JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Metabolite Profiling Approach Reveals the Interface of Primary and Secondary Metabolism in Colored Cauliflowers (*Brassica oleracea* L. ssp. *botrytis*)

Soo-Yun Park,[†] Sun-Hyung Lim,[†] Sun-Hwa Ha,[§] Yunsoo Yeo,[†] Woo Tae Park,[‡] Do Yeon Kwon,[‡] Sang Un Park,*[,‡] and Jae Kwang Kim*[,†]

[†]National Academy of Agricultural Science, Rural Development Administration, Suwon 441-707, Republic of Korea

[‡]Department of Crop Science, Chungnam National University, 220 Gung-dong, Yuseong-gu, Daejeon 305-764, Republic of Korea

[§]Department of Genetic Engineering and Crop Biotech Institute, Kyung Hee University, Yongin 446-701, Republic of Korea

ABSTRACT: In the present study, carotenoids, anthocyanins, and phenolic acids of cauliflowers (*Brassica oleracea* L. ssp. *botrytis*) with various colored florets (white, yellow, green, and purple) were characterized to determine their phytochemical diversity. Additionally, 48 metabolites comprising amino acids, organic acids, sugars, and sugar alcohols were identified using gas chromatography−time-of-flight mass spectrometry (GC-TOFMS). Carotenoid content was considerably higher in green cauliflower; anthocyanins were detected only in purple cauliflower. Phenolic acids were higher in both green and purple cauliflower. Results of partial least-squares discriminant, Pearson correlation, and hierarchical clustering analyses showed that green cauliflower is distinct on the basis of the high levels of amino acids and clusters derived from common or closely related biochemical pathways. These results suggest that GC-TOFMS-based metabolite profiling, combined with chemometrics, is a useful tool for determining phenotypic variation and identifying metabolic networks connecting primary and secondary metabolism.

KEYWORDS: *carotenoids, cauliflower, metabolomics, partial least-squares discriminant analysis, phenolic acids*

INTRODUCTION

In recent years, considerable attention has been paid to edible plants rich in bioactive compounds possessing antioxidant activity and other beneficial physicochemical properties. Brassicaceae vegetables represent a large part of the human diet and are recognized as a rich source of nutrients and health-promoting phytochemicals including vitamins, minerals, carbohydrates, amino acids, carotenoids, tocopherols, glucosinolates, and phenolic compounds.[1] It has been demonstrated that a higher consumption of *Brassica* vegetables reduces the risk of age-related chronic diseases and several types of cancer including lung, prostate, and breast.[2−4]

The main vegetable species of *Brassica*, *Brassica oleracea*, includes kale, cabbage, broccoli, cauliflower, and others. Although the exact mechanism of the chemopreventive action of cruciferous vegetables is unclear, many experimental and clinical studies suggest that the substances present in these vegetables inhibit tumor cell growth and reduce oxidative stress. Johnson showed the potential of glucosinolate breakdown products derived from cruciferous vegetables in protecting against human gastrointestinal and lung cancers,[5] and Ayaz et al. reported the phenolic acid content and antioxidant activity of kale.[6] Several studies have determined total phenol and anthocyanin levels in cauliflower, broccoli, and cabbage and their antioxidant capacities.[7−10] For these reasons, the popularity and consumption of cruciferous vegetables are increasing, and research to identify valuable compounds present in these vegetables is necessary in the food and breeding industries.

Qualitative variations in the nutrients and phytochemical profiles of *Brassica* varieties could contribute to differences in health-promoting properties. Therefore, this study aimed to identify the core primary metabolites and bioactive secondary metabolites from cauliflower (*B. oleracea* L. ssp. *botrytis*) of four varieties with white, yellow, green, and purple florets. The primary metabolite profile is closely related to phenotype and includes important nutritional characteristics.[11,12] Gas chromatography−time-of-flight mass spectrometry (GC-TOFMS)-based metabolic profiling analysis is useful for the rapid and highly sensitive detection of plant metabolites from the central pathways of primary metabolism.[13] In this study, hydrophilic primary metabolites and phenolic acids were profiled using GC-TOFMS, and carotenoids and anthocyanin were quantified using high-performance liquid chromatography (HPLC) from four colored cauliflowers. Data obtained were then subjected to multivariate statistical analyses, partial least-squares discriminant analysis (PLS-DA), Pearson correlation analysis, and hierarchical clustering analysis (HCA) to determine phenotypic variation and relationships between metabolite contents.

MATERIALS AND METHODS

Samples and Chemicals. Seeds of cauliflower (*B. oleracea* L. var. *botrytis*) cv. 'White Sails' (white), 'Cheddar' (yellow), 'Orbit' (green, romanesco), and 'Graffiti' (purple) were purchased from Stokes Seeds Ltd., St. Catharines, Canada, and stored at 4 °C. Four cauliflower cultivar seeds were germinated in a greenhouse, and seedlings were transferred to the experimental farm at Chungnam National University (Daejeon, Korea). After 10 weeks, the cauliflowers were harvested (Figure 1).

Received: March 25, 2013
Revised: June 4, 2013
Accepted: June 20, 2013
Published: June 20, 2013

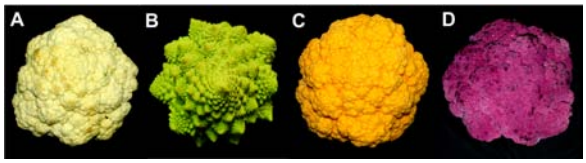

Figure 1. Images of the four cauliflower varieties (*Brassica oleracea* L. var. *botrytis*): (A) cv. White Sails; (B) cv. Orbit (romanesco); (C) cv. Cheddar; (D) cv. Graffiti.

The samples were then freeze-dried at −80 °C for at least 72 h and ground into a fine powder using a mortar and pestle for phytochemical analysis. All chemicals used in this study were of analytical grade. Methanol and chloroform, used as extraction solvents, were purchased from J. T. Baker (Phillipsburg, NJ, USA). Ribitol, sinapic acid, vanillic acid, and *N*-methyl-*N*-trimethylsilyltrifluoroacetamide (MSTFA) were obtained from Sigma Chemical Co. (St. Louis, MO, USA). Methoxyamine hydrochloride was purchased from Thermo Fisher Scientific (Pittsburgh, PA, USA). Syringic, 3,4,5-trimethoxycinnamic, *p*-hydroxybenzoic, ferulic, and acetic acids were acquired from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). *p*-Coumaric acid was obtained from MP Biomedicals (Solon, OH, USA). Pyridine and *N*-(*tert*-butyldimethylsilyl)-*N*-methyltrifluoroacetamide (MTBSTFA) with 1% *tert*-butyldimethylchlorosilane (TBDMCS) were purchased from Thermo Fisher Scientific.

Carotenoid Extraction and Analysis. Carotenoids were extracted and measured using HPLC, as described previously.[14] Briefly, the carotenoids were extracted from cauliflower samples (0.1 g) by adding 3 mL of ethanol containing 0.1% ascorbic acid (w/v), vortex mixing for 20 s, and placing in a water bath at 85 °C for 5 min. The carotenoid extract was saponified with potassium hydroxide (120 μL, 80% w/v) in a water bath at 85 °C for 10 min. After saponification, the samples were immediately placed on ice, and cold deionized water (1.5 mL) was added. β-Apo-8′-carotenal (0.2 mL, 25 μg/mL) was added as an internal standard (IS). To separate the layers, carotenoids were extracted twice with hexane (1.5 mL) by centrifugation at 1200*g*. Aliquots of the extracts were dried under a stream of nitrogen and redissolved in 50:50 (v/v) dichloromethane/methanol before HPLC analysis. The carotenoids were then separated in a C30 YMC column (250 × 4.6 mm, 3 μm; YMC Co., Kyoto, Japan) by an Agilent 1100 HPLC instrument (Massy, France) equipped with a photodiode array detector. Chromatograms were generated at 450 nm. Solvent A consisted of methanol/water (92:8 v/v) with 10 mM ammonium acetate; solvent B consisted of 100% methyl *tert*-butyl ether. Gradient elution was performed at 1 mL/min under the following conditions: 0 min, 90% A/10% B; 20 min, 83% A/17% B; 29 min, 75% A/25% B; 35 min, 30% A/70% B; 40 min, 30% A/70% B; 42 min, 25% A/75% B; 45 min, 90% A/10% B; and 55 min, 90% A/10% B. Carotenoid standards were purchased from CaroteNature (Lupsingen, Switzerland). Calibration curves were drawn for quantification by plotting four concentrations of the carotenoid standards according to the peak area ratios with β-apo-8′-carotenal.

Anthocyanin Extraction and Analysis. Anthocyanin extraction was performed according to a slightly modified method of Kim et al.[15] A total of 20 mg of sample was extracted with 0.95 mL of 85% methanol acidified with 1.0 N HCl followed by sonication for 1 min. Extracts were then incubated at 38 °C for 5 min with a mixing frequency of 500 rpm using a Thermomixer comfort. The crude suspension was centrifuged at 10000*g* at 4 °C for 5 min, and the crude extract was hydrolyzed with 0.2 mL 6 N HCl at 90 °C for 2 h. After centrifugation, the hydrolysate was passed through a 0.22 μm PTFE syringe filter prior to HPLC analysis. Anthocyanin was separated in a C_{18} column (250 × 4.6 mm, 5 μm, Inertsil ODS-3, GL Sciences, Tokyo, Japan) as described above. Elution was performed using a binary gradient of 0.1% formic acid in water (mobile phase A) and 0.1% formic acid in acetonitrile (mobile phase B) according to the following program: 0 min, 95% A/5% B; 20 min, 75% A/25% B; 22 min, 0% A/100% B; 32 min, 0% A/100% B; 34 min, 95% A/5% B; and 44 min, 95% A/5% B. The flow rate was 1.0 mL/min, and the column temperature was 40 °C. The UV−vis detector wavelength was set to 520 nm. Cyanidin was purchased from Extrasynthese (Genay, France). Calibration curves were drawn for quantification by plotting five concentrations of the cyanidin standard.

Phenolic Acid Extraction and Analysis. Soluble (free and esterified forms) and insoluble (bound form) phenolic acids were extracted according to the procedure described by Park et al.[16] with slight modifications. The powdered samples (0.1 g) were extracted twice by water-based sonication for 5 min at room temperature and incubation at 30 °C for 10 min with 1 mL of 85% methanol containing 2 g/L butylated hydroxyanisole (BHA). After centrifugation at 13000 rpm for 10 min at 4 °C, the combined extracts and residue were analyzed to determine the soluble and insoluble phenolic acids, respectively. Fifty microliters of 3,4,5-trimethoxycinnamic acid (100 μg/mL) was added as an IS, and the mixture was hydrolyzed with 1 mL of 5 N NaOH at 30 °C under nitrogen gas for 4 h. Each hydrolyzed sample was adjusted to pH 1.5−2.0 with 6 M HCl, extracted with ethyl acetate, and evaporated in a centrifugal concentrator (Eyela, Tokyo, Japan). For derivatization, 40 μL of MTBSTFA containing 1% TBDMCS and 40 μL of pyridine were added to the dried extracts followed by incubation at 60 °C for 30 min at a mixing frequency of 1200 rpm using a thermomixer comfort (Eppendorf AG, Hamburg, Germany). Each derivatized sample (1 μL) was injected into a 7890A gas chromatograph (Agilent, Atlanta, GA, USA) using a 7683B autosampler (Agilent) with a split ratio of 10 and separated on a 30 m × 0.25 mm i.d. fused silica capillary column coated with 0.25 μm CP-SIL 8 CB low bleed (Varian Inc., Palo Alto, CA, USA). The injector temperature was 230 °C, and the flow rate of helium gas through the column was 1.0 mL/min. The temperature program was set to 150 °C and maintained for 2 min followed by a 15 °C/min oven temperature ramp to 320 °C and a 10 min hold. The column effluent was introduced into a Pegasus HT TOF mass spectrometer (LECO, St. Joseph, MI, USA). The transfer line and ion source temperatures were 250 and 200 °C, respectively. The detected mass range was m/z 85−700, and the detector voltage was set to 1700 V. For quantification purposes, a standard stock solution of six phenolics (ferulic, *p*-coumaric, *p*-hydroxybenzoic, sinapic, syringic, and vanillic acids) and 3,4,5-trimethoxycinnamic acid (used as an IS) was prepared in methanol (100 μg/mL). Calibration samples, ranging from 0.01 to 10.0 μg, were prepared by mixing individual stock solutions of the six phenolic acid standards.

Metabolite Profiling. Metabolite extraction was performed as described previously.[17] A total of 20 mg of ground sample was extracted with 1 mL of a mixed solvent of methanol/water/chloroform (2.5:1:1 by volume). Ribitol (60 μL, 0.2 mg/mL) and 5α-cholestane (50 μL, 0.01 mg/mL) were added as an IS. The extraction was performed at 37 °C with a mixing frequency of 1200 rpm for 30 min using a thermomixer compact. The solutions were then centrifuged at 16000*g* for 3 min. The polar phase (0.8 mL) was transferred into a new tube, and 0.4 mL of water was added. Mixed contents of the tube were centrifuged at

Table 1. Contents (Micromoles per Gram on a Dry Weight Basis) of Carotenoids in Four Cauliflower Varieties (*Brassica oleracea* var. *botrytis*)[a]

color[b]	lutein	α-carotene	(all-*E*)-β-carotene	(9*Z*)-β-carotene	(13*Z*)-β-carotene	total
white	0.50 ± 0.06b		0.22 ± 0.02c	0.05 ± 0.00b		0.97 ± 0.09
green	43.14 ± 3.12a	2.49 ± 0.03a	20.94 ± 1.57a	1.77 ± 0.17a	3.52 ± 0.13a	71.86 ± 5.01
yellow	1.35 ± 0.11b	2.25 ± 0.05b	4.61 ± 1.20b	1.59 ± 0.19a	3.00 ± 0.23b	12.80 ± 1.56
purple	0.91 ± 0.04b		0.35 ± 0.03c	0.08 ± 0.01b	0.24 ± 0.02c	1.58 ± 0.03

[a]Different letters represent significant ($p < 0.05$) differences between means according to ANOVA combined with Duncan's multiple-range test. Each value represents the mean ± standard deviation ($n = 3$). [b]White, White Sails; green, Orbit; yellow, Cheddar; purple, Graffiti.

16000*g* for 3 min. The methanol/water phase was dried in a centrifugal concentrator for 2 h followed by a drying process in a freeze-dryer for 16 h. Methoxime (MO) derivatization was performed by adding 160 μL of methoxyamine hydrochloride (20 mg/mL) in pyridine and shaking at 30 °C for 90 min. Trimethylsilyl (TMS) etherification was performed by the addition of 160 μL of MSTFA at 37 °C for 30 min. A derivatized sample (1 μL) was separated on the same column as described above. The split ratio was set to 1:25 with an injector temperature of 230 °C and a helium gas flow rate of 1.0 mL/min. The temperature program was set as follows: initial temperature of 80 °C for 2 min followed by an increase to 320 °C at 15 °C/min and a 10 min hold at 320 °C. The transfer line and ion source temperatures were 250 and 200 °C, respectively. The scanned mass range was m/z 85−600, and the detector voltage was set to 1700 V.

Statistical Analysis. All analyses were performed at least three times. Experimental data were analyzed by analysis of variance (ANOVA), and significant differences among the means were determined by Duncan's multiple-range test (SAS 9.2; SAS Institute, Cary, NC, USA). In cauliflower, 52 metabolites were identified by GC-TOFMS and HPLC. The data obtained were subjected to PLS-DA (SIMCA-P version 12.0; Umetrics, Umeå, Sweden) to evaluate differences among groups of multivariate data. The PLS-DA output consisted of score plots to visualize the contrast between samples and loading plots to explain the cluster separation. The data file was scaled with unit variance scaling before all variables were subjected to the PLS-DA. Pearson correlation analysis was performed with the SAS 9.2 software package (SAS Institute). Correlation analysis was performed on the relative levels of 52 metabolites with standardization preprocessing. HCA and heatmap visualization of the correlation coefficient were performed using MultiExperiment Viewer version 4.4.0 software (http://www.tm4.org/mev/).

RESULTS AND DISCUSSION

Carotenoid, Anthocyanin, and Phenolic Acid Compositions of Colored Cauliflowers. Color variation in plants could be caused by the quantity and composition of diverse pigment molecules such as chlorophylls, carotenoids, and anthocyanins and may contribute to the health promotion ability of plants due to their antioxidant activity. Thus, HPLC analyses were conducted to investigate carotenoid and anthocyanin contents of various colored cauliflowers. In addition, phenolic acids, well-known antioxidant compounds, were identified using GC-TOFMS. The levels of these compounds in various cauliflowers with florets of different colors have not been evaluated previously. In this study, three carotenoids, lutein and α- and β-carotene, were identified in cauliflower florets. Quantitative results for the compounds from cauliflowers of four varieties (white, yellow, green, and purple) are shown in Table 1. The most abundant carotenoid in all floret extracts was lutein, which comprised >50% of the total carotenoid content, with the exception of yellow cauliflower, wherein β-carotene was the most plentiful. α-Carotene was not detected in either white or purple cauliflower. The total carotenoid level was highest in green cauliflower (71.86 μmol/g) and lowest in white cauliflower (0.97 μmol/g).

Anthocyanins are glycosylated anthocyanidins. Hundreds of naturally occurring anthocyanins have been reported due to the diversity of glycosylation. However, complex glycoslyation patterns make the identification of individual anthocyanins difficult, even with LC-MS analysis.[8] For this reason, anthocyanins in cauliflowers were investigated by analyzing their hydrolysis products, anthocyanidins. Anthocyanin was detected only in purple cauliflower, and the hydrolysis product was nearly exclusively cyanidin, at 161.8 ± 3.5 mg/100 g. This level is higher than the 63.2 mg/100 g level of cyanidin in purple cauliflower, as determined by Li et al.,[8] which could be caused by differences in environmental conditions during cultivation or storage of the respective samples.[18]

Phenolic acid levels in methanol-soluble and -insoluble phenolic fractions of cauliflower florets were also determined. All varieties

Table 2. Contents (Micromoles per Gram on a Dry Weight Basis) of Soluble and Insoluble Phenolic Acids in Four Cauliflower Varieties (*Brassica oleracea* var. *botrytis*)[a]

	soluble				insoluble				total			
phenolic acid[b]	white[c]	green	yellow	purple	white	green	yellow	purple	white	green	yellow	purple
p-OH	62.60 ± 0.35c	46.07 ± 1.83d	73.41 ± 4.61b	259.59 ± 2.95a	12.77 ± 0.66c	14.68 ± 1.65bc	17.95 ± 0.37b	85.33 ± 3.00a	75.36 ± 0.51	60.74 ± 2.65	91.36 ± 4.45	344.93 ± 4.94
Van	47.96 ± 0.11b	47.97 ± 0.12b	46.81 ± 0.15c	52.64 ± 0.11a	46.02 ± 0.11b	46.18 ± 0.10b	45.67 ± 0.08b	52.71 ± 0.78a	93.98 ± 0.03	94.16 ± 0.23	92.48 ± 0.18	105.35 ± 0.77
Syr	1.22 ± 0.05c	2.17 ± 0.12b	3.05 ± 0.16a	0.67 ± 0.07d	1.77 ± 0.09b	1.17 ± 0.04c	2.22 ± 0.04a	1.21 ± 0.16c	2.99 ± 0.13	3.35 ± 0.14	5.27 ± 0.13	1.88 ± 0.23
Cou	828.71 ± 78.28c	2135.51 ± 186.81b	876.11 ± 225.45c	3485.85 ± 345.02a	1060.30 ± 95.83c	1458.13 ± 131.44b	1096.83 ± 115.74c	4413.61 ± 244.85a	1889.01 ± 49.70	3593.64 ± 162.31	1972.94 ± 222.40	7899.46 ± 132.91
Fer	108.61 ± 4.75c	699.84 ± 41.54b	73.87 ± 2.65c	919.09 ± 12.87a	82.80 ± 3.00c	150.31 ± 3.33b	76.93 ± 4.00c	1430.61 ± 63.24a	191.41 ± 6.65	850.15 ± 43.55	150.80 ± 6.57	2349.70 ± 50.55
Sin	97.19 ± 14.88c	3319.25 ± 106.47a	84.78 ± 3.02c	262.03 ± 12.31b	80.72 ± 0.51c	735.43 ± 43.41a	94.33 ± 2.92c	168.33 ± 3.75b	177.91 ± 15.26	4054.68 ± 78.26	179.11 ± 5.90	430.37 ± 12.50
sum	1146.29 ± 67.79	6250.82 ± 254.69	1158.03 ± 220.01	4979.88 ± 364.45	1284.37 ± 98.74	2405.89 ± 178.03	1333.92 ± 122.05	6151.81 ± 297.08	2430.66 ± 67.19	8656.71 ± 192.34	2491.96 ± 223.08	11131.69 ± 93.26

[a]Different letters represent significant ($p < 0.05$) differences between means according to ANOVA combined with Duncan's multiple-range test. Each value represents the mean ± standard deviation ($n = 3$). [b]*p*-OH, *p*-hydroxybenzoic acid; Van, vanillic acid; Syr, syringic acid; Cou, *p*-coumaric acid; Fer, ferulic acid; Sin, sinapic acid. [c]White, White Sails; green, Orbit; yellow, Cheddar; purple, Graffiti.

dx.doi.org/10.1021/jf401330e | J. Agric. Food Chem. 2013, 61, 6999−7007

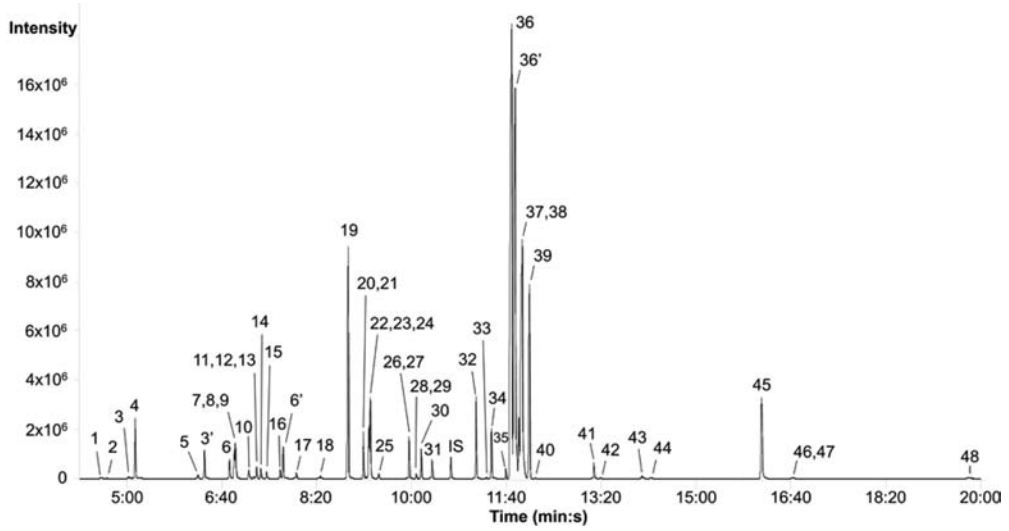

Figure 2. Selected ion chromatograms of hydrophilic metabolites extracted from purple cauliflower (cv. Graffiti) as MO/TMS derivatives separated on a 30 m × 0.25 mm i.d. fused silica capillary column coated with 0.25 μm CP-SIL 8 CB low bleed. Peak identification: 1, pyruvic acid; 2, lactic acid; 3, valine; 4, alanine; 5, glycolic acid; 3′, valine; 6, serine; 7, ethanolamine; 8, glycerol; 9, leucine; 10, isoleucine; 11, proline; 12, nicotinic acid; 13, glycine; 14, succinic acid; 15, glyceric acid; 16, fumaric acid; 6′, serine; 17, threonine; 18, β-alanine; 19, malic acid; 20, salicylic acid; 21, aspartic acid; 22, methionine; 23, pyroglutamic acid; 24, 4-aminobutyric acid; 25, threonic acid; 26, arginine; 27, glutamic acid; 28, phenylalanine; 29, *p*-hydroxybenzoic acid; 30, xylose; 31, asparagine; 32, glutamine; 33, shikimic acid; 34, citric acid; 35, quinic acid; 36, fructose; 36′, fructose; 37, galactose; 38, glucose; 39, mannose; 40, mannitol; 41, inositol; 42, ferulic acid; 43, tryptophan; 44, sinapic acid; 45, sucrose; 46, cellobiose; 47, trehalose; 48, raffinose; IS, internal standard (ribitol).

investigated contained the six phenolic acids (*p*-hydroxybenzoic, vanillic, syringic, *p*-coumaric, ferulic, and sinapic acid), as shown in Table 2. The major portion of phenolic acids was composed of *p*-coumaric, ferulic, and sinapic acids, which accounted for >90% of the total content in all varieties. The total levels of six phenolic acids in green and purple cauliflower were >4-fold higher compared to those in white and yellow cauliflower. A noticeable difference between green and purple cauliflower was observed in the proportions of the three dominant acids. The predominant phenolic acid in purple cauliflower was *p*-coumaric acid (71%), followed by ferulic acid (21%) and sinapic acid (4%), whereas sinapic acid represented the greatest proportion (47%) in green cauliflower, followed by *p*-coumaric acid (42%) and ferulic acid (10%). The compositions of these phenolic acids were in accordance with the results of previous studies. Sikora et al. reported that green cauliflower contained sinapic acid as the primary phenolic acid.[19] Lo Scalzo et al. showed that the *p*-coumaryl and feruloyl esterified forms of cyanidin-3-sophoroside-5-glucoside were predominant in purple cauliflower.[9]

Classification of Cauliflower Varieties by GC-TOFMS-Based Metabolic Profiling. GC-MS-based metabolite profiling was conducted to assess the variation in polar primary metabolites in the various colored cauliflowers. Untargeted metabolomic approaches by GC-MS have been successfully applied to explore metabolic discrimination between various genotypes[14] and to identify associations between phenotypic traits and primary metabolites.[13] In this study, low molecular weight molecules from cauliflower florets were identified by GC-TOFMS. ChromaTOF software was used to assist with peak location; peaks were identified by comparison with reference compounds and use of an in-house library.[17] In addition, the identification of several metabolites was performed using direct comparison of the sample mass chromatogram with those of commercially available standard compounds, which were obtained by a similar MO/TMS derivatization and GC-TOFMS analysis. In total, 48 metabolites including 19 amino acids, 16 organic acids, 9 sugars, 3 sugar alcohols, and 1 amine were detected in cauliflower florets (Figure 2). Quantification was performed using selected ions, as described in Table 3. Quantitative calculations of all analytes were based on peak area ratios relative to that of the IS.

Principal component analysis (PCA) and PLS-DA are the most common chemometric tools for extracting and rationalizing information from any multivariate description of a biological system.[20] Especially, PLS-DA is preferred over PCA for sample discrimination because the dimension reduction provided by PLS is guided explicitly by among-groups variability, here cauliflower varieties, whereas PCA is capable only of identifying gross variability directions and is not capable of distinguishing "among-groups" and "within-groups" variabilities.[21,22] PLS-DA has been applied to identify metabolic differences among varieties in perilla and millet seeds.[12,23] Thus, the quantitative data for these 48 metabolites and 4 pigment compounds identified by HPLC analysis (cyanidin, lutein, and *all-trans*- and 9-*cis*-β-carotene) were subjected to PLS-DA to identify differences in metabolite profiles between varieties (Figure 3). The PLS-DA results demonstrated the absence of significant variance within the same variety. The two highest ranking components accounted for 84.3% of the total variance within the data set. The first components, accounting for 59.5% of the total variation, resolved the measured composition profiles of green cauliflower from the other three varieties. This variation was mainly attributable to amino acids, organic acids, and carotenoids, of which the corresponding loading was negative for all amino acids and carotenoids and positive for all organic acids, with the exceptions of citric and sinapic acid. The loadings plot indicated that amino acids and carotenoids were higher in green cauliflower than in white, yellow, and purple cauliflower, and most organic acids were lower in green cauliflower than in others. In addition, purple cauliflower was separated from yellow cauliflower by the second components, accounting for 24.8% of the total variation. The corresponding loading was positive for all phenolic acids and negative for all carotenoids and sugars with the exception of

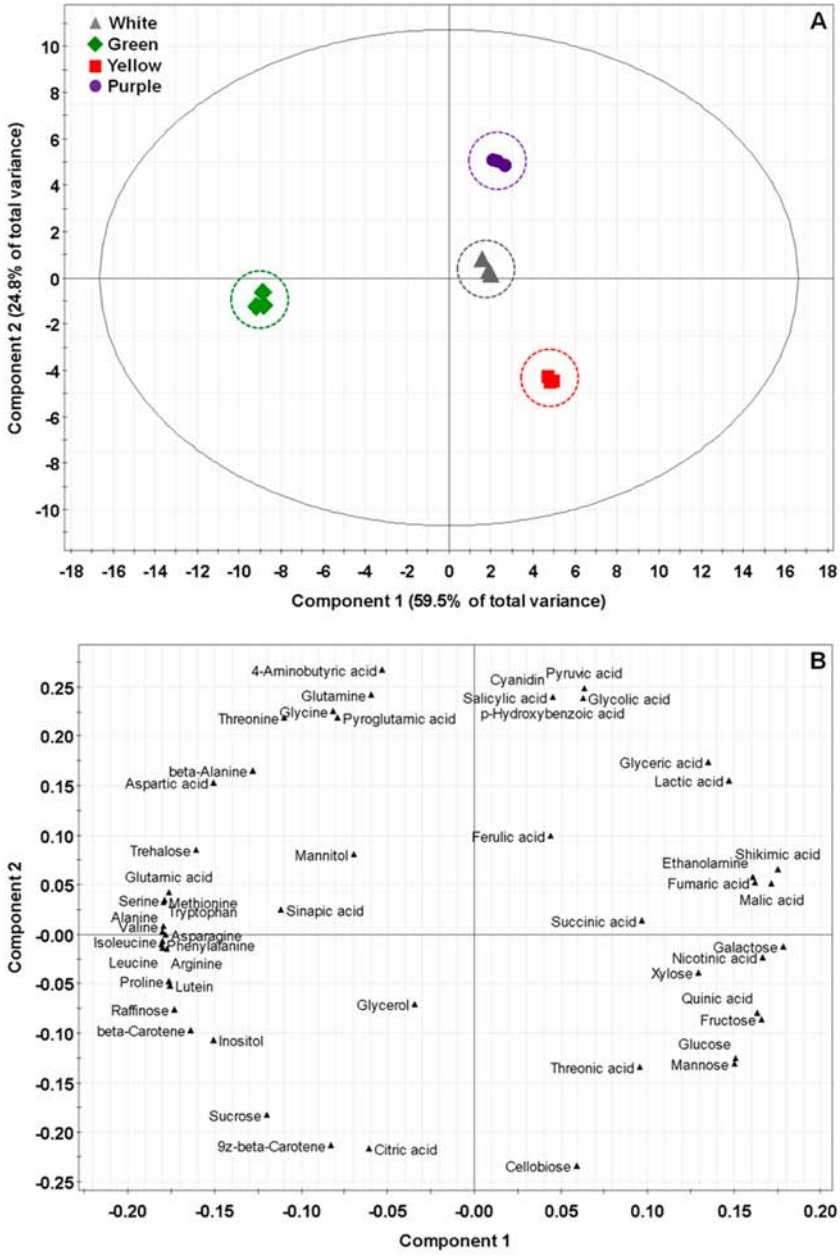

Figure 3. Score (A) and loading (B) plots of principal components 1 and 2 of the PLS-DA results obtained from data on 52 metabolites of four cauliflower varieties.

trehalose, indicating that purple cauliflower had higher phenolic acids and lower sugars and carotenoids compared to yellow cauliflower.

To gain more insight into the metabolic differences between varieties, the variables important in the projection (VIP) scores were examined. VIP is a weighted sum of squares of the PLS weight, and a value >1 is generally used as a criterion to identify the variables important to the model.[17] Among 52 metabolites, the variable that played the greatest role in discriminating between metabolic profiles of colored cauliflowers was 4-aminobutyric acid, followed by pyruvic acid and glutamine (Figure 4). In plant biological systems, the building blocks for secondary metabolites are derived from primary metabolism, and the most important building blocks employed in the biosynthesis of secondary metabolites are derived from the intermediates acetyl coenzyme A (acetyl-CoA), shikimic acid, mevalonic acid, and 1-deoxyxylulose 5-phosphate.[24] The mevalonate and deoxyxylulose phosphate pathways are responsible for the biosynthesis of terpenoid metabolites, including carotenoids. Green cauliflower contained high levels of carotenoids, but low levels of organic acids participated in the tricarboxylic acid cycle produced from acetyl-CoA. Other building blocks based on seven amino acids, alanine, isoleucine, leucine, methionine, phenylalanine, tryptophan, and valine, which were identified at high levels in green cauliflower, are precursors of glucosinolates.[25] Glucosinolates are sulfur-rich secondary metabolites characteristic of the Brassicales order and are well-known for their role in plant resistance to insects and their cancer-preventive properties.[26] They are found largely in the economically and nutritionally important *Brassica* crops, such as oilseed rape, cabbage, broccoli, and cauliflower. Volden et al. reported that green cauliflower had higher total glucosinolate levels compared to white and purple cauliflower.[10] Therefore, these

dx.doi.org/10.1021/jf401330e | J. Agric. Food Chem. 2013, 61, 6999−7007

Table 3. Metabolites Identified in GC-TOFMS Chromatograms of Purple Cauliflower (cv. Graffiti) Extracts

no.[a]	compound	RT[b]	RRT[c]	quantification ion[d]
1	pyruvic acid	4.55	0.426	174
2	lactic acid	4.65	0.435	147
3	valine	5.04	0.471	146
4	alanine	5.15	0.481	116
5	glycolic acid	6.25	0.584	147
6	serine	6.81	0.636	116
7	ethanolamine	6.89	0.644	174
8	glycerol	6.91	0.646	147
9	leucine	6.93	0.648	158
10	isoleucine	7.14	0.668	158
11	proline	7.23	0.676	142
12	nicotinic acid	7.26	0.679	180
13	glycine	7.28	0.681	174
14	succinic acid	7.36	0.688	147
15	glyceric acid	7.46	0.697	147
16	fumaric acid	7.70	0.720	245
17	threonine	7.98	0.746	219
18	β-alanine	5.40	0.786	174
19	malic acid	8.49	0.831	147
20	salicylic acid	9.15	0.856	267
21	aspartic acid	9.16	0.856	100
22	methionine	9.21	0.861	176
23	pyroglutamic acid	9.25	0.865	156
24	4-aminobutyric acid	9.28	0.868	174
25	threonic acid	9.43	0.882	147
26	arginine	9.93	0.928	142
27	glutamic acid	9.96	0.931	246
28	phenylalanine	10.09	0.943	218
29	*p*-hydroxybenzoic acid	10.11	0.945	223
30	xylose	10.18	0.951	103
31	asparagine	10.36	0.969	116
IS	ribitol	10.70	1.000	217
32	glutamine	11.14	1.041	156
33	shikimic acid	11.32	1.058	204
34	citric acid	11.41	1.067	273
35	quinic acid	11.66	1.090	345
36	fructose	11.76	1.100	103
37	galactose	11.89	1.112	147
38	glucose	11.95	1.118	147
39	mannose	12.08	1.129	147
40	manitol	12.18	1.139	319
41	inositol	13.21	1.235	305
42	ferulic acid	13.33	1.246	338
43	tryptophan	14.05	1.314	202
44	sinapic acid	14.22	1.329	338
45	sucrose	16.15	1.510	217
46	cellobiose	16.66	1.557	204
47	trehalose	16.69	1.560	191
48	raffinose	19.79	1.850	217

[a]Numbers represent the compound index for the chromatogram peaks shown in Figure 2. [b]Retention time (min). [c]Relative retention time (retention time of the analyte/retention time of the IS). [d]Specific mass ion used for quantification.

results suggest that primary metabolite profiling could provide important information concerning compositional metabolite differences between genotypes and possible links between primary and secondary metabolites. Recently, using MS-based metabolite profiling and multivariate analyses (PCA and orthogonal PLS-DA),

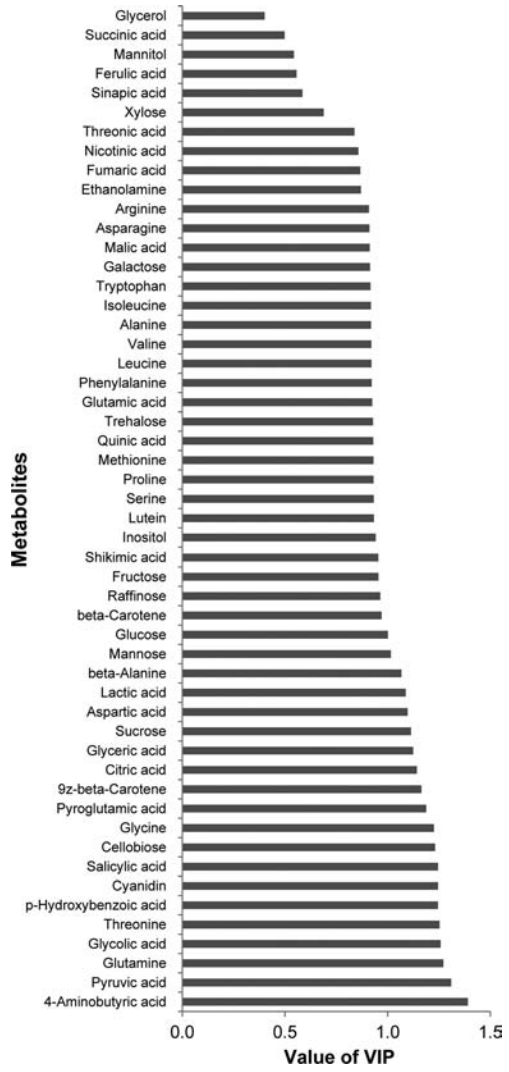

Figure 4. Variable importance in the prediction (VIP) value of variables from the PLS-DA analysis with data on 52 metabolites of four cauliflower varieties.

Cho et al. identified the quantitative relationship between the levels of phenolic amino acids and phenylpropanoids in tobacco responses to pathogen inoculation,[27] and Lee et al. characterized the association between the antioxidant activity and related metabolites in different sizes of *Aloe vera*, in which larger size plants with the highest antioxidant activity had relatively high contents of sugars and anthraquinones derivatives (as phenolic metabolites) modified by the addition of a sugar molecule.[28] Our recent study demonstrated that hydrophilic metabolite profiling, combined with chemometrics, is a useful tool for the identification of metabolic networks connecting primary and secondary metabolism in rice grain.[17]

Correlations between Levels of Metabolites in Colored Cauliflowers. To examine the relationship between levels of 52 metabolites in cauliflower florets, Pearson's correlation analysis and HCA of the accessions were performed (Figure 5). Correlation analysis is a useful technique for determining the strength of a relationship between two quantitative samples and can be applied to the discovery of associations between metabolites belonging to a biological system. Steuer demonstrated that intrinsic fluctuations of a metabolic system induce a characteristic pattern of interdependencies between metabolites,

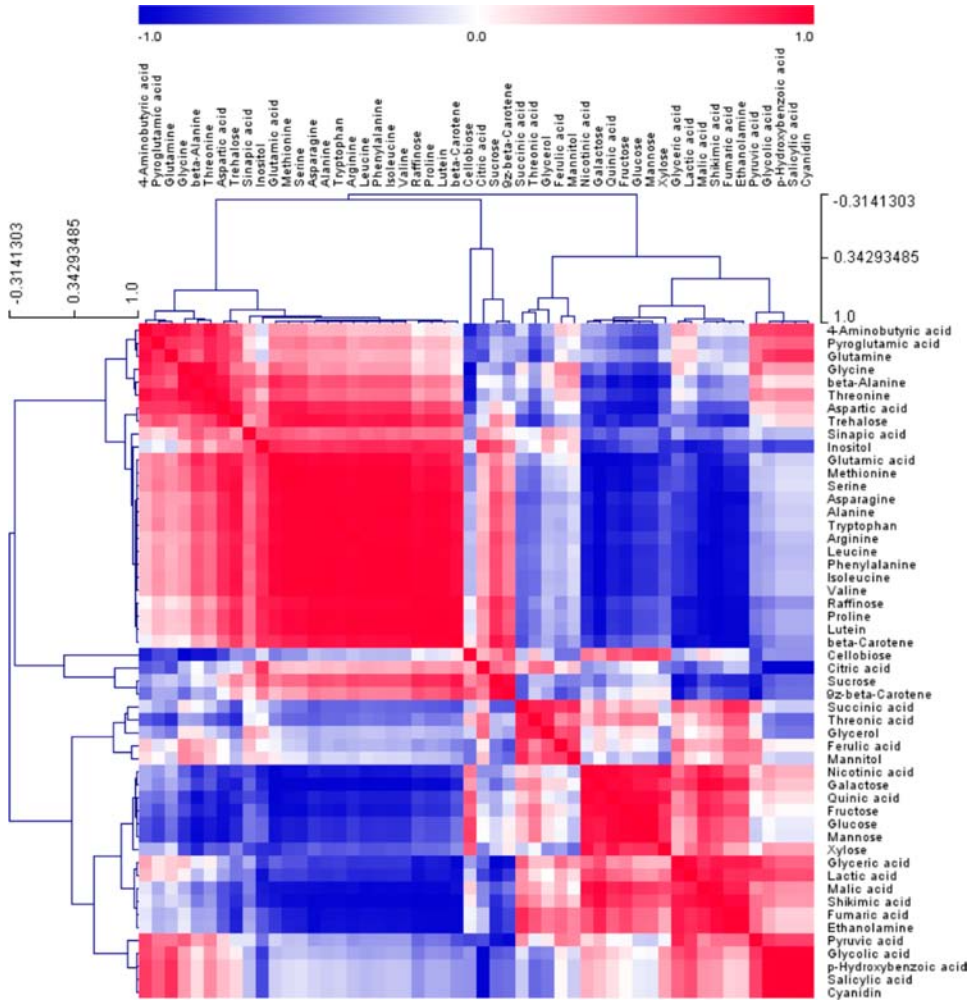

Figure 5. Correlation matrix and cluster analysis of results obtained from data on 52 metabolites of four cauliflower varieties. Each square indicates the Pearson's correlation coefficient of a pair of compounds, and the value for the correlation coefficient is represented by the intensity of the blue or red color, as indicated on the color scale. Hierarchical clusters are represented by a cluster tree.

depending on genetic and experimental backgrounds, which can be exploited by comparative correlation analysis.[29] As shown in Figure 4, metabolites from the same class, such as amino acids, organic acids, or sugars, are generally clustered together. The most distinct aspect was found in amino acids, which revealed significant positive correlations among each other, but negative correlations with most compounds of other classes, such as organic acids and sugars. In contrast, citric acid, sinapic acid, inositol, raffinose, sucrose, and trehalose clustered within a group including amino acids rather than groups including organic acids or sugars. These results were consistent with the findings from PLS-DA loading plots (Figure 3B), indicating that PLS-DA can be used to visualize complex data. Within the amino acid cluster, isoleucine contents were positively correlated with the branched amino acids, leucine (r = 0.9985, p < 0.0001) and valine (r = 0.9967, p < 0.0001). Likewise, significant positive relationships were apparent between threonine and glycine (r = 0.9496, p < 0.0001) or threonine and aspartic acid (r = 0.9357, p < 0.0001) that are biologically linked amino acids.[13,30] These results revealed a high correlation between metabolites that participate in closely related pathways, demonstrating the robustness of the present experimental system.

To further explore relationships among the genotypes and levels of 52 metabolites, a heatmap was created to visualize metabolic differences among samples (Figure 6). All samples were correctly classified according to their genotype, which indicated that green cauliflower was separated from the others. The heatmap showed differences in the relative metabolite levels of the four cauliflower varieties. The most notable differences were observed in metabolite levels of green cauliflower, which showed higher levels of amino acids and carotenoids and lower levels of organic acids compared to the other three varieties. These results are in agreement with the PLS-DA score plots (Figure 3A). All multivariate tools used showed similar classification of various cauliflowers based on genotype, which indicates that metabolic profiling, combined with chemometrics, can discriminate between plants at the genotypic level, as well as track metabolic links.

In conclusion, our data demonstrate the diversity of health beneficial substances (carotenoids, anthocyanins, and phenolic acids) and identify 48 hydrophilic metabolites including amino acids, organic acids, sugar, and sugar alcohols, in various colored cauliflowers (white, yellow, green, and purple). Carotenoid contents were considerably higher in green cauliflower, and anthocyanin was detected only in purple cauliflower. Total phenolic acid levels comprising free, esterified, and bound forms were relatively higher in both green and purple cauliflower. These results provide valuable information regarding future

dx.doi.org/10.1021/jf401330e | J. Agric. Food Chem. 2013, 61, 6999−7007

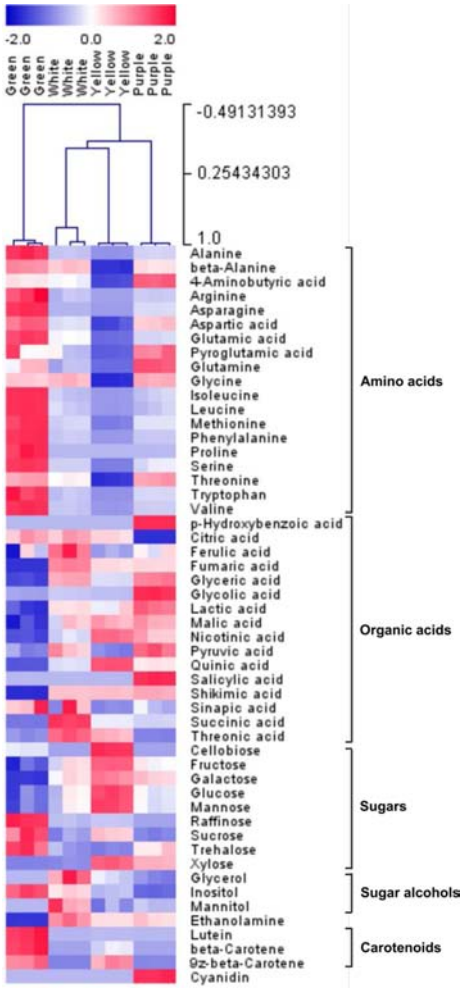

Figure 6. Heatmap representing differences in relative metabolite levels of four cauliflower varieties. Red or blue indicates that the metabolite content is decreased or increased, respectively. Metabolites were categorized into amino acids, organic acids, sugars, sugar alcohols, and carotenoids.

breeding programs for cauliflower or new vegetables with increased health benefits, including carotenoids, anthocyanins, and phenolic acids. Using the PLS-DA, Pearson correlation, and HCA multivariate analyses, it was possible to identify meaningful compositional differences and relationships between core primary metabolites in four cauliflower varieties. A strong correlation between the metabolites was observed in those that participated in closely related pathways. Green cauliflower had higher amino acid and lower organic acid levels compared with white, yellow, and purple cauliflowers, which were considerably related to secondary metabolism. The results of the present study suggest the usefulness of GC-TOFMS-based metabolite profiling, combined with chemometrics, as a tool for determining phenotypic variation and identifying metabolic links between primary and secondary metabolism.

AUTHOR INFORMATION

Corresponding Author

*(S.-U.P.) Phone: +82-42-821-5730. Fax: +82-42-822-2631. E-mail: supark@cnu.ac.kr. (J.K.K.) Phone: +82-31-299-1154. Fax: +82-31-299-1122. E-mail: kjkpj@korea.kr.

Funding

This study was supported by the National Academy of Agricultural Science (Code PJ009654) and the Next-Generation BioGreen 21 Program (SSAC, PJ009520), Rural Development Administration, Republic of Korea.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Cartea, M. E.; Francisco, M.; Soengas, P.; Velasco, P. Phenolic compounds in *Brassica* vegetables. *Molecules* **2011**, *16*, 251−280.

(2) Lampe, J. W.; Peterson, S. Brassica, biotransformation and cancer risk: genetic polymorphisms alter the preventive effects of cruciferous vegetables. *J. Nutr.* **2002**, *132*, 2291−2994.

(3) Podsędek, A. Natural antioxidants and antioxidant capacity of *Brassica* vegetables: a review. *LWT−Food Sci. Technol.* **2007**, *40*, 1−11.

(4) Jahangir, M.; Kim, H. K.; Choi, Y. H.; Verpoorte, R. Health-affecting compounds in *Brassicaceae*. *Compr. Rev. Food Sci. Food Saf.* **2009**, *8*, 31−43.

(5) Johnson, T. I. Glucosinolates in the human diet. Bioavailability and implications for health. *Phytochem. Rev.* **2002**, *1*, 183−188.

(6) Ayaz, F. A.; Hayırlıoglu-Ayaz, S.; Alpay-Karaoglu, S.; Grúz, J.; Valentová, K.; Ulrichová, J.; Strnad, M. Phenolic acid contents of kale (*Brassica oleraceae* L. var. *acephala* DC.) extracts and their antioxidant and antibacterial activities. *Food Chem.* **2008**, *107*, 19−25.

(7) Balouchi, Z.; Peyvast, G. A.; Ghasemnezhad, M.; Saadatian, M. Changes of antioxidant compounds of broccoli (*Brassica oleracea* L. var. *Italica*) during storage at low and high temperatures. *South−West. J. Hortic. Biol. Environ.* **2011**, *2*, 193−212.

(8) Li, H.; Deng, Z.; Zhu, H.; Hu, C.; Liu, R.; Young, J. C.; Tsao, R. Highly pigmented vegetables: anthocyanin compositions and their role in antioxidant activities. *Food Res. Int.* **2012**, *46*, 250−259.

(9) Lo Scalzo, R.; Genna, A.; Branca, F.; Chedin, M.; Chassaigne, H. Anthocyanin composition of cauliflower (*Brassica oleracea* L. var. botrytis) and cabbage (*B. oleracea* L. var. capitata) and its stability in relation to thermal treatments. *Food Chem.* **2008**, *107*, 136−144.

(10) Volden, J.; Bengtsson, G. B.; Wicklund, T. Glucosinolates, L-ascorbic acid, total phenols, anthocyanins, antioxidant capacities and colour in cauliflower (*Brassica oleracea* L. ssp. *botrytis*); effects of long-term freezer storage. *Food Chem.* **2009**, *112*, 967−976.

(11) Hoekenga, O. A. Using metabolomics to estimate unintended effects in transgenic crop plants: problems, promises, and opportunities. *J. Biomol. Tech.* **2008**, *19*, 159−166.

(12) Kim, J. K.; Park, S. Y.; Yeo, Y.; Cho, H. S.; Kim, Y. B.; Bae, H.; Park, C. H.; Lee, J. H.; Park, S. U. Metabolic profiling of millet (*Panicum miliaceum*) using gas chromatography-time-of-flight mass spectrometry (GC-TOFMS) for quality assessment. *Plant Omics* **2013**, *6*, 73−78.

(13) Carreno-Quintero, N.; Acharjee, A.; Maliepaard, C.; Bachem, C. W.; Mumm, R.; Bouwmeester, H.; Visser, R. G.; Keurentjes, J. J. Untargeted metabolic quantitative trait loci analyses reveal a relationship between primary metabolism and potato tuber quality. *Plant Physiol.* **2012**, *158*, 1306−1318.

(14) Park, S. Y.; Park, W. T.; Park, Y. C.; Ju, J. I.; Park, S. U.; Kim, J. K. Metabolomics for the quality assessment of *Lycium chinense* fruits. *Biosci., Biotechnol., Biochem.* **2012**, *76*, 2188−2194.

(15) Kim, J. K.; Lee, S. Y.; Chu, S. M.; Lim, S. H.; Suh, S. C.; Lee, Y. T.; Cho, H. S; Ha, S. H. Variation and correlation analysis of flavonoids and carotenoids in Korean pigmented rice (*Oryza sativa* L) cultivars. *J. Agric. Food Chem.* **2010**, *58*, 12804−12809.

(16) Park, S. Y.; Ha, S. H.; Lim, S. H.; Jung, J. Y.; Lee, S. M.; Yeo, Y.; Kim, J. K. Determination of phenolic acids in Korean rice (*Oryza sativa* L.) cultivars using gas chromatography-time-of-flight mass spectrometry. *Food Sci. Biotechnol.* **2012**, *21*, 1141−1148.

(17) Kim, J. K.; Park, S. Y.; Lim, S. H.; Yeo, Y.; Cho, H. S.; Ha, S. H. Comparative metabolic profiling of pigmented rice (*Oryza sativa* L.) cultivars reveals primary metabolites are correlated with secondary metabolites. *J. Cereal Sci.* **2013**, *57*, 14−20.

(18) Lu, Y.; Du, J.; Tang, J.; Wang, F.; Zhang, J.; Huang, J.; Liang, W.; Wang, L. Environmental regulation of floral anthocyanin synthesis in *Ipomoea purpurea*. *Mol. Ecol.* **2009**, *18*, 3857−3871.

(19) Sikora, E.; Cieślik, E.; Filipiak-Florkiewicz, A.; Leszczyńska, T. Effect of hydrothermal processing on phenolic acids and flavonols contents in selected brassica vegetables. *Acta Sci. Pol., Technol. Aliment.* **2012**, *11*, 45−51.
(20) Ramadan, Z.; Jacobs, D.; Grigorov, M.; Kochhar, S. Metabolic profiling using principal component analysis, discriminant partial least squares, and genetic algorithms. *Talanta* **2006**, *68*, 1683−1691.
(21) Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* **2003**, *17*, 166−173.
(22) Brereton, R. G. Validation and optimization. In *Chemometrics for Pattern Recognition*; Wiley: Chichester, UK, 2009; pp 311−391.
(23) Kim, J. K.; Park, S. Y.; Na, J. K.; Seong, E. S.; Yu, C. Y. Metabolite profiling based on lipophilic compounds for quality assessment of perilla (*Perilla frutescens*) cultivars. *J. Agric. Food Chem.* **2012**, *60*, 2257−2263.
(24) Dewick, P. M. Secondary metabolism: the building blocks and construction mechanisms. In *Medicinal Natural Products: A Biosynthetic Approach*, 2nd ed.; Wiley: Chichester, UK, 2002; pp 7−12.
(25) Grubb, C. D.; Abel, S. Glucosinolate metabolism and its control. *Trends Plant Sci.* **2006**, *11*, 89−100.
(26) Sønderby, I. E.; Geu-Flores, F.; Halkier, B. A. Biosynthesis of glucosinolates − gene discovery and beyond. *Trends Plant Sci.* **2010**, *15*, 283−290.
(27) Cho, K.; Kim, Y.; Wi, S. J.; Seo, J. B.; Kwon, J.; Chung, J. H.; Park, K. Y.; Nam, M. H. Nontargeted metabolite profiling in compatible pathogen-inoculated tobacco (*Nicotiana tabacum* L. cv. Wisconsin 38) using UPLC-Q-TOF/MS. *J. Agric. Food Chem.* **2012**, *60*, 11015−11028.
(28) Lee, S.; Do, S. G.; Kim, S. Y.; Kim, J.; Jin, Y.; Lee, C. H. Mass spectrometry-based metabolite profiling and antioxidant activity of *Aloe vera* (*Aloe barbadensis* Miller) in different growth stages. *J. Agric. Food Chem.* **2012**, *60*, 11222−11228.
(29) Steuer, R. On the analysis and interpretation of correlations in metabolomic data. *Brief. Bioinf.* **2006**, *7*, 151−158.
(30) Baxter, C. J.; Liu, J. L.; Fernie, A. R.; Sweerlove, L. J. Determination of metabolic fluxes in a non-steady-state system. *Phytochemistry* **2007**, *68*, 2313−2319.