Compositional analyses of diverse phytochemicals and polar metabolites from different-colored potato (*Solanum tubersum* L.) tubers

Wonhui Lee¹ · Yunsoo Yeo¹ · Seonwoo Oh¹ · Kwang-Soo Cho² · Young-Eun Park² · Soon Ki Park³ · Si Myung Lee¹ · Hyun Suk Cho¹ · Soo-Yun Park¹

Received: 24 January 2017 / Revised: 13 June 2017 / Accepted: 27 June 2017 / Published online: 29 August 2017
© The Korean Society of Food Science and Technology and Springer Science+Business Media B.V. 2017

Abstract Lipophilic bioactive compounds and hydrophilic primary metabolites from potato (*solanum tubersum* L.) tubers with different-colored flesh (white-, yellow-, red-, and purple) were characterized. The carotenoid content was relatively higher in red-colored potatoes, in which lutein was most plentiful. Among the other lipophilic compounds analyzed, including policosanols, tocopherols, and phytosterols, octacosanol was measured in the largest amount, followed by β -sitosterol, irrespective of color variations. Forty-three hydrophilics consisting of amino acids, organic acids, sugars, and sugar alcohols and 18 lipophilics were subjected to data-mining processes. The results of multivariate statistical analyses clearly distinguished the different varieties and separated red-fleshed potatoes from other color-fleshed potatoes according to abundance of amino acids, sugars, and carotenoids. This study confirmed the metabolic association-related biochemical pathway between metabolite characteristic and color differences in potato tubers. These results can facilitate understanding the metabolic differences among diverse colored potatoes and provide fruitful information for genetic engineering of potato cultivars.

Keywords Metabolomics · Multivariate analysis · Phytochemicals · Potato · Primary metabolites

Introduction

Bioactive compounds have attracted attention for their effects on human health. Antioxidant activity and phytochemical properties are of particularly great interest, as they may reduce the risks of cardiovascular disease, cancers, and degenerative diseases [1]. Numerous bioactive compounds, consisting of a great variety of chemical structures and functions, are present in all plants. Some epidemiologic studies have demonstrated that a diet rich in whole grains, fruits, and vegetables has favorable effects on chronic diseases by promoting the scavenging of free radicals [2]. For these reasons, dietary intake of high levels of antioxidant materials is regarded as necessary to improve the quality of a healthy lifestyle.

Potatoes are one of the most-consumed staple food materials and are grown worldwide. They are recognized as a good source of carbohydrates and of vitamins C and B, in addition to minerals such as potassium, phosphorous, and calcium. Furthermore, potato contains high concentrations of phytonutrients, which provide antioxidant activities [3, 4]. There are many different phytonutrients in potato, including carotenoids, flavonoids, phenolics, and anthocyanins. These compounds, which are present as small molecules and secondary metabolites, are important because of their beneficial role in many aspects of human health [5, 6]. Thus, potato is an edible crop with abundant health-beneficial bioactive compounds. Qualitative variations in the crops of different cultivars could potentially contribute to increasing the availability of nutrients and phytochemicals to the breeding and food industries. Advances in analytical technology through time will make it possible to identify and understand the functions of a variety of phytochemicals, particularly their antioxidant properties.

✉ Soo-Yun Park
psy22@korea.kr

¹ National Institute of Agricultural Sciences, Rural Development Administration, Jeonju, Jeonbuk 54874, Korea

² National Institute of Crop Science, Rural Development Administration, Pyeongchang, Gangwon 25342, Korea

³ College of Agriculture and Life Sciences, Kyungpook National University, Daegu 41566, Korea

Fig. 1 Images of the four potato varieties (*Solanum tuberosum* L.). (A) cv. Atlantic, (B) cv. Haryeong, (C) cv. Hongyoung, (D) cv. Jayoung

In this study, as bioactive secondary metabolites, carotenoids were quantified using high-performance liquid chromatography (HPLC), and the quality of lipophilic composition, including policosanols, tocopherols, and phytosterols, was evaluated using gas chromatography-time-of-flight mass spectrometry (GC-TOFMS) in potatoes (*solanum tuberosum* L.) with peculiar flesh colors (white, yellow, red, and purple). In addition, hydrophilic primary metabolites were profiled using GC-TOFMS to identify phenotypic differences between core metabolite contents. GC-TOFMS analysis is a helpful tool because of its highly sensitive reproducibility. In addition, quick spectra accumulation times enable high handling capacity for analysis [7]. The data obtained were subjected to multivariate statistical analyses, principal component analysis (PCA), partial least-squares discriminant analysis (PLS-DA), orthogonal projections to latent structures discriminant analysis (OPLS-DA), Pearson correlation analysis, and hierarchical clustering analysis (HCA) to demonstrate relationships among metabolite contents. These data will facilitate future potato breeding strategies and be useful for developing new functional potatoes with high levels of health benefits.

Materials and methods

Plant material and chemicals

Four potato (*Solanum tuberosum* L.) cultivars, ‘Atlantic’ (white), ‘Haryoung’ (yellow), ‘Hongyoung’ (red), and ‘Jayoung’ (purple) bred by Highland Agriculture Research Institute (HARI) were used in this study. After storage at 4 °C for 4 months, four seed potatoes were kept in a greenhouse with scattered light condition for sprouting. Sprouted seed potatoes were planted at HARI (Pyeongchang, Korea; 37°40.50'N, 128°43'47.6"E) on May 13, 2015. After 120 days cultivation, the potatoes were harvested on September 13, 2015 and stored in the room that

steadily maintains 95% humidity and 4 °C for 30 days (Fig. 1). After finely chopped, the lyophilized samples at −80 °C for 4 days were pulverized into a very fine powder using a planetary mono mill (pulverisette 6; Fritsch, Idar-Oberstein, Germany) for componential analysis. And then, the samples milled were stored at −80 °C prior to metabolite extraction.

Violaxanthin, antheraxanthin, lutein, zeaxanthin, and (all-*E*)- β -carotene were obtained from CaroteNature (Lupsingen, Switzerland). Eicosanol (c20), heneicosanol (c21), Docosanol (c22), tricosanol (c23), tetracosanol (c24), hexacosanol (c26), heptacosanol (c27), octacosanol (c28), triacontanol (c30), cholesterol, stigmasterol, β -sitosterol, 5 α -cholestane, β -apo-8'-carotenal, ribitol, pyridine, ascorbic acid, and *N*-methyl-*N*-trimethylsilyltrifluoro-actamide (MSTFA) were from Sigma (St. Louis, MO, USA). α -Tocopherol and methoxyamine hydrochloride were purchased from Merck (Darmstadt, Germany) and Thermo Fisher Scientific (Pittsburgh, PA, USA), respectively.

Lipophilic metabolite analysis

Carotenoids were analyzed by HPLC according to the previous method [8]. The carotenoid extracts from potato samples (0.3 g), which were extracted using hexane (3 mL) and dried under a stream of nitrogen, were redissolved in dichloromethane/methanol (50:50 v/v, 125 μ L). The extracts were injected to a C30 YMC column (250 $\times$ 4.6 mm, 3 μ m; YMC Co., Kyoto, Japan) by an Agilent 1100 HPLC instrument (Massy, France) combined with a photodiode array detector (450 nm). Solvent A consisting of methanol/water (92:8 v/v) with 10 mM ammonium acetate and solvent B of 100% methyl *tert*-butyl ether were used for gradient elution under the following conditions (1 mL/min): 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. Carotenoids were quantified using calibration curves by

plotting four concentrations of the commercial standards. β -Apo-8'-carotenal was used as an internal standard (IS).

Identification of policosanols, tocopherol, and phytosterols were carried out according to the published method [9]. Lipophilic metabolites were extracted from the milled samples (0.2 g) combining with 5 α -cholestane as an IS. The lyophilized extracts were derivatized with 30 μ L of MSTFA and 30 μ L of pyridine followed by shaking incubation (1200 $\times$ g) at 60 °C for 30 min using a thermomixer comfort (Eppendorf AG, Hamburg, Germany). The derivatized extracts (1 μ L) was separated on a 30 m $\times$ 0.25-mm i.d. fused silica capillary column coated with 0.25- μ m CP-SIL 8 CB low bleed (Varian Inc., Palo Alto, CA, USA) with a split ratio of 10 by a 7890 A gas chromatograph (Agilent, Atlanta, GA, USA) with a 7683B autosampler (Agilent). The injector temperature was 290 °C and the flow rate of helium gas was 1.0 mL/min. The temperature program was set to 200 °C and held for 2 min followed by a 10 °C/min to 310 °C and 10 min maintenance. A Pegasus HT TOF mass spectrometer (LECO, St. Joseph, MI, USA) was used for mass analysis of the samples and ranged for m/z 50–600. The transfer line, the ion-source temperatures, and detector voltage were 280, 230 °C, and 1800 V, respectively. The 13 lipophilic standards ranging from 0.01 to 10.0 μ g and 5 α -cholestane as an IS were used for preparing calibration curve.

Hydrophilic metabolites analysis

Hydrophilic compounds were analyzed based on previous method [8]. The derivatization was conducted by mixing methoxyamine hydrochloride and pyridine followed by shaking at 30 °C for 90 min. And then MSTFA was added in the mixture for trimethylsilyl (TMS) etherification followed by the shaking incubation at 37 °C for 30 min. A derivatized sample (1 μ L) was injected into the same column with GC-TOFMS as used for policosanols, tocopherol, and phytosterols analyses. 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 initial temperature of 80 °C for 2 min followed by an rise to 320 °C at 15 °C/min and 10 min maintenance at 320 °C. The transfer line temperatures, ion-source temperatures, and scanned mass range were 250, 200 °C, and m/z 85–600 respectively. The mass detector voltage was set to 1800 V.

Statistical analysis

All analyses were carried out with five independent samples for biological replicates. Data obtained were subjected to analysis of variance (ANOVA), and Duncan's multiple-range test (SAS 9.2; SAS Institute, Cary, NC, USA).

Quantification data achieved from GC-TOFMS and HPLC were applied to PCA and PLS-DA (SIMCA-P version 12.0; Umetrics, Umeå, Sweden) to visualize the correlations in complex data. OPLS-DA was subsequently performed to gain clear separation. The dataset scaled with unit variance scaling were subjected to PCA, PLS-DA and OPLS-DA. Pearson's correlation analysis was conducted using the SAS 9.2 software package (SAS Institute, Cary, NC, USA). HCA using the correlation coefficients were carried out by the software Multi-Experiment Viewer version 4.9.0 (<http://www.tm4.org/mev/>).

Result and discussion

Carotenoid, policosanol, tocopherol, and phytosterol contents of colored potatoes

HPLC analysis was carried out to determine the carotenoid contents from distinct colored potato tubers. In this study, violaxanthin/neoxanthin, antheraxanthin, lutein, zeaxanthin, and β -carotene were identified in potato tubers. The two compounds violaxanthin and neoxanthin could not be distinguished by the analytical method applied in this study. Carotenoids were identified according to retention times of commercial standards and compared with previous results [10]. Among the carotenoids identified, the most abundant component was lutein, which constituted more than 50% of the total carotenoid accumulation in all samples, except for yellow potato, in which it was about 38% (Table 1). Antheraxanthin was not detected in white or purple potatoes, whereas violaxanthin/neoxanthin was identified in only yellow potatoes. Zeaxanthin was exceptionally abundant in red potatoes, and β -carotene was detected in small amounts in all varieties of different-colored potatoes.

Carotenoids are accumulated in most vegetables, fruits, and flowers as an essential group of organic pigments responsible for coloration, especially red, orange, and yellow colors [11]. Because animals and humans cannot synthesize carotenoids *de novo*, these need to be provided from the diet. α - and β -carotene supply the principal dietary sources of vitamin A and serve a variety of biological activities as antioxidants, which improve the immune system and reduce the risks of cardiovascular and degenerative diseases [12, 13]. Lutein and zeaxanthin occur in the major components of the human retina and function to protect against age-related macular degeneration [14]. Previous studies demonstrating the profile of carotenoid contents in potatoes showed varied compositions in accordance with color, genotype, season, and storage characteristics. Morris et al. [15] reported that antheraxanthin, lutein, violaxanthin, and zeaxanthin are

Table 1 Contents ($\mu\text{g/g}$; dry weight basis) of lipophilic bioactive phytochemicals in colored potatoes

	White	Yellow	Red	Purple
<i>Carotenoids</i>				
Viola + Neo	ND	0.12 ± 0.01	ND	ND
Antheraxanthin	ND	$0.17 \pm 0.02\text{a}$	$0.06 \pm 0.02\text{b}$	ND
Lutein	$0.15 \pm 0.04\text{c}$	$0.20 \pm 0.06\text{c}$	$0.83 \pm 0.19\text{a}$	$0.36 \pm 0.01\text{b}$
Zeaxanthin	ND	ND	0.08 ± 0.02	ND
β -Carotene	$0.04 \pm 0.01\text{b}$	$0.04 \pm 0.01\text{b}$	$0.10 \pm 0.03\text{a}$	$0.05 \pm 0.01\text{b}$
Sum	$0.22 \pm 0.11\text{c}$	$0.64 \pm 0.26\text{b}$	$1.31 \pm 0.67\text{a}$	$0.49 \pm 0.16\text{b}$
<i>Policosanols</i>				
C20	$1.00 \pm 0.25\text{a}$	$1.35 \pm 0.49\text{a}$	$1.56 \pm 0.55\text{a}$	$1.42 \pm 0.56\text{a}$
C21	$4.39 \pm 0.77\text{a}$	$6.17 \pm 2.55\text{a}$	$5.48 \pm 1.76\text{a}$	$4.71 \pm 0.99\text{a}$
C22	$4.94 \pm 1.27\text{a}$	$6.30 \pm 2.60\text{a}$	$5.56 \pm 1.84\text{a}$	$5.45 \pm 1.32\text{a}$
C23	$1.98 \pm 0.48\text{ab}$	$1.13 \pm 0.49\text{b}$	$2.16 \pm 1.04\text{a}$	$1.21 \pm 0.69\text{ab}$
C24	$4.90 \pm 1.22\text{a}$	$4.61 \pm 1.93\text{a}$	$4.12 \pm 1.38\text{a}$	$3.97 \pm 0.82\text{a}$
C26	$7.09 \pm 1.76\text{a}$	$5.63 \pm 2.48\text{a}$	$6.52 \pm 2.02\text{a}$	$6.92 \pm 1.05\text{a}$
C27	$2.17 \pm 0.58\text{a}$	$1.63 \pm 0.71\text{ab}$	$1.00 \pm 0.44\text{b}$	$1.87 \pm 0.58\text{a}$
C28	$34.17 \pm 7.86\text{a}$	$32.88 \pm 15.32\text{a}$	$35.35 \pm 10.84\text{a}$	$38.20 \pm 6.24\text{a}$
C30	$5.56 \pm 0.99\text{bc}$	$3.74 \pm 1.57\text{c}$	$6.53 \pm 2.74\text{ab}$	$8.73 \pm 1.18\text{a}$
Sum	$65.19 \pm 16.24\text{a}$	$63.06 \pm 28.37\text{a}$	$66.78 \pm 18.18\text{a}$	$70.95 \pm 13.98\text{a}$
<i>Tocopherol and sterols</i>				
α -Tocopherol	$0.78 \pm 0.37\text{b}$	$1.37 \pm 0.39\text{a}$	$0.66 \pm 0.29\text{b}$	$0.93 \pm 0.43\text{ab}$
Cholesterol	$3.25 \pm 0.54\text{b}$	$2.91 \pm 0.37\text{b}$	$5.14 \pm 0.14\text{a}$	$4.83 \pm 0.77\text{a}$
Stigmasterol	$7.62 \pm 1.38\text{b}$	$8.02 \pm 1.46\text{b}$	$11.69 \pm 3.07\text{a}$	$10.58 \pm 2.11\text{ab}$
β -Sitosterol	$26.51 \pm 4.28\text{a}$	$16.75 \pm 1.80\text{bc}$	$20.17 \pm 8.05\text{ab}$	$11.02 \pm 4.40\text{c}$
Sum	$40.96 \pm 9.95\text{a}$	$31.96 \pm 8.43\text{ab}$	$41.86 \pm 18.57\text{a}$	$31.29 \pm 5.27\text{b}$

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 $\pm$ standard deviation ($n = 5$). White, Atlantic; yellow, Haryeong; red, Hongyoung; purple, Jayoung; viola, violaxanthin; neo, neoxanthin; ND, not detectable; c20, eicosanol; c21, heneicosanol; c22, docosanol; c23, tricosanol; c24, tetracosanol; c26, hexacosanol; c27, heptacosanol; c28, octacosanol; c30, triacontanol

predominant carotenoids in potatoes. Zhou et al. [13] examined distinct differences in color depending on the degree of carotenoid accumulation, ranging from white to deep orange-yellow. In white potato, carotenoid could not be detected. In contrast, orange-yellow potato had the highest amount of carotenoid. Valcarcel et al. [16], who conducted analyses of the total carotenoid content of 60 varieties of potatoes, confirmed that the dark-yellow varieties had the highest level of carotenoid and that light-yellow varieties had the second highest. White, cream, and blue-colored varieties had lower levels. The results of the present study indicated that red-colored potato had the highest total carotenoid content ($1.31 \mu\text{g/g}$), followed by yellow- and purple-colored potatoes (0.64 and $0.49 \mu\text{g/g}$, respectively). White-colored potato had the lowest amount ($0.22 \mu\text{g/g}$).

As other lipophilic bioactive compounds, policosanols, tocopherols, and phytosterols were determined in four colored potato tubers using GC-TOFMS. Previously, we analyzed 17 lipophilic metabolites from perilla seeds using

Fig. 2 (A) Selected ion chromatograms of hydrophilic metabolites extracted from purple-fleshed potatoes (cv. Jayoung) as MO/TMS derivatives separated on a 30-m $\times$ 0.25-mm (i.d.) fused silica capillary column coated with a 0.25- μm CP-SIL 8 CB low bleed. (B) Score and (C) loading plots of principal components 1 and 2 of the PCA results obtained from data on 61 metabolites of four colored potatoes. 1–43, the compounds listed in Table 2; 44, c20; 45, c21; 46, c22; 47, c23; 48, c24; 49, c26; 50, c27; 51, c28; 52, c30; 53, cholesterol; 54, β -sitosterol; 55, stigmasterol; 56, α -tocopherol; 57, violaxanthin/neoxanthin; 58, antheraxanthin; 59, lutein; 60, zeaxanthin; 61, β -carotene

GC-TOFMS [17]. However, the contents of these components in different-colored potatoes had not been evaluated before. In total, thirteen metabolites, including nine policosanols, one tocopherol, and three phytosterols, were detected in the samples (Table 1). Policosanols have a variety of effective physiological functions, such as reducing platelet aggregation and endothelial damage [18]. Phytosterols provide advantages for the human blood by reducing serum cholesterol and low-density lipoprotein

while increasing high-density lipoprotein [19]. Tocopherols act as antioxidants that (especially α -tocopherol) can relieve oxidative stress [20]. In the present study, the levels of these compounds in potato tubers were evaluated, and the results showed no specific differences related to color. Octacosanol (c28) was the most abundant policosanol in all potato samples, accounting for more than 30% of the total lipophilic compounds including carotenoids. Octacosanol is involved in several important functions, such as energy metabolism, cholesterol biosynthesis, and anti-fatigue [21, 22], and has value for the development of health foods and medicines [23]. Therefore, for lipophilic bioactive materials, potato tubers are a good dietary component irrespective of color.

Classification of colored potatoes by GC-TOFMS-based metabolic profiling and multivariate analysis

Gas chromatography-mass spectrometry (GC-MS) has been used to explore metabolic differentiation among various genotypes [24, 25]. Roessner et al. [26] demonstrated metabolic differences between soil- and in vitro-grown potato tubers and between wild and transgenic potato tubers using GC-MS. We used GC-TOFMS to investigate the contents and composition of low-molecular-weight hydrophilics from the four different-colored potato tubers and defined the levels of 43 core metabolites, such as amino acids, organic acids, and sugars (Fig. 2A). ChromaTOF software, which provided peak location information by comparison with reference compounds and an in-house library, was used to assist with peak location [25]. Several compounds were identified by direct comparison of the sample mass chromatogram with those of commercial standards. The metabolites identified were quantified by selected ions (Table 2).

The quantification data of 43 primary hydrophilic metabolites identified in this study which normalized based on IS signal intensity and 18 secondary lipophilic metabolites (carotenoids, policosanols, tocopherols, and phytosterols), were applied to PCA to explore diversity in metabolic composition between the analyzed samples. As the most typical chemometric tool, PCA is useful for extracting and interpreting information [27]. Although PCA can only identify gross variability, it has been used successfully as a preliminary step in multivariate analysis to gain an overview and form a good outline in complicated results [9, 28]. Representation plots based on scores and loadings provided an overview of the patterns in the dataset and allowed examination of correlations among the 61 metabolites (Fig. 2B, C). In the score plot (Fig. 2B), the two highest-ranking components accounted for 59.1% of the total variation. All of the samples exhibiting an identical color were clustered together, except that white and

Table 2 Metabolites identified in GC-TOFMS chromatograms of purple-colored potatoes (cv. Jayoung) extracts

No. ^a	Compound	RT ^b	RRT ^c	Quantification ^d
1	Pyruvic acid	4.62	0.430	174
2	Lactic acid	4.71	0.437	147
3	Alanine	5.21	0.484	116
4	Glycolic acid	6.32	0.587	147
5	Valine	6.43	0.597	144
6	Serine	6.87	0.638	116
7	Ethanolamine	6.95	0.646	174
8	Glycerol	6.97	0.648	147
9	Leucine	6.99	0.649	158
10	Isoleucine	7.20	0.669	158
11	Proline	7.29	0.677	142
12	Nicotinic acid	7.34	0.682	180
13	Glycine	7.34	0.682	174
14	Succinic acid	7.42	0.689	147
15	Glyceric acid	7.52	0.699	147
16	Fumaric acid	7.76	0.721	245
17	Threonine	8.04	0.747	219
18	β -alanine	8.46	0.786	174
19	Malic acid	8.94	0.831	147
20	Salicylic acid	9.22	0.856	267
21	Aspartic acid	9.22	0.857	100
22	Methionine	9.27	0.861	176
23	Pyroglutamic acid	9.32	0.866	156
24	4-aminobutyric acid	9.35	0.869	174
25	Threonic acid	9.50	0.882	147
26	Arginine	9.99	0.928	142
27	Glutamic acid	10.03	0.932	246
28	Phenylalanine	10.16	0.944	218
29	Xylose	10.24	0.952	103
30	Asparagine	10.44	0.970	116
IS	Ribitol	10.76	1.000	217
31	Glutamine	11.21	1.042	156
32	Shikimic acid	11.35	1.055	204
33	Citric acid	11.49	1.068	273
34	Quinic acid	11.73	1.090	345
35	Fructose	11.80	1.097	103
36	Galactose	11.94	1.110	147
37	Glucose	11.98	1.113	147
38	Mannose	12.13	1.127	147
39	Mannitol	12.20	1.133	319
40	Inositol	13.28	1.234	305
41	Tryptophane	14.13	1.313	202
42	Sucrose	16.23	1.508	217
43	Trehalose	16.76	1.558	191

^a Number represents the compound index for the chromatogram peaks shown in Fig. 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

Fig. 3 Score and loading plots of principal components 1 and 2 of the PLS-DA results obtained from the metabolite profiles of *four colored* potatoes. The plots represent the PLS-DA results obtained from data on all of the metabolites identified in this study (A, B), the

low-molecular-weight hydrophilic compounds listed in Table 2 (C, D), and the lipophilic bioactive compounds listed in Table 1 (E, F). The numbers of the loading plots represent the same compounds as for Fig. 2

purple samples were not differentiated. Especially, red-fleshed potatoes were separated from the others by the first component, accounting for 44.6% of the total variation. The loading plot showed positive values for all amino acids and sugars. In addition, organic acids and sugar alcohols were mostly positive, with the exception of lactic acid and inositol. The PCA results indicated that the red-fleshed potatoes had higher levels of primary metabolite compounds than the other three varieties had and that the metabolic profiles for the 61 compounds for the white and

the purple potatoes did not differ significantly. Yellow-fleshed potatoes were apart from white- and purple-fleshed potatoes by the second component accounting for 14.5% of the total variation. The loadings indicated that yellow-fleshed potatoes had higher violaxanthin/neoxanthin and antheraxanthin contents compared to the white- and purple-fleshed potatoes.

To obtain additional understanding about the compositional diversity among the cultivars, the dataset was subjected to PLS-DA, which was conducted using three

Fig. 4 (A) OPLS-DA score plots derived from all of the metabolites identified in this study for *red-fleshed* potatoes and the other three varieties. (B) Top 27 metabolites ranked by variable importance by the prediction (VIP) value obtained from the OPLS-DA analysis of data on 61 metabolites of *four potatoes*. (C) Correlation matrix and cluster analysis of results obtained from data on 61 metabolites of

different *flesh-color potatoes*. Each *square* indicates the Pearson's correlation coefficient for 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

different datasets individually: all of the 61 metabolites, the 43 hydrophilic metabolites, and the 18 lipophilic metabolites. PLS-DA is preferable to PCA for sample distinction by reducing the dimension [28]. The same separation pattern shown by the PCA results (Fig. 2B) was observed in the score plots of the PLS-DA results for all metabolites (Fig. 3A) and hydrophilic metabolites (Fig. 3C). Variation in the hydrophilic metabolites was mainly attributable to amino acids and organic acids. Amino acids such as valine,

serine, glycine, and threonine grouped on the very left of the loadings plot, whereas organic acids including pyruvic, lactic, and quinic acids clustered on the right side (Fig. 3D). The loadings plot revealed higher levels of amino acids in red potatoes than in white, yellow, or purple potatoes and lower levels of organic acids in the red potatoes than in the others. These results confirmed that the core hydrophilic metabolites offered helpful information for metabolic distinction among genotypes. Furthermore,

Fig. 4 continued

the four cultivars could be distinguished by combining components 1 and 2 of the PLS-DA for the 18 lipophilic metabolite profiles, which showed that yellow potato tubers were significantly different from white, red, and purple potato tubers (Fig. 3E). This difference was mainly attributable to violaxanthin/neoxanthin, antheraxanthin, and α -tocopherol, which were highly detected in the yellow-fleshed potato tubers (Fig. 3F).

To confirm the metabolite's contributions to the separation between the red potato tubers and the other three varieties, the OPLS-DA method was applied to the dataset (Fig. 4A). OPLS can be considered a method for addressing the problem presented in PLS-DA, namely indirect variation correlation, especially as the number of classes

increases, by incorporating an orthogonal signal correction (OSC) filter into a PLS model [29]. The advanced separation model facilitates the interpretability of the data by extracting information on changes in the metabolite composition of samples [30]. The contribution of each variable to the separation was confirmed by the importance of the variable in the projection (VIP) value (Fig. 4B). Generally, variable with a VIP value of more than 1 is considered as the most significant variables in a dataset [31]. Twenty-four primary metabolites and 3 carotenoids had significant VIP values (>1). These findings suggest that the primary metabolites, particularly amino acids and sugars, differed significantly among the differently colored potato tubers. Zhang et al. [32] reported that in vitro, sucrose and

mannitol administration up-regulate carotenoid levels in citrus juice sacs. Here, it was shown that red-fleshed potatoes have higher sucrose, mannitol, and carotenoids. Therefore, these results are indicative of the capability of core hydrophilics profiling for confirming metabolic diversity between varieties and identifying potential connections between primary and secondary metabolites. Our recent studies showed that hydrophilic metabolite profiling incorporating chemometrics is an effective method for determining metabolic links associating primary and secondary metabolism in diverse edible crops [7, 31].

Correlations among metabolite levels in colored potatoes

To explore specific networks among the levels of the 61 components identified in potato tubers, we conducted Pearson's correlation analysis and HCA on the dataset. Correlation analysis is a handy tool for discovering the intensity of an association between two quantitative variables and can be applied to determine relationships among metabolites in a biological system [33]. Using the method, strong correlations can be found from compounds located in biologically tightly connected pathways.

Among the amino acids, isoleucine contents were definitely related with the amino acids leucine ($r = 0.9475$, $p < 0.0001$) and valine ($r = 0.9076$, $p < 0.0001$). Similarly, a remarkable relationship between serine and glycine was detected ($r = 0.9421$, $p < 0.0001$). The Pearson correlation coefficients among γ -aminobutyric acid, glutamine, and proline, which are closely linked biosynthetically as members of the glutamic acid family, were higher than 0.7 [34]. Another noticeable association was appeared between phenylalanine and tryptophan ($r = 0.8912$, $p < 0.0001$), which are participated in biologically linked pathway as aromatic amino acids. Shikimic acid content was significantly associated with content of salicylic acid ($r = 0.8539$, $p < 0.0001$) that is produced from the shikimate pathway as phenolic acid. The HCA results gained from the Pearson's correlation coefficients showed three distinctive components groups that are indicated using a dotted box (Fig. 4C). The most distinctive aspect was observed in policosanols, all of which in this study formed a group, with the exception of triacontanol (c30-ol). A group contained aromatic amino acids (phenylalanine and tryptophan) and sugars (galactose, fructose, glucose, mannose, and sucrose). Another group included all amino acids, except aromatic amino acids, as was also reported in previous metabolomics studies of diverse potato varieties [34, 35]. These results showed that most metabolites formed clusters based on their biochemical nature (i.e., amino acids, sugars, and policosanols).

In conclusion, we identified 61 metabolites, including hydrophilic core primary metabolites and health-benefiting lipophilic metabolites, in potato tubers with different-colored flesh (white, yellow, red, and purple). These substances beneficial to health were quantified to explore compositional diversity according to color variation. Carotenoid level was fairly higher in red-fleshed potatoes than in other varieties, and lutein was the most plentiful carotenoid. Other lipophilic compounds, policosanols, tocopherols, and phytosterols did not differ significantly among different-colored potato varieties. These results provide valuable information for future traditional breeding programs or genetic engineering approaches for developing new foods containing higher levels of carotenoids. Nevertheless, potato tubers can be used irrespective of color as a good dietary component of lipophilic bioactive materials. The metabolite profiles obtained in this study were applied to data-mining means such as PCA, PLS-DA, OPLS-DA, Pearson correlation, and HCA multivariate analyses, which made it possible to verify significant metabolic variation and associations between core primary metabolites in the four potato cultivars. A powerful relationship was detected for metabolites that involved in tightly linked pathways of biological systems. All multivariate analyses revealed a comparable potato categorization according to genotype, in which red-fleshed potatoes had higher amino acid, sugar, and carotenoid content than the other three potatoes. This study confirmed metabolite profiling in combination with chemometrics are useful techniques for determining core metabolites in the central pathways of plant metabolism and for identifying phenotypic diversity and metabolic networks among primary and secondary metabolism.

Acknowledgements This study was supported by the National Institute of Agricultural Sciences (Code PJ011752), Rural Development Administration, Republic of Korea.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

1. Prakash D, Gupta KR. The antioxidant phytochemicals of nutraceutical importance. *Open Nutraceuticals J.* 2: 20–35 (2009)
2. Kris-Etherton PM, Hecker KD, Bonanome A, Coval SM, Binkoski AE, Hilpert KF, Griel AE, Etherton TD. Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer. *Am. J. Med.* 113: 71–88 (2002)
3. Jansen G, Flamme W, Schüller K, Vandrey M. Tuber and starch quality of wild and cultivated potato species and cultivars. *Potato Res.* 44: 137–146 (2001)
4. Camire ME, Kubow S, Donnelly DJ. Potatoes and human health. *Crit. Rev. Food Sci. Nutr.* 49: 823–840 (2009)

5. Friedman M. Chemistry, biochemistry, and dietary role of potato polyphenols. A review. *J. Agric. Food Chem.* 45: 1523–1540 (1997)
6. Fernandez-Orozco R, Gallardo-Guerrero L, Hornero-Méndez D. Carotenoid profiling in tubers of different potato (*Solanum* sp) cultivars: accumulation of carotenoids mediated by xanthophyll esterification. *Food Chem.* 141: 2864–2872 (2013)
7. Kim JK, Park SY, Lim SH, Yeo Y, Cho HS, Ha SH. Comparative metabolic profiling of pigmented rice (*Oryza sativa* L.) cultivars reveals primary metabolites are correlated with secondary metabolites. *J. Cereal Sci.* 57: 14–20 (2013)
8. Park SY, Lee SY, Yang JW, Lee JS, Oh SD, Oh S, Lee SM, Lim MH, Park SK, Jang JS, Cho HS, Yeo Y. Comparative analysis of phytochemicals and polar metabolites from colored sweet potato (*Ipomoea batatas* L.) tubers. *Food Sci. Biotechnol.* 25: 283–291 (2016)
9. Kim TJ, Lee KB, Baek SA, Choi J, Ha SH, Lim SH, Park SY, Yeo Y, Park SU, Kim JK. Determination of lipophilic metabolites for species discrimination and quality assessment of nine leafy vegetables. *J. Korean Soc. Appl. Biol. Chem.* 58: 909–918 (2015)
10. Park SY, Choi SR, Lim SH, Yeo Y, Kweon SJ, Bae YS, Kim KW, Im KH, Ahn SK, Ha SH, Park SU, Kim JK. Identification and quantification of carotenoids in paprika fruits and cabbage, kale, and lettuce leaves. *J. Korean Soc. Appl. Biol. Chem.* 57: 355–358 (2014)
11. Lu S, Li L. Carotenoid metabolism: biosynthesis, regulation, and beyond. *J. Integr. Plant Biol.* 50: 778–785 (2008)
12. Rao AV, Rao LG. Carotenoids and human health. *Pharmacol. Res.* 55: 207–216 (2007)
13. Zhou X, McQuinn R, Fei Z, Wolters AM, VAN Eck J, Brown C, Giovannoni JJ, Li L. Regulatory control of high levels of carotenoid accumulation in potato tubers. *Plant Cell. Environ.* 34: 1020–1030 (2011)
14. Ahmed SS, Lott MN, Marcus DM. The macular xanthophylls. *Surv. Ophthalmol.* 50: 183–193 (2005)
15. Morris WL, Ducreux L, Griffiths DW, Stewart D, Davies HV, Taylor MA. Carotenogenesis during tuber development and storage in potato. *J. Exp. Bot.* 55: 975–982 (2004)
16. Valcarcel J, Reilly K, Gaffney M, O'Brien N. Total carotenoids and L-ascorbic acid content in 60 varieties of potato (*Solanum tuberosum* L.) grown in Ireland. *Potato Res.* 58: 29–41 (2015)
17. Kim JK, Park SY, Na JK, Seong ES, Yu CY. Metabolite profiling based on lipophilic compounds for quality assessment of perilla (*Perilla frutescens*) cultivars. *J. Agric. Food Chem.* 60: 2257–2263 (2012)
18. Irmak S, Dunford NT. Policosanol contents and compositions of wheat varieties. *J. Agric. Food Chem.* 53: 5583–5586 (2005)
19. Arruzazabala ML, Molina V, Mas R, Fernández L, Carbajal D, Valdés S, Castaño G. Antiplatelet effects of policosanol (20 and 40 mg/day) in healthy volunteers and dyslipidaemic patients. *Clin. Exp. Pharmacol. Physiol.* 29: 891–897 (2002)
20. Haffner SM. Clinical relevance of the oxidative stress concept. *Metabolism* 49: 30–34 (2000)
21. Kim H, Park S, Han DS, Park T. Octacosanol supplementation increases running endurance time and improves biochemical parameters after exhaustion in trained rats. *J. Med. Food* 6: 345–351 (2003)
22. Oliaro-Bosso S, Calcio Gaudino E, Mantegna S, Giraudo E, Meda C, Viola F, Cravotto G. Regulation of HMGCoA reductase activity by policosanol and octacosadienol, a new synthetic analogue of octacosanol. *Lipids* 44: 907–916 (2009)
23. Long L, Wu S, Sun J, Wang J, Zhang H, Qi G. Effects of octacosanol extracted from rice bran on blood hormone levels and gene expressions of glucose transporter protein-4 and adenosine monophosphate protein kinase in weaning piglets. *Anim. Nutr.* 1: 293–298 (2015)
24. Park SY, Park WT, Park YC, Ju JI, Park SU, Kim JK. Metabolomics for the quality assessment of *Lycium chinense* fruits. *Biosci. Biotechnol. Biochem.* 76: 2188–2194 (2012)
25. Kim YB, Park SY, Park CH, Park WT, Kim SJ, Ha SH, Arasu MV, Al-Dhabi NA, Kim JK, Park SU. Metabolomics of differently colored *Gladiolus* cultivars. *Appl. Biol. Chem.* 59: 597–607 (2016)
26. Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L. Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J.* 23: 131–142 (2000)
27. Ramadan Z, Jacobs D, Grigorov M, Kochhar S. Metabolic profiling using principal component analysis, discriminant partial least squares, and genetic algorithms. *Talanta* 68: 1683–1691 (2006)
28. Barker M, Rayens W. Partial least squares for discrimination. *J. Chemometrics* 17: 166–173 (2003)
29. Worley B, Powers R. Multivariate analysis in metabolomics. *Curr. Metabolomics* 1: 92–107 (2013)
30. Madala NE, Piater LA, Steenkamp PA, Dubery IA. Multivariate statistical models of metabolomics data reveals different metabolite distribution patterns in isonitrosoacetophenone-elicited *Nicotiana tabacum* and *Sorghum bicolor* cells. *Springerplus* 3: 254 (2014)
31. Park SY, Lim SH, Ha SH, Yeo Y, Park WT, Kwon DY, Park SU, Kim JK. Metabolite profiling approach reveals the interface of primary and secondary metabolism in colored cauliflowers (*Brassica oleracea* L. ssp. *botrytis*). *J. Agric. Food Chem.* 61: 6999–7007 (2013)
32. Zhang L, Ma G, Kato M, Yamawaki K, Takagi T, Kiriwa Y, Ikoma Y, Matsumoto H, Yoshioka T, Nesumi H. Regulation of carotenoid accumulation and the expression of carotenoid metabolic genes in citrus juice sacs *in vitro*. *J. Exp. Bot.* 63: 871–886 (2012)
33. Camacho D, Fuente A, Mendes P. The origin of correlations in metabolomics data. *Metabolomics* 1: 53–63 (2005)
34. Carreno-Quintero N, Acharjee A, Maliepaard C, Bachem CW, Mumm R, Bouwmeester H, Visser RG, Keurentjes JJ. Untargeted metabolic quantitative trait loci analyses reveal a relationship between primary metabolism and potato tuber quality. *Plant Physiol.* 158: 1306–1318 (2012)
35. Dobson G, Shepherd T, Verrall SR, Conner S, McNicol JW, Ramsay G, Shepherd LVT, Davies HV, Stewart D. Phytochemical diversity in tubers of potato cultivars and landraces using a GC–MS metabolomics approach. *J. Agric. Food Chem.* 56: 10280–10291 (2008)