is depicted in **Figure 1**. Twenty-two participants were recruited at the beginning of the study. One participant was removed due to the missing of more than two appointments. Seventeen participants completed all six blood collections and plasma from these 17 participants were analyzed. One participant was evaluated to have low compliance based on questionnaire response. Data from the rest sixteen participants were included in the results. Nine of them received cranberry juice for the first treatment session, the rest 7 received placebo. All of them switched to the other type of juice for the second treatment session. Characteristics of participants including weight, BMI and age were presented in **Table S2**. There was no significant difference between groups.

All QC samples clustered on the PCA score plots (**Figure S3**), indicating the consistency and stability of UHPLC-Q-Orbitrap-HRMS analyses. L-leucine-D_{10}, L-tryptophan-2, 3, 3-D_3, creatine-D_3 and caffeine-D_3 in injected standards had low relative standard deviation shown in supplemental **Table S3**, suggesting the stability and validity of instrumental and data acquisition.

Multivariate analysis

Separation of two treatment groups were observed after both 3 days and 21 days juice consumption in OPLS-DA models (**Figure 2A, 2B** and **3A, 3B**). Parameters of OPLS-DA from positive and negative ionization were listed in **Table 1**. With 1500 X-variables R^2Y are over 0.9 for both positive and negative ionization 3-day and 21-day treatment, indicating at least 90% of the variance of Y variables was explained by these OPLS-DA models. Q^2 calculated from sevenfold cross-validation indicated goodness of prediction of OPLS-DA models. The cross-validated score plots showed a clear separation for two groups (**Figure 2C, 2D** and **3C,**

3D). The results after 3-days juice consumption were consistent with a previous open-label study. A 200-time permutation test was conducted to evaluate the goodness of fit of models (**Figure 4**). X-axis is the correlation coefficient between permuted and original response variables (denoted as "r"), which represents the degree of randomization of response variable y. At r = 0, y variables are completely randomized; while at r = 1, there is no permutation of y variables. Decrease of Q^2 was observed along with the decrease of r, suggesting the model did not overfit [15]. R^2 remained nearly unchanged with r, which is a normal phenomenon when there are large number of predictors and small number of observations in the models in biological sciences [16].

Predictive models comparing cranberry juice with placebo consumers

S-plots were applied to identify discriminant metabolites (**Figure 5 and 6**). Eighty-two discriminant metabolites were found in 3-day positive ionization model while 55 discriminant metabolites were found in 21-day model. Among them, 18 discriminant metabolites observed in the present study were shared with a previous 3-day open label study in positive ionization (**Table 2**) [8]. Fifteen of 18 shared discriminant metabolites increased after cranberry juice consumption in previous open labeled study, therefore were selected to build OPLS-DA model with better specificity for cranberry juice.

One hundred and seven discriminant metabolites were found in 3-day model and 85 in 21-day model from negative ionization. Among them, eighteen discriminant metabolites from current study were shared with those identified in a previous 3-day open-label study (**Table 3**). All these metabolites were found to increase after cranberry juice consumption.

For 21-day juice consumption groups, 9 shared discriminant metabolites increased after cranberry consumption were observed in positive ionization and another 12 shared discriminant metabolites were found in negative ionization. All these metabolites were included in **Table 2** and **3**. There was no new shared discriminant metabolites discovered in 21-day juice consumption groups. It is reasonable since previous data was obtained on 3-day juice consumption trial, which matched better with the present 3-day consumption data.

Predictive OPLS-DA models were built with the shared discriminant metabolites mentioned above using data from previous study. Cranberry juice consumers and apple juice consumers clustered on the right and left side of OPLS-DA score plots. For predictive OPLS-DA models with 15 and 18 X-variables (***Model PvC15+ and PvC18-***), Q^2 were larger than 0.7 for both positive and negative ionization models, indicating good predictability (**Table 1**). A 200-time permutation test was conducted to investigate the validity and predictability of OPLS-DA model (**Figure 7A, 7B**). The regression lines of R^2 and Q^2 dramatically decreased when the correlation coefficients between permuted and original response variables decreased, suggesting that the predictive OPLS-DA models did not overfit. Classification scatters (**Figure 8**) showed the prediction results of the present study data by the predictive models. Previous 3-day open labeled study data was used as training set, and current study data was used as test set. Cranberry juice group was classified as cranberry juice group (on the right side) while placebo group was classified as apple juice group (on the left side). Correct classification rates were 96.9% for 3-day juice consumption data acquired by positive and negative ionization, 93.8% for 21-day juice consumption data acquired by

positive ionization. The 21-day juice consumption data acquired by negative ionization had a correct classification rate of 90.6%.

Predictive models comparing cranberry consumers with baseline

Multivariate data analysis was done between cranberry juice group and baseline group using the same strategy used between placebo and cranberry juice groups. Seventy-five discriminant metabolites were identified after 3-day juice consumption in positive ionization, among them, 14 were shared with previous study and were used to build the predictive model in positive ionization (**Figure 9A**). All these 14 metabolites were included in **Table 2**. Seventy-eight discriminant metabolites were putatively identified after 21-day juice consumption (**Figure 9B**). In negative ionization, 92 discriminant metabolites were tentatively identified after 3-day juice consumption and 65 discriminant metabolites were identified after 21-day juice consumption (**Figure 10**). Fifteen metabolites were shared with those in previous study and was used to build the predictive model in negative ionization, all these 15 were included in **Table 3**. Therefore, no new shared discriminant metabolites were identified in the analysis between baseline samples and cranberry juice consumers. Predictive OPLS-DA models were built up with these 15 shared discriminant metabolites mentioned above using data from previous 3-day open label study with the baseline group on the left side and cranberry juice consumption group on the right. Parameters of the above two models were listed in **Table 1** as models ***BvC14+*** and ***BvC15-***. A 200-time permutation test was conducted and proved the validity and predictability of OPLS-DA model (**Figure 7C, 7D**). According to the classification scatters (**Figure 11**), both positive and negative ionization models showed better performance when classifying samples after 3-day cranberry juice consumption with correct classification rate of 93.8%. For the

classification of samples after 21-day cranberry juice consumption, the correct classification rates were 84.4% for negative ionization and 90.6% for positive ionization. The lower correct classification rates of the models ***BvC14+*** and ***BvC15-*** between baseline and cranberry juice consumer were not surprising. With less variables used, these models might not perform as well as models ***PvC15+*** and ***PvC18-***.

To help evaluating the general classification ability, models ***PvC15+***, ***BvC14+***, ***PvC18-*** and ***BvC15-*** were used to classify all samples. For model ***PvC15+ and PvC18-***, baseline samples were classified into apple juice consumption group (on the left side), the total correct classification rate of all 80 samples was 93.8% for both models (**Figure 12A, 12B**). For models ***BvC14+*** and ***BvC15-***, placebo group was classified as baseline (on the left), the total classification rate of 80 samples was 88.8% for positive ionization (**Figure 12C**) and 93.8% for negative ionization (**Figure 12D**).

Differences in study design and participants might explain some discriminant metabolites observed in previous study were not identified as discriminant metabolites in present study. Discriminant metabolites found in open label study consist of intake and effect biomarkers according to the FoodBALL consortium [17]. The effect biomarkers refer to the functional responses of human body to an exposure of food compounds which is influenced by the heterogeneity in biotransformation. Effect biomarkers were not expected in a double-blinded study design because placebo effect was minimized. Different sets of participants would also result in variation of discriminant metabolites. Other than a few procyanidin-rich foods, participants were not asked to avoid other foods during both studies. The diet differences may also contribute to variation of discriminant metabolites. Juices used in previous and present studies were prepared by the same manufacturer using same

protocol. However, Table 1S showed that cranberry juice used the present study contained less procyanidins and anthocyanins compared to the juices used in open-label study because phenolic content is affected by cultivar, growth location, and year of harvesting [18]. Nevertheless, the shared discriminant metabolites observed across studies had significant importance and could be used as biomarkers for future cranberry clinical studies. Our study showed the probability that cranberry consumers can be correctly identified by monitoring discriminant metabolites only instead of using global metabolomic approaches, which save time and cost.

Discriminant metabolites identification

Numbers of discriminant metabolites tentatively identified, unidentified, and used in the predictive models were summarized in **Table S4**. Discriminant metabolites were identified using accurate mass and/ or product ion spectrum and database searching according to a recommended approach [4]. Among 15 metabolites used in the ***PvC15+*** model (**Table 2**), 10 metabolites were putatively identified (4 were identified previously, 6 were newly identified). They included exogenous metabolites from cranberry consumption: quinic acid, 3-(hydroxyphenyl) propionic acid, (*S*)-Homostachydrine, ethyl (methylthio)methyl disulfide; and endogenous metabolites from cranberry consumption: glycerol 3-phosphate, dihydroxyquinoline, hippuric acid, hydroxypyruvic acid, 3,4-dihydroxyphenylglycol and guanidoacetic acid. Six metabolites were putatively identified (5 were identified previously, 1 were newly identified) among 18 metabolites used in the ***PvC18-*** model (**Table 3**). Quinic acid, catechol sulfate and vanilloloside were exogenous metabolites increased after cranberry consumption, while 2-chloromaleylacetate, 2-phenylacetamide, hippuric acid, and tyrosine were increased endogenous metabolites after cranberry consumption. Twelve

additional discriminant metabolites found in the present study (but not in previous study) were putatively identified *via* mass spectral analysis and database searching on HMDB. These metabolites were listed in **Table 4**. For positive ionization, pyroglutamic acid, as an endogenous metabolite was found to increase after apple juice consumption. For negative ionization, *S*-acetyl dihydroasparagusic acid, pyrocatechol, guaiacol were found to increase after cranberry consumption as exogenous metabolites. 3-isopropylmalate, 2-chloromaleylacetate, lanthionine ketimine, prolyl-hydroxyproline, tyramine-*O*-sulfate, (3,4,5,6-tetrahydroxyoxan-2-yl)methyl 4-hydroxybenzoate and {4-[2,3-dioxo-3-(2,4,6-trihydroxy-3-methoxyphenyl)propyl]-2-hydroxy-6-methoxyphenyl}oxidanesulfonic acid increased after cranberry consumption as endogenous metabolites. 4-deoxythreonic acid was found to increase after apple consumption as an endogenous metabolite. Additional discriminant metabolites not identified were listed in **Table S5** and **S6**.

Interestingly, among the shared discriminant metabolites (with previous open label study), four of them: 3-(hydroxyphenol) propionic acid, hippuric acid, hydroxyhippuric acid and catechol sulfate were from the microbial catabolism of procyanidins. This indicates that potentially there might be a very important role of gut microbiota on the degradation of procyanidins and these metabolites "stayed" as discriminant metabolites for both studies suggesting the consistence of microbial catabolism of procyanidins. There were 19 new discriminant metabolites putatively identified. Catechol, also known as pyrocatechol, naturally occur in fruits and vegetables, but also found to be a metabolite in some bacteria including *Escherichia*, *Mycobacterium* and *Pseudomonas* [19, 20]. Guaiacol, as a monomethyl ether of catechol, has the property of inducing cell proliferation by being a potent scavenger of reactive oxygen radical s[21]. Both were exogenous metabolites

increased after cranberry consumption. Anthocyanins contained a catechol ring in its structure and account for 50% of the oxygen radical scavenging capacity of cranberries [22]. It was unknown whether the increase of guaiacol was a result of anthocyanin metabolism, but our finding might reveal a potential mechanism of cranberries' antioxidant effect *in vivo*. Cranberry consumption increased 3,4-dihydroxyphenylglycol, which was a normal endogenous metabolite of norepinephrine, a catecholamine [23]. Free 3,4-dihydroxyphenylglycol was observed to decrease along with aging. One study also suggested the plasmatic level of 3,4-dihydroxyphenylglycol decreased among depressed patients [24]. Lanthionine ketimine is thought to possibly serve as a neurochemicals by binding specifically to brain membrane with high affinity [25]. Some recent study even suggested its ethyl ester derivative might play a neuroprotective role against neurodegenerative diseases like Alzheimer's disease [26]. Some endogenous metabolites were involved in amino acid metabolism like hydroxypyruvic acid, guanidoacetic acid, 3-isopropylmalate, prolyl-hydroxyproline and 2-phenylacetamide, which were intermediates in glycine, serine, glutamic acid and threonine metabolism. Tyrosine was found to increase after cranberry consumption in both studies. Along with the increase of tyrosine, tyramine-*O*-sulfate, which was formed from decarboxylation of tyrosine was found to increase as well. *S*-acetyl dihydroasparagusic acid was tentatively identified because it was the only match in the HMDB database by molecular ion. Ethyl (methylthio)methyl disulfide has been detected but not quantified in fruits making it a potential biomarker for fruit consumption according to the FooDB database (www.FooDB.ca). *(S)*-Homostachydrine is found in the fruits and seeds making it a potential marker of fruit consumption [13, 27]. 2-Chloromaleylacetate is a derivative of medium-chain Keto Acids [13]. (3,4,5,6-tetrahydroxyoxan-2-yl)methyl 4-hydroxybenzoate was generated by enzyme *via* glycoside-hydrolysis, and {4-[2,3-dioxo-3-

(2,4,6-trihydroxy-3-methoxyphenyl)propyl]-2-hydroxy-6-methoxyphenyl} oxidanesulfonic acid was generated by enzyme *via* 4-*O*-sulfation of phenolic compounds [13]. Both were predicted metabolites by BioTransformer [28]. Quinic acid, glycerol 3-phosphate, dihydroxyquinoline, vanilloloside, hippuric acid were identified and discussed in previous study. It is somehow too early to draw conclusions that any discriminant metabolite related to cranberry juice consumption contributed to health benefits, since all identifications were putative. However, these findings may provide information about the metabolic pathways impacted by cranberry juice consumption.

Hippuric acid was proposed as a potential biomarker of fruits and vegetable consumption [29, 30]. Exogenous metabolites derived from microbial catabolism of procyanidins found in the current study were not exclusive of the cranberry because they were also found after consumption of other foods such as cocoa [31]. Identification of biomarkers specific to cranberry juice intake remain a difficult task to be addressed in future research. Without specific biomarkers, specificity and prediction accuracy was achieved by multivariate predictive models built upon multiple metabolomic features.

CONCLUSIONS

The present study demonstrated for the first time that predictive models could be built with a small number of plasma metabolites to identify cranberry juice consumers and non-consumers. The high correct rate of identification suggested the possibility of using these models to evaluate the compliance of cranberry consumption among clinical trial participants. The objective evaluation of compliance may significantly improve the reliability of cranberry clinical trial results by excluding data from participants who did not consume

cranberry juices. Additional discriminant metabolites were putatively identified and could serve as key pieces in a puzzle to explore the mechanisms of cranberries to mitigate urinary tract infections. The discriminant metabolites identified in the present also provided clues on degradation of procyanidins and oligosaccharides in cranberries by gut microbiota in colon.

ACKNOWLEDGEMENT

Conflict of Interest: Shaomin Zhao and Liwei Gu reported grant support from Ocean Spray Cranberries Inc. Co-authors Christina Khoo and Haiyan Liu are employees of Ocean Spray Cranberries Inc. Zhihua Su has nothing to disclose.

REFERENCE

[1] Al-Badr, A., Al-Shaikh, G., Recurrent Urinary Tract Infections Management in Women: A review. *Sultan Qaboos University medical journal* 2013, *13*, 359-367.

[2] Zhao, S., Liu, H., Gu, L., American cranberries and health benefits - an evolving story of 25 years. *Journal of the science of food and agriculture* 2018.

[3] Jepson, R. G., Williams, G., Craig, J. C., Cranberries for preventing urinary tract infections. *Cochrane Database Syst Rev* 2012, *10*, CD001321.

[4] Ulaszewska, M. M., Weinert, C. H., Trimigno, A., Portmann, R., Andres Lacueva, C., Badertscher, R., Brennan, L., Brunius, C., Bub, A., Capozzi, F., Cialie Rosso, M., Cordero, C. E., Daniel, H., Durand, S., Egert, B., Ferrario, P. G., Feskens, E. J. M., Franceschi, P., Garcia-Aloy, M., Giacomoni, F., Giesbertz, P., Gonzalez-Dominguez, R., Hanhineva, K., Hemeryck, L. Y., Kopka, J., Kulling, S. E., Llorach, R., Manach, C., Mattivi, F., Migne, C., Munger, L. H., Ott, B., Picone, G., Pimentel, G., Pujos-Guillot, E., Riccadonna, S., Rist, M. J., Rombouts, C., Rubert, J., Skurk, T., Sri Harsha, P. S. C., Van Meulebroek, L., Vanhaecke, L., Vazquez-Fresno, R., Wishart, D., Vergeres, G., Nutrimetabolomics: An Integrative Action for Metabolomic Analyses in Human Nutritional Studies. *Molecular nutrition & food research* 2019, *63*, e1800384.

[5] P, S. C. S. H., Abdul Wahab, R., Cuparencu, C., Dragsted, L. O., Brennan, L., A Metabolomics Approach to the Identification of Urinary Biomarkers of Pea Intake. *Nutrients* 2018, *10*.

[6] Llorach, R., Garrido, I., Monagas, M., Urpi-Sarda, M., Tulipani, S., Bartolome, B., Andres-Lacueva, C., Metabolomics study of human urinary metabolome modifications after intake of almond (Prunus dulcis (Mill.) D.A. Webb) skin polyphenols. *J Proteome Res* 2010, *9*, 5859-5867.

[7] Rebholz, C. M., Lichtenstein, A. H., Zheng, Z., Appel, L. J., Coresh, J., Serum untargeted metabolomic profile of the Dietary Approaches to Stop Hypertension (DASH) dietary pattern. *Am J Clin Nutr* 2018, *108*, 243-255.

[8] Liu, H. Y., Garrett, T. J., Su, Z. H., Khoo, C., Gu, L. W., UHPLC-Q-Orbitrap-HRMS-based global metabolomics reveal metabolome modifications in plasma of young women after cranberry juice consumption. *Journal of Nutritional Biochemistry* 2017, *45*, 67-76.

[9] Lachin, J. M., Matts, J. P., Wei, L. J., Randomization in Clinical-Trials - Conclusions and Recommendations. *Control Clin Trials* 1988, *9*, 365-374.

[10] Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M. Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S. L., Nuwaysir, L. M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E. W., Moritz, R.

L., Katz, J. E., Agus, D. B., MacCoss, M., Tabb, D. L., Mallick, P., A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* 2012, *30*, 918-920.

[11] Pluskal, T., Castillo, S., Villar-Briones, A., Oresic, M., MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *Bmc Bioinformatics* 2010, *11*.

[12] Eriksson, L., Byrne, T., Johansson, E., Trygg, J., Vikström, C., *Multi-and megavariate data analysis basic principles and applications*, Umetrics Academy 2013.

[13] Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., Sajed, T., Johnson, D., Li, C. R., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y. J., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y. F., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., Scalbert, A., HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res* 2018, *46*, D608-D617.

[14] Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., McLean, J. A., Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *J Am Soc Mass Spectr* 2016, *27*, 1897-1905.

[15] Mahadevan, S., Shah, S. L., Marrie, T. J., Slupsky, C. M., Analysis of metabolomic data using support vector machines. *Anal Chem* 2008, *80*, 7562-7570.

[16] Cook, R. D., Weisberg, S., *Applied regression including computing and graphics*, John Wiley & Sons 2009.

[17] Gao, Q., Pratico, G., Scalbert, A., Vergeres, G., Kolehmainen, M., Manach, C., Brennan, L., Afman, L. A., Wishart, D. S., Andres-Lacueva, C., Garcia-Aloy, M., Verhagen, H., Feskens, E. J. M., Dragsted, L. O., A scheme for a flexible classification of dietary and health biomarkers. *Genes Nutr* 2017, *12*, 34.

[18] Liu, H. Y., Tayyari, F., Khoo, C., Gu, L. W., A H-1 NMR-based approach to investigate metabolomic differences in the plasma and urine of young women after cranberry juice or apple juice consumption. *J Funct Foods* 2015, *14*, 76-86.

[19] Yam, K. C., D'Angelo, I., Kalscheuer, R., Zhu, H., Wang, J. X. , Snieckus, V., Ly, L. H., Converse, P. J., Jacobs, W. R., Jr., Strynadka, N., Eltis, L. D., Studies of a ring-cleaving dioxygenase illuminate the role of cholesterol metabolism in the pathogenesis of Mycobacterium tuberculosis. *PLoS pathogens* 2009, *5*, e1000344.

[20] Balderas-Hernandez, V. E., Trevino-Quintanilla, L. G., Hernandez-Chavez, G., Martinez, A., Bolivar, F., Gosset, G., Catechol biosynthesis from glucose in Escherichia coli anthranilate-overproducer strains by heterologous expression of anthranilate 1,2-dioxygenase from Pseudomonas aeruginosa PAO1. *Microbial cell factories* 2014, *13*, 136.

[21] Mimurai, T., Yazaki, K., Sawaki, K., Ozawa, T., Kawaguchi, M., Hydroxyl radical scavenging effects of guaiacol used in traditional dental pulp sedation: reaction kinetic study. *Biomedical research* 2005, *26*, 139-145.

[22] Zheng, W., Wang, S. Y., Oxygen radical absorbing capacity of phenolics in blueberries, cranberries, chokeberries, and lingonberries. *J Agric Food Chem* 2003, *51*, 502-509.

[23] Elsworth, J. D., Roth, R. H., Redmond, D. E., Jr., Relative importance of 3-methoxy-4-hydroxyphenylglycol and 3,4-dihydroxyphenylglycol as norepinephrine metabolites in rat, monkey, and humans. *Journal of neurochemistry* 1983, *41*, 786-793.

[24] Loo, H., Scatton, B., Dennis, T., Benkelfat, C., Gay, C., Poirier-Littre, M. F., Garreau, M., Vanelle, J. M., Olie, J. P., Deniker, P., Study of noradrenaline metabolism in depressed patients by the determination of plasma dihydroxyphenylethylene glycol. *L'Encephale* 1983, *9*, 297-316.

[25] Cavallini, D., Ricci, G., Dupre, S., Pecci, L., Costa, M., Matarese, R. M., Pensa, B., Antonucci, A., Solinas, S. P., Fontana, M., Sulfur-containing cyclic ketimines and imino acids. A novel family of endogenous products in the search for a role. *European journal of biochemistry* 1991, *202*, 217-223.

[26] Koehler, D., Shah, Z. A., Hensley, K., Williams, F. E., Lanthionine ketimine-5-ethyl ester provides neuroprotection in a zebrafish model of okadaic acid-induced Alzheimer's disease. *Neurochemistry international* 2018, *115*, 61-68.

[27] Servillo, L., Giovane, A., Balestrieri, M. L., Ferrari, G., Cautela, D., Castaldo, D., Occurrence of pipecolic acid and pipecolic acid betaine (homostachydrine) in Citrus genus plants. *J Agric Food Chem* 2012, *60*, 315-321.

[28] Feunang, Y. D., *Cheminformatics Tools for Enabling Metabolomics*, University of Alberta 2017.

[29] Guerra, A., Folesani, G., Mena, P., Ticinesi, A., Allegri, F., Nouvenne, A., Pinelli, S., Del Rio, D., Borghi, L., Meschi, T., Hippuric acid in 24 h urine collections as a biomarker of fruits and vegetables intake in kidney stone formers. *Int J Food Sci Nutr* 2014, *65*, 1033-1038.

[30] van Dorsten, F. A., Grun, C. H., van Velzen, E. J., Jacobs, D. M., Draijer, R., van Duynhoven, J. P., The metabolic fate of red wine and grape juice polyphenols in humans assessed by metabolomics. *Molecular nutrition & food research* 2010, *54*, 897-908.

[31] Mayorga-Gross, A. L., Esquivel, P., Impact of Cocoa Products Intake on Plasma and Urine Metabolites: A Review of Targeted and Non-Targeted Studies in Humans. *Nutrients* 2019, *11*.

Table 1. Summary of parameters for OPLS-DA for human plasma after drinking cranberry juice and placebo juice.

Name and type of model	Ionization mode	Number of x-variables	Time and groups	N[a]	$R^2X(cum)$[b]	$R^2Y(cum)$[b]	$Q^2(cum)$[c]
OPLS-DA models built upon the present double-blinded study							
BLD PvC 3D+	Positive	1500	Placebo *vs.* Cranberry 3-day	1P[d]+1O[e]	0.253	0.950	0.815
BLD PvC 21D+			Placebo *vs.* Cranberry 21-day	1P[d]+1O[e]	0.186	0.984	0.844
BLD BvC 3D+			Baseline *vs.* Cranberry 3-day	1P[d]+1O[e]	0.248	0.970	0.859
BLD BvC 21D+			Baseline *vs.* Cranberry 21-day	1P[d]+1O[e]	0.239	0.950	0.837
BLD PvC 3D-	Negative	1500	Placebo *vs.* Cranberry 3-day	1P[d]+1O[e]	0.254	0.965	0.860
BLD PvC 21D-			Placebo *vs.* Cranberry 21-day	1P[d]+1O[e]	0.227	0.968	0.829
BLD BvC 3D-			Baseline *vs.* Cranberry 3-day	1P[d]+1O[e]	0.286	0.948	0.853
BLD BvC 21D-			Baseline *vs.* Cranberry 21-day	1P[d]+1O[e]	0.221	0.962	0.850
Predictive models built upon an open-label study using discriminant metabolites shared with double-blinded study							
OPL PvC15+	Positive	15	Placebo *vs.* Cranberry juice	1P[d]+1O[e]	0.929	0.767	0.725
OPL PvC18-	Negative	18	Placebo *vs.* Cranberry juice	1P[d]+1O[e]	0.778	0.902	0.870
OPL BvC14+	Positive	14	Baseline *vs.* Cranberry juice	1P[d]+1O[e]	0.890	0.825	0.772
OPL BvC15-	Negative	15	Baseline *vs.* Cranberry juice	1P[d]+1O[e]	0.839	0.891	0.862

[a] Number of components; [b] R^2X and R^2Y are the cumulative modeled variations in the X and Y matrix, respectively; [c] Q^2 is the cumulative predicted variation in the Y matrix; [d] Predictive component; [e] Orthogonal component.

Codes of models: "BLD PvC 3D+" denotes OPLS-DA model built upon the present double-blinded study with metabolites after 3-day juice consumption from positive ionization. "BLD BvC 3D-" represents OPLS-DA model built upon the present double-blinded study with metabolites after 3-day juice consumption from negative ionization. "OPL PvC15+" denotes predictive OPLS-DA models built upon an open-label study using discriminant metabolites shared with double-blinded study found increased after cranberry juice consumption from positive ionization mode. "OPL BvC14+" represents predictive OPLS-DA models built upon an open-label study using discriminant metabolites shared with double-blinded study from negative ionization. Shared discriminant metabolites were found in the present double-blinded study and a previous open-label study.

Table 2. List of shared discriminant metabolites in positive ionization; Metabolites No. 1 and 4-18 were used in predictive OPLS-DA models to identify cranberry juice consumers; Shared discriminant metabolites were found in the present double-blinded study and a previous open-label study.

No.	Retention time (min)	Detected Mass $[M-H]^-$	p[1] (contribution) [a]	p(corr)[1] (confidence) [b]	Increased after cranberry/placebo juice consumption	Putative identification in the present study	Identification level [c]
1	0.82	193.0705	-0.086	-0.717	Cranberry	Quinic Acid*	level 2
2	1.47	118.5258	0.075	0.818	Placebo	No match	
3	1.48	220.9737	0.086	0.839	Placebo	No match	
4	1.48	205.0015	0.077	0.822	Placebo	No match	
5	5.52	167.0693	-0.109	-0.880	Cranberry	3-(Hydroxyphenyl) propionic acid*	level 2
6	6.24	158.1173	-0.066	-0.596	Cranberry	(*S*)-Homostachydrine	level 2
7	7.92	159.5415	-0.070	-0.641	Cranberry	No match	
8	7.92	292.0140	-0.055	-0.624	Cranberry	No match	
9	7.92	173.0231	-0.061	-0.645	Cranberry	Glycerol 3-phosphate	level 2
10	7.92	219.5518	-0.119	-0.622	Cranberry	No match	
11	7.92	162.0547	-0.084	-0.678	Cranberry	Dihydroxyquinoline*	level 2
12	7.92	155.0124	-0.061	-0.624	Cranberry	Ethyl (methylthio)methyl disulfide	level 2
13	7.92	196.0390	-0.060	-0.611	Cranberry	No match	
14	7.92	180.0652	-0.084	-0.655	Cranberry	Hippuric acid*	level 2
15	7.92	105.0337	-0.081	-0.655	Cranberry	Hydroxypyruvic acid	level 2
16	7.92	150.5362	-0.057	-0.606	Cranberry	No match	
17	7.92	171.0494	-0.062	-0.615	Cranberry	3,4-Dihydroxyphenyl glycol	level 2
1	7.92	118.065	-0.080	-0.669	Cranberry	Guanidoacetic	level 2

8	2	acid

[a] Number inside the parentheses is the p[1] value obtained from OPLS-DA model based on cranberry juice *vs*. placebo.
[b] Number inside the parentheses is the p (corr)[1] value obtained from OPLS-DA model based on cranberry juice *vs.* placebo.
[c] Confidence levels of identification based on Schrimpe-Rutledge et al. [14].
*Identified in a previous study [8].

Table 3. List of shared discriminant metabolites in negative ionization; All were used in predictive OPLS-DA model to identify cranberry juice consumers; Shared discriminant metabolites were found in the present double-blinded study and a previous open-label study.

No.	Retention time	Detected Mass $[M-H]^-$	p[1] (contribution)[a]	p(corr)[1] (confidence)[b]	Increased after cranberry/placebo juice consumption	Putative identification in the present study
1	0.83	191.0553	-0.061	-0.697	Cranberry	Quinic acid*
2	1.90	159.0662	-0.075	-0.798	Cranberry	No match
3	2.69	159.0662	-0.061	-0.752	Cranberry	No match
4	4.71	161.9865	-0.117	-0.857	Cranberry	No match
5	6.60	272.9385	-0.064	-0.805	Cranberry	No match
6	6.60	256.9737	-0.063	-0.804	Cranberry	No match
7	6.60	400.9611	-0.085	-0.698	Cranberry	No match
8	6.60	190.9819	-0.057	-0.790	Cranberry	No match
9	6.60	188.9861	-0.057	-0.790	Cranberry	Catechol sulfate*
10	7.82	315.1083	-0.055	-0.635	Cranberry	Vanilloloside*
11	7.92	276.0271	-0.054	-0.765	Cranberry	No match
12	7.92	179.0555	-0.058	-0.752	Cranberry	No match
13	7.92	379.0906	-0.070	-0.722	Cranberry	No match
14	7.92	134.0610	-0.058	-0.756	Cranberry	2-Phenylacetamide
15	7.92	178.0505	-0.058	-0.752	Cranberry	Hippuric acid*
16	7.92	224.0557	-0.061	-0.762	Cranberry	No match
17	8.00	180.0664	-0.058	-0.698	Cranberry	Tyrosine*
18	8.28	182.0819	-0.055	-0.642	Cranberry	No match

[a] Number inside the parentheses is the p[1] value obtained from OPLS-DA model based on cranberry juice *vs.* placebo.

[b] Number inside the parentheses is the p (corr)[1] value obtained from OPLS-DA model based on cranberry juice *vs.* placebo.

[c] Confidence level of identification based on Schrimpe-Rutledge et al. [14].

*Identified in a previous study [8].

Table 4. Tentatively identified additional discriminant metabolites from present study (but not in previous study).

No.	Retention time	m/z	Identification [a]	Increased after cranberry/placebo juice consumption	Ionization	Placebo vs. Cranberry juice 3-day	Placebo vs. Cranberry juice 21-day	Baseline vs. Cranberry juice 3-day	Baseline vs. Cranberry juice 21-day
1	1.00	175.0609	3-Isopropylmalate	Cranberry	negative	1[b]	1	1	1
2	1.47	130.0336	Pyroglutamic acid	Placebo	positive	1	1	0	0
3	1.72	119.0353	4-Deoxythreonic acid	Placebo	negative	1	1	0	0
4	2.47	190.9653	2-Chloromaleylacetate	Cranberry	negative	0[b]	0	1	0
5	6.60	187.9787	Lanthionine ketimine	Cranberry	negative	1	0	0	0
6	6.60	192.9863	S-Acetyl dihydroasparagusic acid	Cranberry	negative	1	0	0	0
7	6.60	109.0287	Pyrocatechol	Cranberry	negative	1	0	0	0
8	7.50	227.0911	Prolyl-Hydroxyproline	Cranberry	negative	1	0	1	0
9	7.64	123.	Guaiacol	Cranberry	nega	1	0	0	0

10	7.92	04 49 21 6.02 47	Tyramine-O-sulfate	Cranberry	negative	1	1	1	1
11	8.03	29 9.07 71	(3,4,5,6-tetrahydroxyoxan-2-yl)methyl 4-hydroxybenzoate	Cranberry	negative	1	1	1	1
12	8.17	44 3.00 95	{4-[2,3-dioxo-3-(2,4,6-trihydroxy-3-methoxyphenyl)propyl]-2-hydroxy-6-methoxyphenyl}+oxidanesulfonic acid	Cranberry	negative	1	0	1	1

[a] All were level 2 identification.

[b] 1 denotes detection of a discriminant metabolite in a model. 0 indicates the absence of a discriminant metabolite in a model.

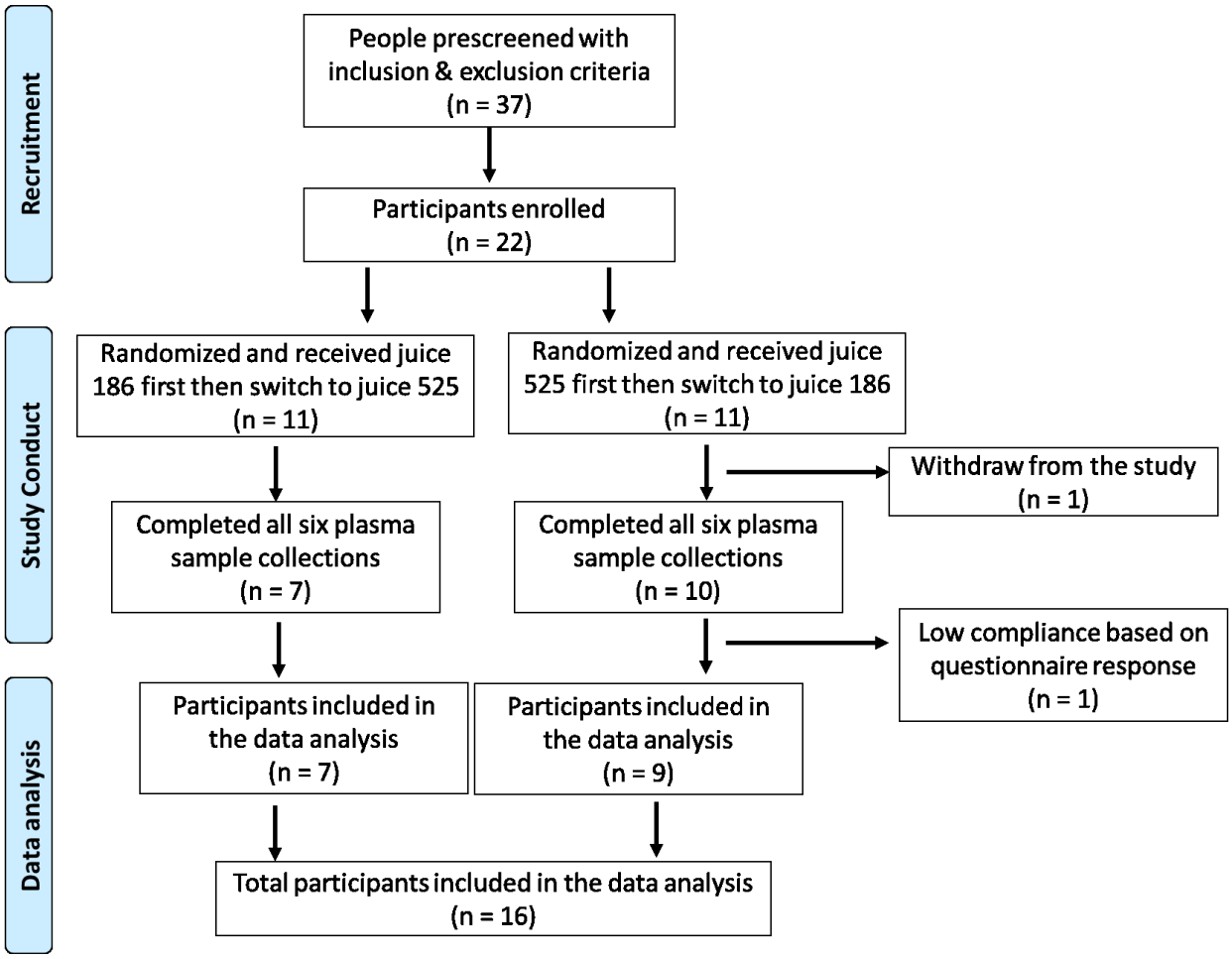

Figure 1. Study flow diagram.

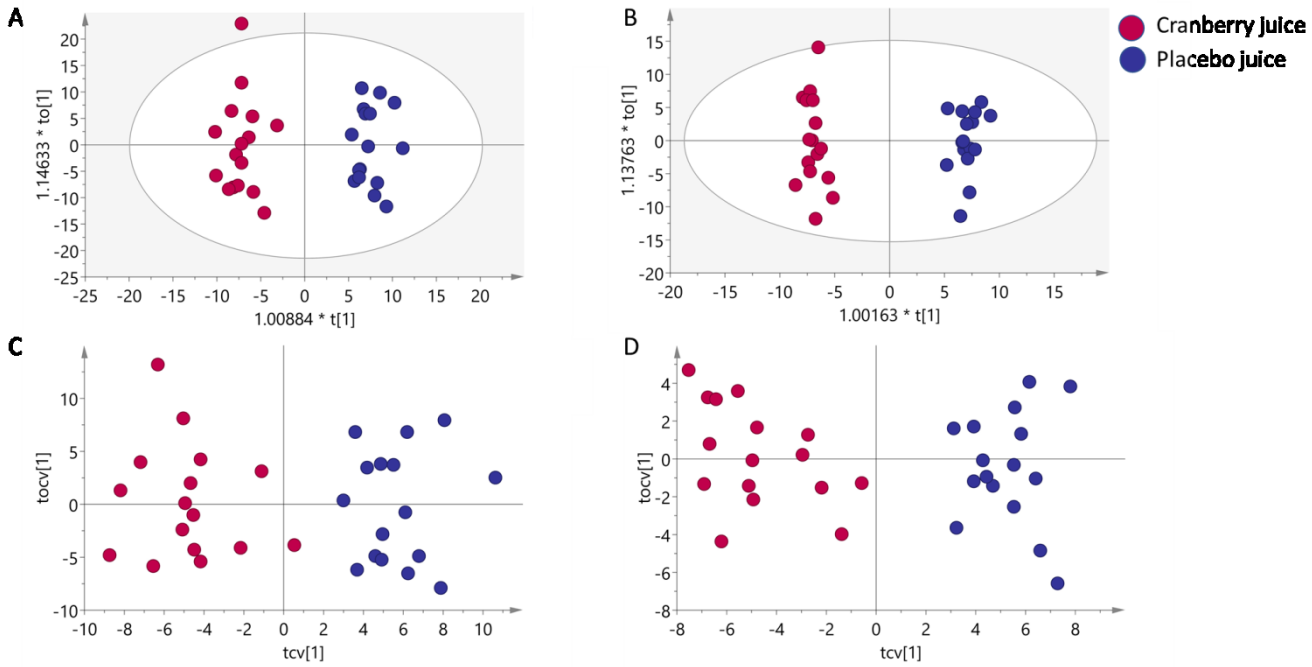

Figure 2. The score plots (A, B) and cross-validated score plots (C, D) acquired by positive ionization derived from OPLS-DA models of plasma after cranberry juice or placebo consumption. Panel A and C: Plasma after drinking juice for 3 days; Panel B and D: Plasma after drinking juice for 21 days. Red dots: plasma after drinking cranberry juice. Blue dots: plasma after drinking placebo juice.

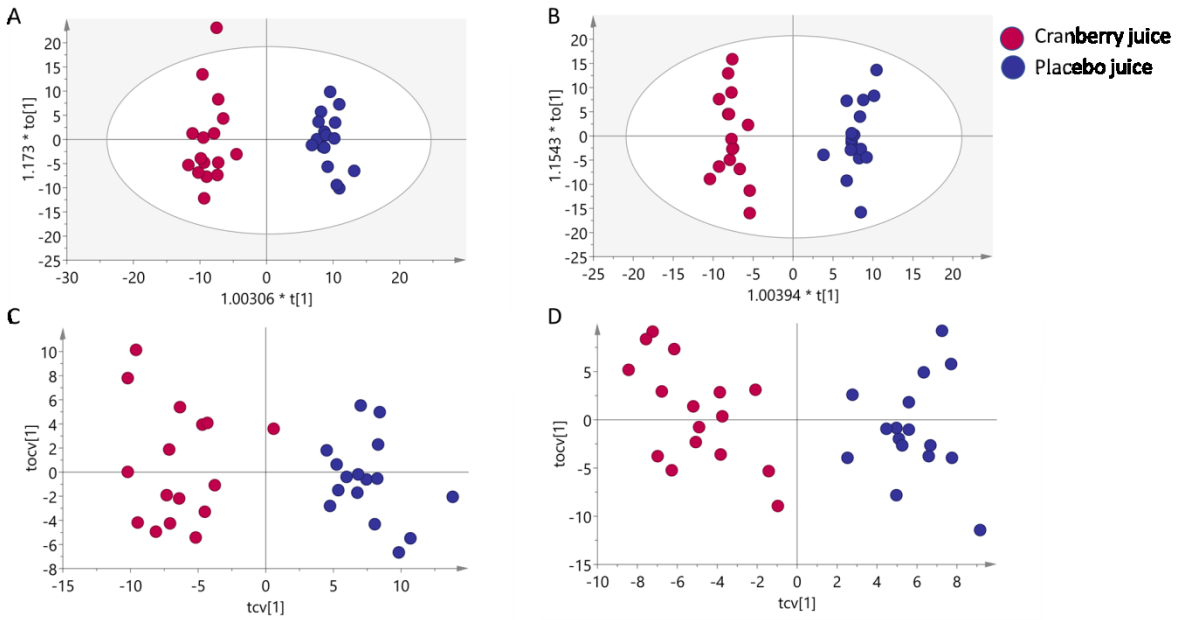

Figure 3. The score plots (A, B) and cross-validated score plots (C, D) acquired by negative ionization derived from OPLS-DA models of plasma after cranberry juice or placebo consumption. Panel A and C: Plasma after drinking juice for 3 days; Panel B and D: Plasma after drinking juice for 21 days. Red dots: plasma after drinking cranberry juice. Blue dots: plasma after drinking placebo juice.

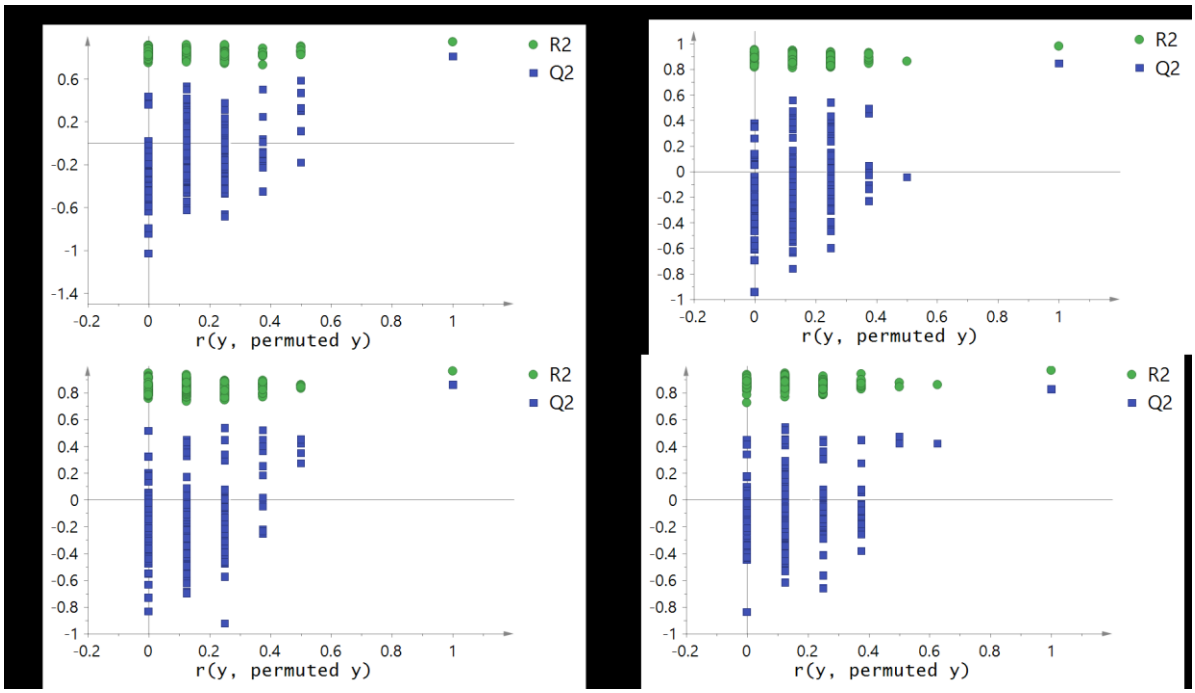

Figure 4. Validation plots obtained from 200 permutation tests for the OPLS-DA models with 1500 X variables detected in plasma after cranberry or placebo consumption. 3-day, 21-day consumption acquired by positive ionization (A, B); 3-day, 21-day consumption acquired by negative ionization (C, D).

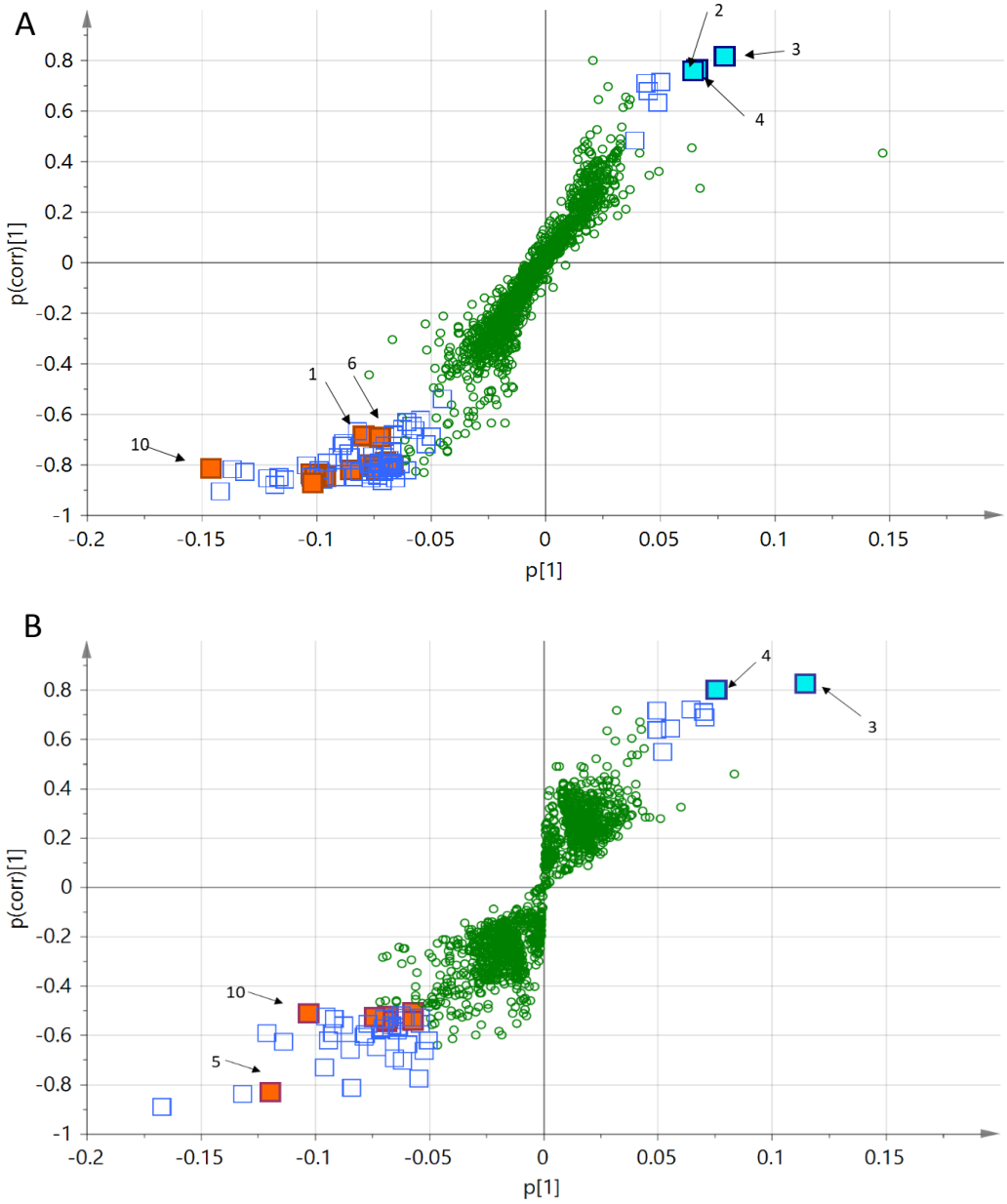

Figure 5. S-plots associated with the OPLS-DA score plot of data derived from UHPLC-HRMS of human plasma after cranberry juice or placebo consumption by positive ionization (A. after 3-day consumption; B. after 21-day consumption). p[1] is the loading vector of covariance in the first principal component. p(corr)[1] is loading vector of correlation in the first principal component. Variables with $|p| \geq 0.05$ and $|p(corr)| \geq 0.5$ are considered statistically significant. Statistically significant variables were hollow blue squares. Filled orange boxes were metabolites increased after cranberry juice consumption in previous study. Filled light blue boxes were metabolites increased after apple juice consumption in a previous study (numbers labeled are corresponded to the numbers of discriminant metabolites in Table 2).

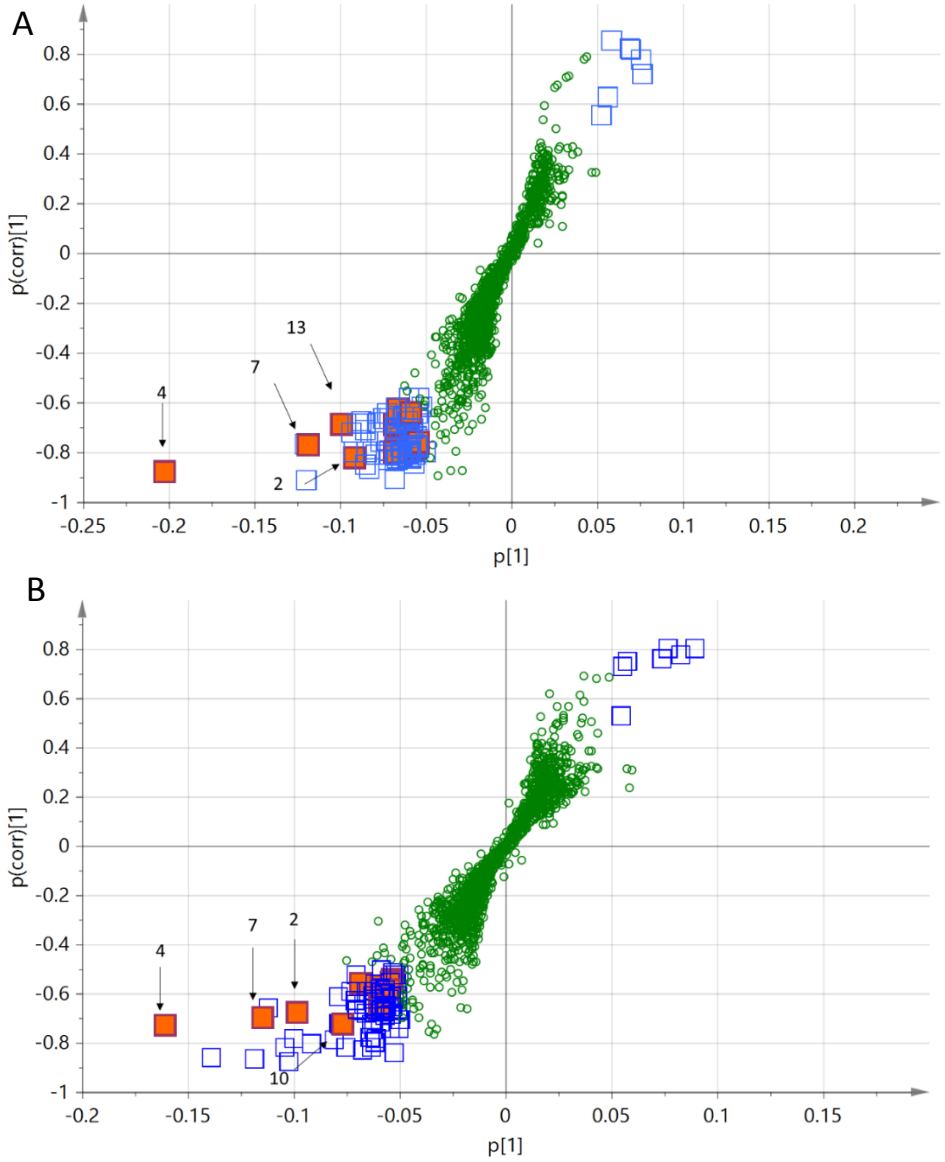

Figure 6. S-plots associated with the OPLS-DA score plot of data derived from UHPLC-HRMS of human plasma after cranberry juice or placebo consumption by negative ionization (A. after 3-day consumption; B. after 21-day consumption). p[1] is the loading vector of covariance in the first principal component. p(corr)[1] is loading vector of correlation in the first principal component. Variables with $|p| \geq 0.05$ and $|p(corr)| \geq 0.5$ are considered statistically significant. Statistically significant variables were hollow blue squares. Filled orange boxes were metabolites increased after cranberry juice consumption in a previous study (numbers labeled are corresponded to the numbers of discriminant metabolites in Table 3).

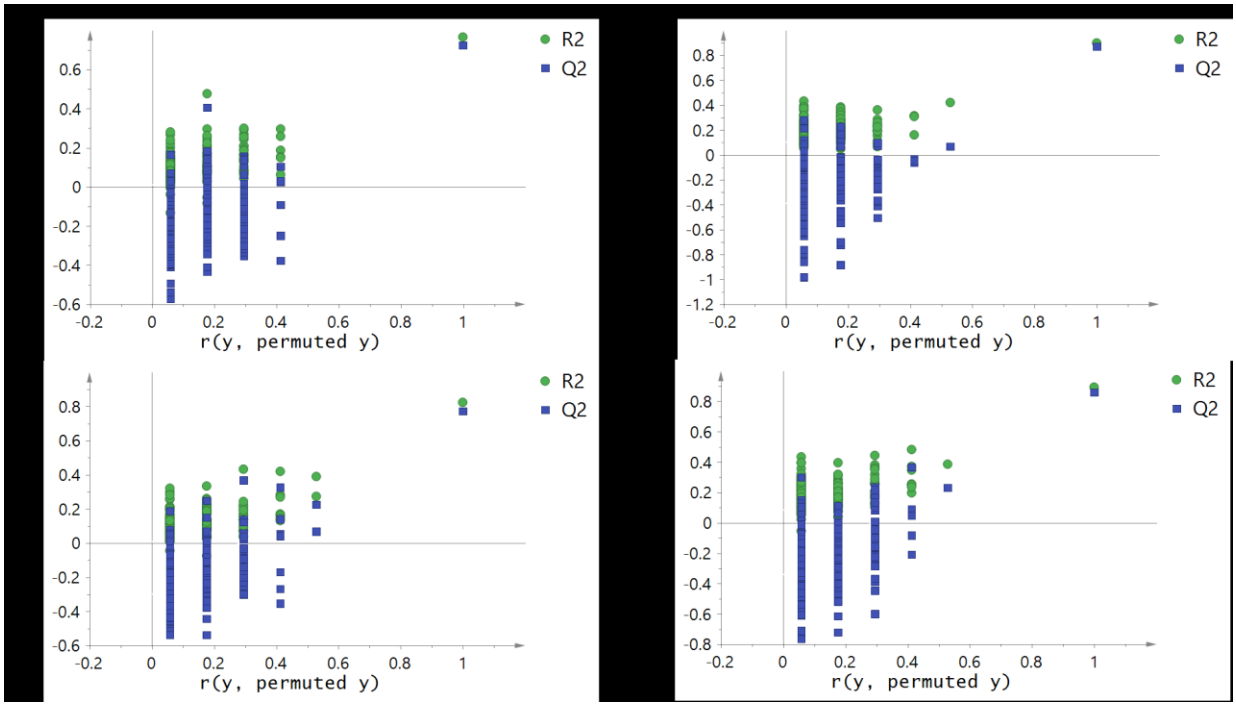

Figure 7. Validation plots obtained from 200 permutation tests for the predictive OPLS-DA models ***OPL PvC15+*** (A), ***OPL PvC18-*** (B), ***OPL BvC14+***(C), ***OPL BvC15-***(D).

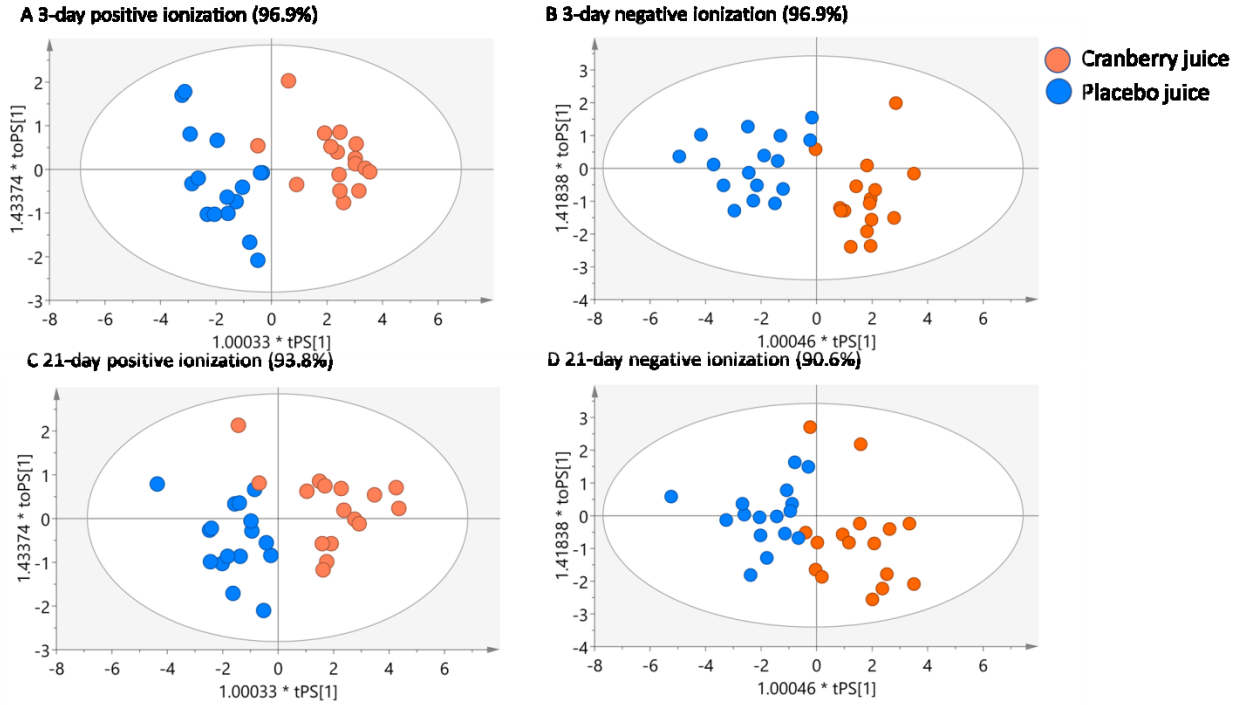

Figure 8. Classification scatters of cranberry juice and placebo consumer identification with plasma metabolites. According to the predictive model, plasma samples after drinking cranberry were classified to cranberry juice consumer (orange dots), plasma samples after drinking placebo were classified to apple juice consumer (light blue dots). In positive ionization mode, 15 discriminant metabolites listed in Table 2 were used to identify placebo and cranberry juice consumer after 3 days consumption with 96.9% correct rate (A); after 21 days consumption with 93.8% correct rate (C). In negative ionization mode, 18 discriminant metabolites listed in Table 3 were used to identify placebo and cranberry juice consumer after 3 days consumption with 96.9% correct rate (B); after 21 days consumption with 90.6% correct rate (D).

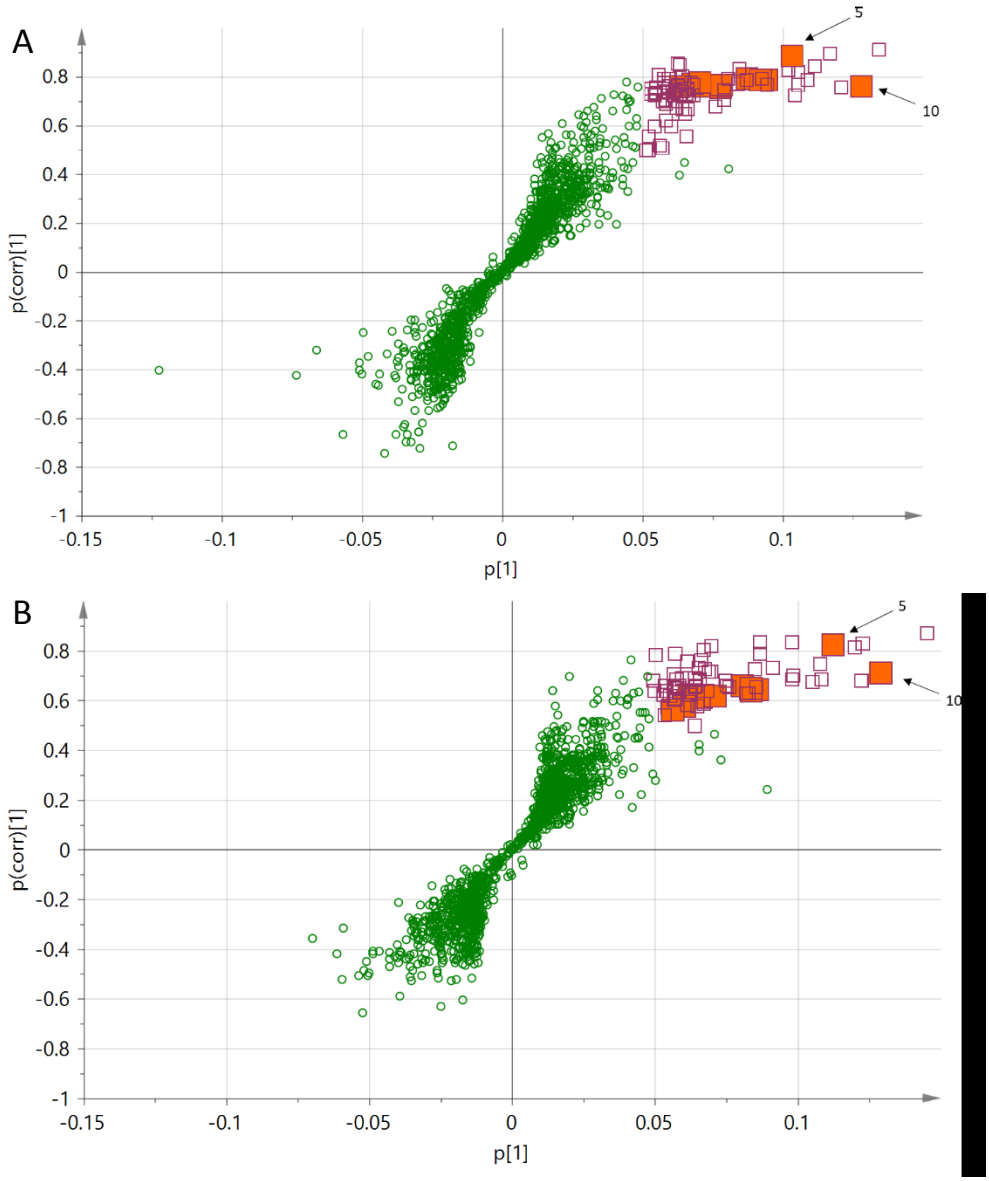

Figure 9. S-plots associated with the OPLS-DA score plot of data derived from UHPLC-HRMS of human plasma after cranberry juice consumption and baseline by **positive** ionization (A. after 3-day consumption; B. after 21-day consumption). p[1] is the loading vector of covariance in the first principal component. p(corr)[1] is loading vector of correlation in the first principal component. Variables with $|p| \geq 0.05$ and $|p(corr)| \geq 0.5$ are considered statistically significant. Statistically significant variables were hollow purple squares. Filled orange boxes were metabolites increased after cranberry juice consumption in a previous study (numbers labeled are corresponded to the numbers of discriminant metabolites in Table 2).

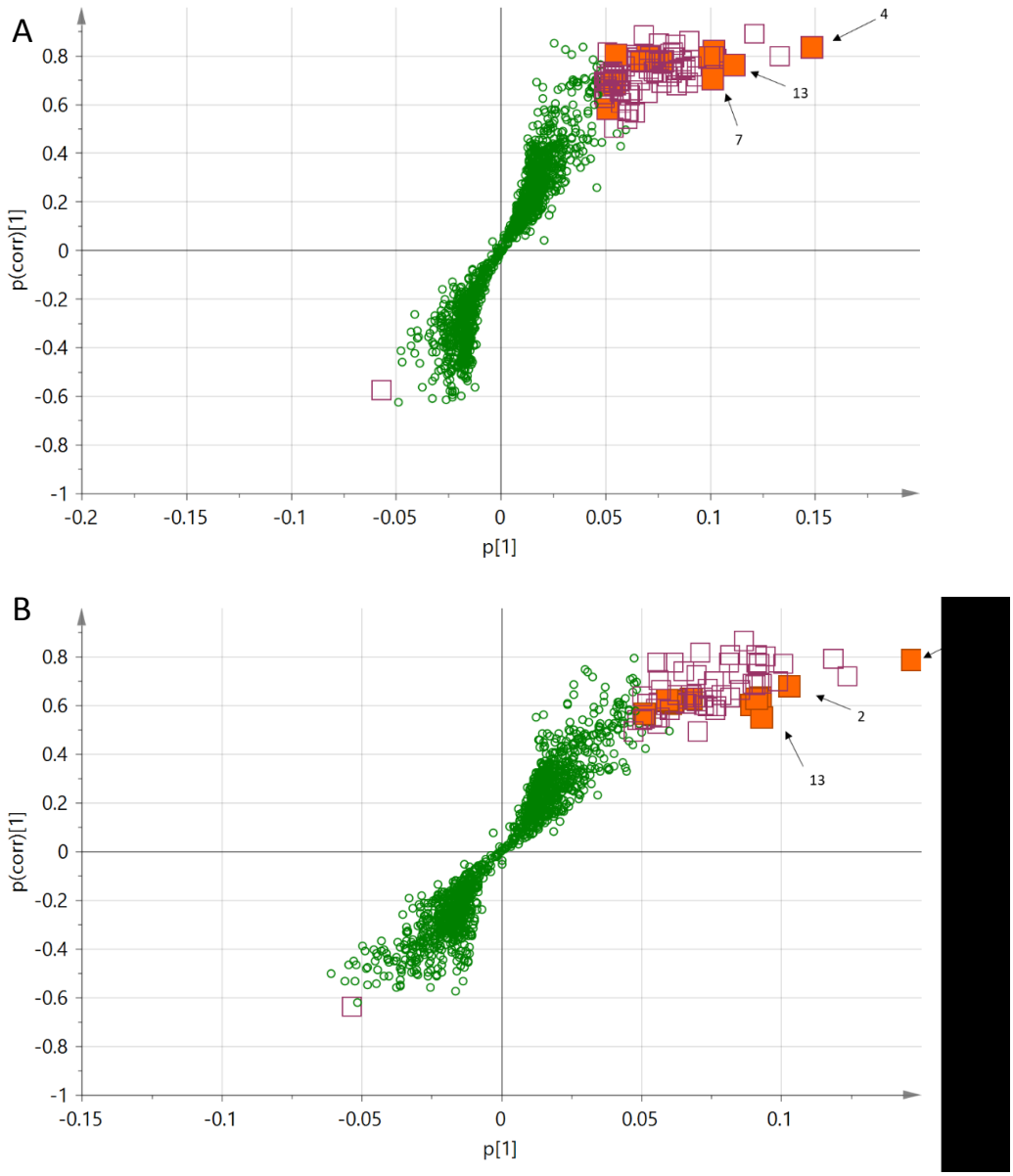

Figure 10. S-plots associated with the OPLS-DA score plot of data derived from UHPLC-HRMS of human plasma after cranberry juice consumption and baseline by **negative** ionization (A. after 3-day consumption; B. after 21-day consumption). p[1] is the loading vector of covariance in the first principal component. p(corr)[1] is loading vector of correlation in the first principal component. Variables with $|p| \geq 0.05$ and $|p(corr)| \geq 0.5$ are considered statistically significant. Statistically significant variables were hollow purple squares. Filled orange boxes were metabolites increased after cranberry juice consumption in previous study (numbers labeled are corresponded to the numbers of discriminant metabolites in Table 3).

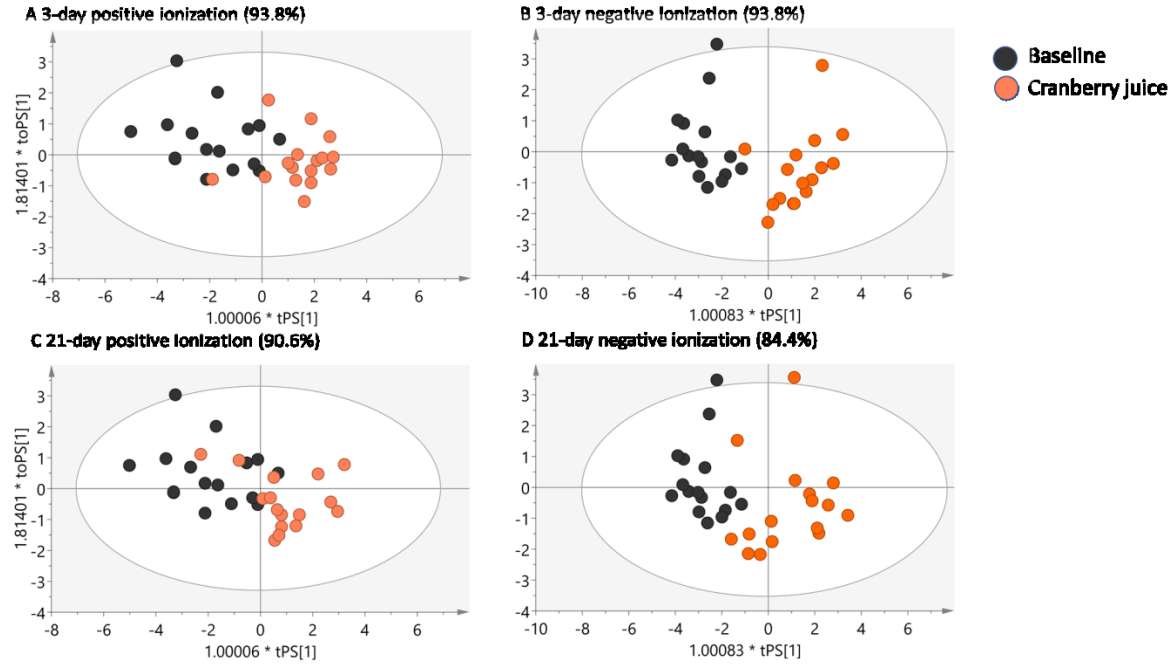

Figure 11. Classification scatters of cranberry juice consumer and non-consumer identification with plasma metabolites. According to the predictive model, plasma samples after drinking cranberry were classified to cranberry juice consumer (orange dots), plasma samples at baseline were classified to non-consumer (black dots). In positive ionization mode, 15 discriminant metabolites were used to identify cranberry juice consumer and non-consumer after 3 days consumption with 93.8% correct rate (A); after 21 days consumption with 90.6% correct rate (C). In negative ionization mode, 15 discriminant metabolites were used to identify cranberry juice consumer and non-consumer after 3 days consumption with 91.2% correct rate 93.8% (B); after 21 days consumption with 84.4% correct rate (D).

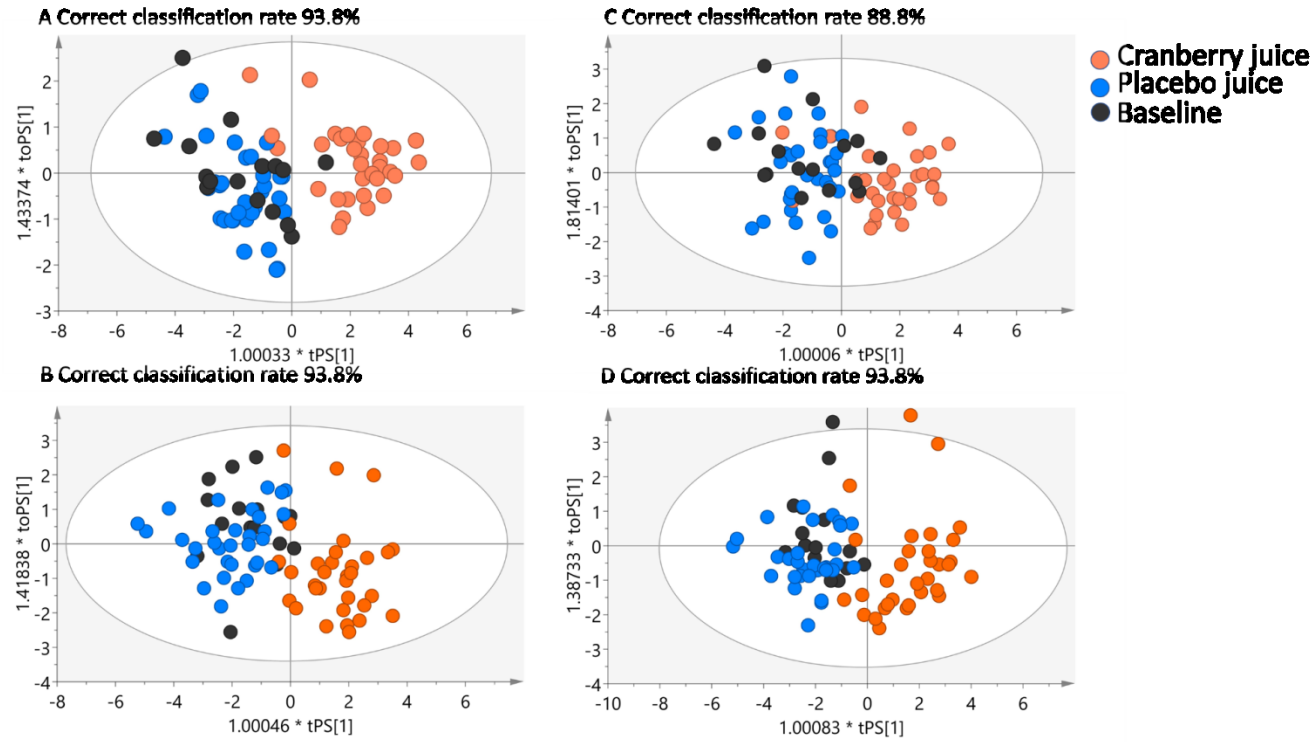

Figure 12. Classification scatters of all samples by ***Model PC***. Positive ionization (A); Negative ionization (B). Classification scatters of all samples by ***Model BC***. Positive ionization (C); Negative ionization (D).

This double-blinded study was aimed to build models with validated biomarkers to verify human consumption of cranberry juice. Discriminant metabolites from a previous 3-day open-label study were re-discovered and predictive OPLS-DA models were built to identify cranberry juice consumers and non-consumers. The models were able to identify cranberry juice consumers over placebo juice group with up to 96.9% correction rates.