J Ginseng Res 38 (2014) 187–193

Contents lists available at ScienceDirect

Journal of Ginseng Research

journal homepage: http://www.ginsengres.org

Research article

Discrimination of white ginseng origins using multivariate statistical analysis of data sets

Hyuk-Hwan Song [1], Ji Young Moon [2], Hyung Won Ryu [1], Bong-Soo Noh [3], Jeong-Han Kim [4], Hyeong-Kyu Lee [1], Sei-Ryang Oh [1,*]

[1] Natural Medicine Research Center, Korea Research Institute of Bioscience and Biotechnology, Ochang, Korea
[2] Experiment Research Institute of National Agricultural Products Quality Management Service, Gimcheon, Korea
[3] Department of Food Science and Technology, Seoul Women's University, Seoul, Korea
[4] Department of Agricultural Biotechnology, Seoul National University, Seoul, Korea

ARTICLE INFO

Article history:
Received 18 November 2013
Received in Revised form
20 March 2014
Accepted 21 March 2014
Available online 3 April 2014

Keywords:
ginsenoside
metabolomics
Panax ginseng Meyer
white ginseng

ABSTRACT

Background: White ginseng (*Panax ginseng* Meyer) is commonly distributed as a health food in food markets. However, there is no practical method for distinguishing Korean white ginseng (KWG) from Chinese white ginseng (CWG), except for relying on the traceability system in the market.
Methods: Ultra-performance liquid chromatography quadrupole time-of-flight mass spectrometry combined with orthogonal partial least squares discrimination analysis (OPLS-DA) was employed to discriminate between KWG and CWG.
Results: The origins of white ginsengs in two test sets (1.0 μL and 0.2 μL injections) could be successfully discriminated by the OPLS-DA analysis. From OPLS-DA S-plots, KWG exhibited tentative markers derived from ginsenoside Rf and notoginsenoside R3 isomer, whereas CWG exhibited tentative markers derived from ginsenoside Ro and chikusetsusaponin Iva.
Conclusion: Results suggest that ultra-performance liquid chromatography quadrupole time-of-flight mass spectrometry coupled with OPLS-DA is an efficient tool for identifying the difference between the geographical origins of white ginsengs.

1. Introduction

Ginseng (*Panax ginseng* Meyer) is a multifunctional therapeutic herb that is commonly used throughout the world. Primarily in East Asia, ginseng has been used as traditional medicine to enhance the immune system, control blood pressure, and strengthen the cardiovascular system [1]. The ginseng herb is processed using various methods. For example, peeled ginseng root turns white when dried in the sun, which has led to it being called white ginseng, whereas red ginseng is produced by steaming and drying. A wide variety of pharmacological properties have been reported for ginseng, such as anti-oxidant, anti-stress, neuroprotective, hypoglycemic, and anti-tumor effects [2–5]. The ginseng herb and ginseng-derived products include multiple secondary metabolites, such as protopanaxadiol (PPD)-type (e.g., ginsenoside Rb1, Rb2, Rc, Rd, and Rg3), protopanaxatriol (PPT)-type (e.g., ginsenoside Rg1, Re, Rf, and Rg2), and oleanane (OCO)-type ginsenosides (e.g., ginsenoside Ro) [6]. Different ginsenoside ratios have been reported for different species, geographical origins, and processing methods, and such ratios are considered to be responsible for the different bioactivities [7,8].

Metabolomics primarily focuses on comprehensive and quantitative profiling for small-molecule metabolites in a biological system. It has been applied to a variety of areas, such as plant toxicology, nutrition, and systems biology [9–11]. Multiple analytical methods, including nuclear magnetic resonance, gas chromatography-mass spectrometry, and liquid chromatography-mass spectrometry, have been applied in metabolic profiling in

* Corresponding author. Natural Medicine Research Center, Korea Research Institute of Bioscience and Biotechnology, 30-Yeongudanji-ro, Ochang-eup, Chungcheongbuk-do 363-883, Korea.
E-mail address: seiryang@kribb.re.kr (S.-R. Oh).

1226-8453/$ – see front matter Copyright © 2014, The Korean Society of Ginseng, Published by Elsevier. All rights reserved.
http://dx.doi.org/10.1016/j.jgr.2014.03.002

order to differentiate *Panax* species [12–14]. Among the various analytical methods, ultra-performance liquid chromatography quadrupole time-of-light mass spectrometry (UPLC-QTOF/MS) is used in comprehensive and reliable ginsenoside profiling for various ginseng products [15–17]. In certain studies, morphological and chemical methods were used to discriminate Korean ginseng from other *P. ginseng* sources [14,18]. Recently, metabolomics research has been used to discriminate the origin of ginseng products [19]. Despite this, ginsenosides have not been fully investigated as chemical markers despite their pharmacological importance. In our study, a metabolomics approach, combining a UPLC-QTOF/MS-based analysis with orthogonal partial least squares discrimination analysis (OPLS-DA), is used to determine the geographical origin of white ginsengs. The present study manifested that the statistical model (OPLS-DA) would facilitate the discrimination of Korean white ginseng (KWG) and Chinese white ginseng (CWG) origins in concert with the UPLC-QTOF/MS. Furthermore, the prediction model exhibited statistical reliability and could be applied to discriminate samples in the market.

2. Materials and methods

2.1. Chemicals and materials

High-performance liquid chromatography-grade acetonitrile and methanol were obtained from SK Chemicals Co. (Seongnam, Korea). The aqueous solutions were prepared using ultrapure water from a Milli-Q system (18.2 MΩ, Millipore, Bedford, MA, USA). Leucine-enkephalin and formic acid were purchased from Sigma-Aldrich (St. Louis, MO, USA). The white ginseng samples were provided by the Experiment Research Institute of National Agricultural Products Quality Management Service. KWG (53 samples) was obtained from several Korean markets in 2008–2009. CWG (10 samples from China and eight samples from Korea) was purchased from several vendors in China and Korea during 2006–2009 (Table 1). All samples were verified by the National Agricultural Products Quality Management Service and were used for origin identification. Reference standards of ginsenoside Rg1 (5), ginsenoside Re (6), ginsenoside Rf (9), 20(*R*)-ginsenoside Rh1 (11), ginsenoside Ra2 (14), ginsenoside Rb1 (15), ginsenoside Rc (17), ginsenoside Ra1 (18), ginsenoside Rb3 (22), ginsenoside Rb2 (23), and ginsenoside Rd (28) were provided by Fleton Natural Products Co., Ltd. (Chengdu, China). The standards were dissolved in methanol to obtain stock solutions at approximately 1.0 mg/mL and were stored at 4°C.

Table 1
Details of the white ginseng samples

No.	Year	Market place	No.	Year	Market place
K02	2009	Imsil	K31–K35	2009	Chungcheongbuk-do
K03, K04	2009	Gunsan	K36–K43	2009	Yeongju
K05, K06	2009	Geochang	K44	2009	Muan
K07	2009	Seoul	K45, K46, K49	2009	Hamyang
K08	2009	Gimje	K47	2009	Gochang
K09	2009	Seocheon	K48	2009	Dangjin
K10	2009	Gumi	K50	2009	Hampyeong
K11	2009	Boryeong	K51, K52	2009	Jeollabuk-do
K12, K13	2009	Miryang	K53	2009	Gangjin
K14	2009	Jeongeup	K54	2009	Daejeon
K15	2009	Buan	S01–S03, S06–S12	2009	China
K16–K21	2008	Yeongju	S13	2006	Gunsan (made in China)
K22–K24	2009	Geumsan	S14, S21	2008	Seoul (made in China)
K25, K27	2009	Hapcheon	S15	2008	Gimhae (made in China)
K26	2009	Inje	S17	2008	Daegu (made in China)
K28	2009	Iksan	S18	2008	Naju (made in China)
K29	2009	Damyang	S19	2008	Iksan (made in China)
K30	2009	Hongcheon	S20	2008	Suwon (made in China)

2.2. Sample preparation

The ginseng samples were dried and pulverized to powder using a mill and passed through a 40-mesh sieve. The fine ginseng powder was weighed (0.4 g) and extracted with 5 mL of 70% methanol in an ultrasonic waterbath for 60 min [13]. The extract was filtered through a syringe filter (0.22 μm) and injected directly into the UPLC system.

2.3. UPLC-QTOF/MS analysis

Ginseng metabolite profiling was performed using the ACQUITY UPLC system (Waters Corporation, Milford, MA, USA), which was equipped with a binary solvent delivery manager and a sample manager coupled to a Micromass Q-TOF Premier mass spectrometer (Waters Corporation, Milford, MA, USA) with an electrospray interface. Chromatographic separation was performed using an ACQUITY BEH C_{18} chromatography column (Waters Corporation; 2.1 mm × 100 mm, 1.7 μm). The column temperature was maintained at 35°C, and the mobile Phases A and B were water with 0.1% formic acid and acetonitrile with 0.1% formic acid, respectively. The gradient elution program to get the ginsenoside profile was as follows: 0 min, 10% B; 0–7 min, 10–33% B; 7–14 min, 33–56% B; 14–21 min, 56–100% B; wash for 23.5 min with 100% B; and a 1.5 min recycle time. The injection volumes were 1.0 μL and 0.2 μL for each test set, and the flow rate was 0.4 mL/min. The mass spectrometer was operated in positive ion mode. N_2 was used as the desolvation gas. The desolvation temperature was 350°C, the flow rate was 500 L/h, and the source temperature was 100°C. The capillary and cone voltages were 2700V and 27V, respectively. The data were collected for each test sample from 200 Da to 1,500 Da with 0.25-s scan time and 0.01-s interscan delay over a 25-min analysis time. Leucine-enkephalin was used as the reference compound (*m/z* 556.2771 in the positive mode).

2.4. Chemometric data analysis

The raw mass data were normalized to total intensity (area) and analyzed using the MarkerLynx Applications Manager version 4.1 (Waters, Manchester, UK). The parameters included a retention time range of 4.0–19.0 min, a mass range from 200 Da to 1,500 Da, and a mass tolerance of 0.04 Da. The isotopic data were excluded, the noise elimination level was 10, and the mass and retention time windows were 0.04 min and 0.1 min, respectively. After creating a suitable processing method, the dataset was processed through the Create Dataset window. The resulting two-dimensional matrix for the measured mass values and intensities for each sample was further exported to SIMCA-P^+ software 12.0 (Umetrics, Umeå, Sweden) using both unsupervised principal component analysis and supervised OPLS-DA.

3. Results and discussion

3.1. Mass spectrometry data analysis of white ginseng ginsenosides

As shown in previous articles [13,16], the ACQUITY BEH C_{18} column (Waters Corporation) has frequently been used to separate ginsenosides from various *Panax* herbs. As presented in Fig. 1A (CWG) and Fig. 1B (KWG), 11 compounds were assigned by comparing them to standard ginsenosides and 19 ginsenosides were identified by comparing their retention time and mass spectra with the reference compounds. The compounds were further

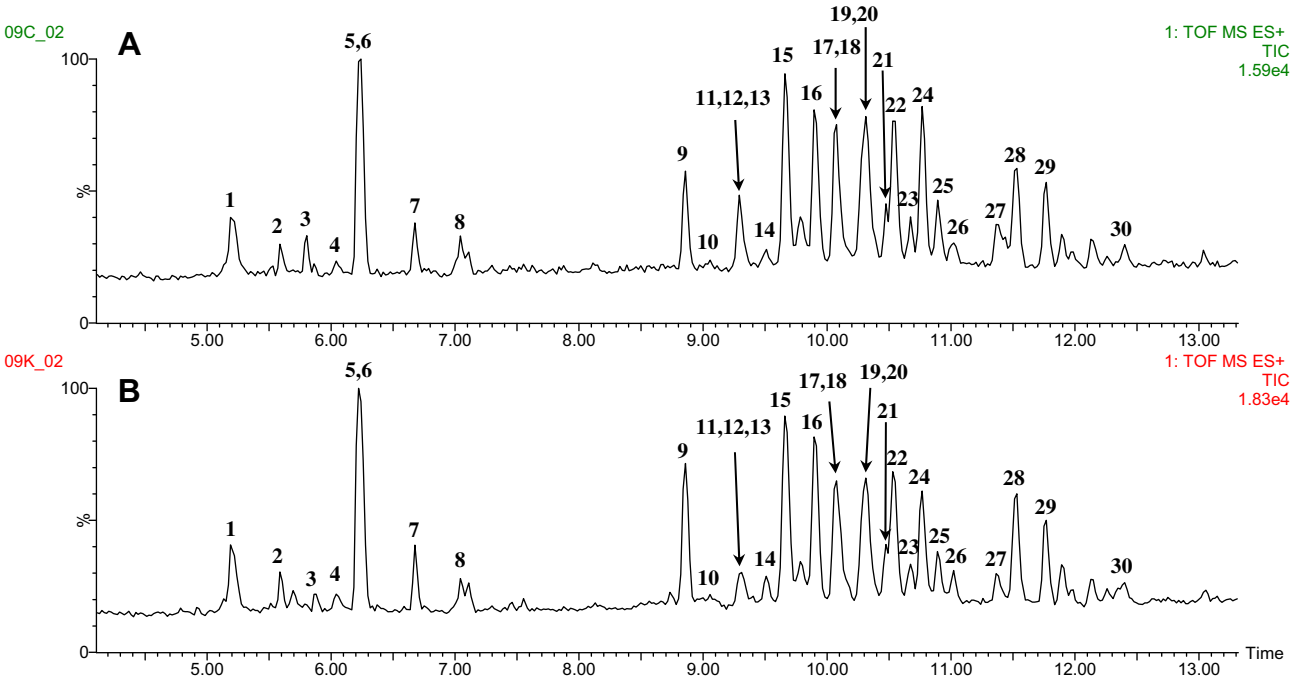

Fig. 1. Total ion current chromatograms of white ginseng extract (1.0 μL) using UPLC-QTOF/MS. (A) Chinese White Ginseng and (B) Korean White Ginseng. (1–4, notoginsenoside R3 isomer; 5, ginsenoside Rg1; 6, ginsenoside Re; 7, malonyl ginsenoside Rg1; 8, unknown; 9, ginsenoside Rf; 10, notoginsenoside R2; 11, 20(*R*)-ginsenoside Rh1; 12; notoginsenoside R4/Fa; 13, ginsenoside Ra0; 14, ginsenoside Ra2; 15, ginsenoside Rb1; 16, malonyl ginsenoside Rb1; 17, ginsenoside Rc; 18, ginsenoside Ra1; 19, ginsenoside Ro; 20, malonyl ginsenoside Rc; 21, malonyl ginsenoside Rb1 isomer; 22, ginsenoside Rb3; 23, ginsenoside Rb2; 24, malonyl ginsenoside Rb3; 25, malonyl ginsenoside Rb2; 26, quinquenoside R1; 27, chikusetsusaponin Iva; 28, ginsenoside Rd; 29, malonyl ginsenoside Rd; 30, gypenoside XV.)

Table 2
Characterization of ginsenosides in white ginseng using ultra-performance liquid chromatography quadrupole time-of-flight mass spectrometry

No.	t_R (min)	Precursor ion and/or adduct ions	Exact mass $[M+H]^+$	Error (ppm)	Formula	Identification
1	5.20	963.5590$[M+H]^+$, 985.5554$[M+Na]^+$	963.5529	−0.3	$C_{48}H_{82}O_{19}$	Notoginsenoside R3 isomer
2	5.61	963.5604$[M+H]^+$, 985.5532$[M+Na]^+$	963.5529	7.8	$C_{48}H_{82}O_{19}$	Notoginsenoside R3 isomer
3	5.69	963.5582$[M+H]^+$, 985.5528$[M+Na]^+$	963.5529	5.5	$C_{48}H_{82}O_{19}$	Notoginsenoside R3 isomer
4	6.04	963.5582$[M+H]^+$, 985.5528$[M+Na]^+$	963.5529	6.4	$C_{48}H_{82}O_{19}$	Notoginsenoside R3 isomer
5	6.22	801.5018$[M+H]^+$, 823.4845$[M+Na]^+$	801.5000	3.9	$C_{42}H_{72}O_{14}$	Ginsenoside Rg1[1)]
6	6.22	947.5637$[M+H]^+$, 969.5613$[M+Na]^+$	947.5579	6.4	$C_{48}H_{82}O_{18}$	Ginsenoside Re[1)]
7	6.68	887.5004$[M+H]^+$, 904.5289$[M+NH_4]^+$, 909.4954$[M+Na]^+$	887.5004	0	$C_{45}H_{74}O_{17}$	Malonyl ginsenoside Rg1
8	7.04	981.5855$[M+H]^+$, 998.5974$[M+NH_4]^+$, 1003.5397$[M+Na]^+$	981.5787	6.9	$C_{52}H_{84}O_{17}$	Unknown
9	8.86	801.5105$[M+H]^+$, 823.4924$[M+Na]^+$	801.5000	13.1	$C_{42}H_{72}O_{14}$	Ginsenoside Rf[1)]
10	9.06	771.4827$[M+H]^+$, 793.4720$[M+Na]^+$	771.4895	−8.8	$C_{41}H_{70}O_{13}$	Notoginsenoside R2
11	9.31	1277.9528 $[2M+H]^+$	639.447		$C_{36}H_{62}O_9$	20(*R*)- Ginsenoside Rh1[1)]
12	9.31	1241.6694$[M+H]^+$	1241.6530	13.2	$C_{59}H_{100}O_{27}$	Notoginsenoside R4/Fa
13	9.31	1271.6882$[M+H]^+$, 1293.6697$[M+Na]^+$	1271.6636	19.3	$C_{60}H_{102}O_{28}$	Ginsenoside Ra0
14	9.51	1211.6556$[M+H]^+$, 1233.6558$[M+Na]^+$	1211.6425	10.8	$C_{58}H_{98}O_{26}$	Ginsenoside Ra2[1)]
15	9.66	1109.6155$[M+H]^+$	1109.6108	4.2	$C_{54}H_{92}O_{23}$	Ginsenoside Rb1[1)]
16	9.90	1195.6158$[M+H]^+$	1195.6112	3.8	$C_{57}H_{94}O_{26}$	Malonyl ginsenoside Rb1
17	10.08	1079.6074$[M+H]^+$	1079.6002	6.7	$C_{53}H_{90}O_{22}$	Ginsenoside Rc[1)]
18	10.08	1211.6473$[M+H]^+$, 1228.6910$[M+NH_4]^+$	1211.6425	4.0	$C_{58}H_{98}O_{26}$	Ginsenoside Ra1[1)]
19	10.28	957.6287$[M+H]^+$, 974.5645$[M+NH_4]^+$	957.6210	−7.8	$C_{48}H_{92}O_{18}$	Ginsenoside Ro
20	10.31	1165.6062$[M+H]^+$, 1187.6073$[M+Na]^+$	1165.6006	4.8	$C_{56}H_{92}O_{25}$	Malonyl ginsenoside Rc
21	10.47	1195.6171$[M+H]^+$	1195.6112	4.9	$C_{57}H_{94}O_{26}$	Malonyl ginsenoside Rb1 isomer
22	10.53	1079.6013$[M+H]^+$	1079.6002	1.0	$C_{53}H_{90}O_{22}$	Ginsenoside Rb3[1)]
23	10.67	1079.6063$[M+H]^+$	1079.6002	5.7	$C_{53}H_{90}O_{22}$	Ginsenoside Rb2[1)]
24	10.77	1165.6035$[M+H]^+$	1165.6006	2.5	$C_{56}H_{92}O_{25}$	Malonyl ginsenoside Rb3
25	10.89	1165.6056$[M+H]^+$	1165.6006	10.5	$C_{56}H_{92}O_{25}$	Malonyl ginsenoside Rb2
26	11.02	1151.6244$[M+H]^+$, 1168.6555$[M+NH_4]^+$, 1173.6216$[M+Na]^+$	1151.6213	2.7	$C_{56}H_{94}O_{24}$	Quinquenoside R1
27	11.36	812.4812$[M+NH_4]^+$, 817.4389$[M+Na]^+$	795.453		$C_{42}H_{66}O_{14}$	Chikusetsusaponin Iva
28	11.53	947.5610$[M+H]^+$, 969.5427$[M+Na]^+$	947.5579	3.3	$C_{48}H_{82}O_{18}$	Ginsenoside Rd[1)]
29	11.77	1033.5590$[M+H]^+$, 1055.5431$[M+Na]^+$	1033.5583	0.7	$C_{51}H_{84}O_{21}$	Malonyl ginsenoside Rd
30	12.40	947.5607$[M+H]^+$, 969.5450$[M+Na]^+$	947.5579	3.0	$C_{48}H_{82}O_{18}$	Gypenoside XVII[1)]

[1)] Confirmed by comparison with reference standards.

confirmed through ion fragmentation patterns [20,21]. As illustrated in Table 2, white ginseng saponins were detected as protonated ions $[M+H]^+$, sodium adduct ions $[M+Na]^+$, and/or ammonium adduct ions $[M+NH_4]^+$ in the positive ion mode. The pathway for the specific fragmentation pattern supports the classification of 30 ginsenosides into three groups according to the following structures: (1) 11 compounds (peak 1–11) were identified as protopanaxatriol (PPT) type with sugar moieties attached to the C-6 and/or C-20; (2) two ginsenosides (peaks 19 and 27) were identified as OCO-type ginsenosides; and (3) the rest of compounds were identified as PPD-type with sugar moieties attached to the C-3 and/or C-20. Three types showed their own diagnostic ions in fragmentation. PPT- and PPD-type ginsenosides showed characteristic fragment ions at *m/z* 441.37 and *m/z* 425.37, respectively, indicating the losses of sugar moieties, whereas OCO-type ginsenosides showed fragment ion at *m/z* 439.36 corresponding to their aglycone. The cleaved pathways of three types were reported in previous researches [21,22].

3.2. Discrimination of white ginsengs' origin

The extracts from KWG (53 samples) and CWG (18 samples) were continuously and randomly injected into the UPLC-QTOF/MS system with a 25-min run time. Given the peaks' complexity in the UPLC chromatograms, it was difficult to distinguish between KWG and CWG through visual chromatogram observation, which indicated that the major components in the ginseng from the two origins were similar.

In this case, an effective approach for discerning differences is multivariate statistical analysis. Multivariate analysis has been widely used in the metabolomics field in recent years for extremely complex samples [23]. First, we performed principal component analysis, which is widely used as a metabolomics profiling technique for plant metabolites [24,25]. After Pareto (Par) scaling with mean-centering, the data were displayed as a score plot in a coordinate system with latent variables, "principal components" (data not shown).

Recently, supervised OPLS-DA has been widely used to study the differences between two similar groups [26]. OPLS-DA model quality can be estimated using the cross-validation parameters Q^2 (model predictability) and $R^2(y)$ (total explained variation for the X matrix). OPLS-DA for the samples produced one predictive as well as one orthogonal (1 + 3) component and showed that the cross-validated predictive ability Q^2 was 0.877, and the variance related to the differences between the two origins $R^2(y)$ was 0.992 (Fig. 2A) and cross validated analysis of variation (CV-ANOVA) $p = 2.52 \times 10^{-25}$.

Validation of an analysis model is critical for statistical multivariate analyses. We validated the analysis model by excluding certain data (a test data set) and reconstructing a new model with the remaining data (a training data set). The Y-predicted score plot indicated a confident prediction between two groups through the first predicted score (*tPS*), which summarized the X variation orthogonal to Y for the prediction set. The predicted assignment for each sample was compared to the original value, and thereby the model was evaluated for prediction accuracy and reliability. This method has been used to predict drug toxicity and geographical origin in recent metabolomics studies [27,28]. For the prediction test confidence, one-third of the samples (18 Korean and six Chinese samples) were randomly excluded and re-analyzed using the OPLS-DA model. The model for predicting their origins was established using one predictive component and one orthogonal component with $R^2(y) = 0.930$ and $Q^2 = 0.796$. The samples from the blind test were correctly assigned to their origin cluster, and the 24 analyzed samples were well predicted as shown in Fig. 2B, which indicates that the OPLS-DA model can discriminate between KWG and CWG. A variety of concentrations of ginsenosides in the sample, however, can cause difficulty in generating quantitative ion intensity for a compound in the UPLC-QTOF/MS system. As major peaks of ginsenosides were frequently saturated at a high concentration, we applied two sample sets (0.2 μL and 1.0 μL) for optimal analysis. The 0.2 μL test set model produced similar results to the 1.0 μL test set with $R^2(y) = 0.954$, $Q^2 = 0.792$, and CV-ANOVA $p = 5.37 \times 10^{-20}$ (Fig. 2C). The OPLS-DA model for predicting the ginseng origins was established using one predictive and two orthogonal components with $R^2(y) = 0.973$ and $Q^2 = 0.775$. In addition, the blind test samples were correctly assigned to their origin's cluster (Fig. 2D).

3.3. Assignment of tentative markers of white ginseng origins

A useful tool for comparing a variables' magnitude and reliability is the S-plot from the OPLS-DA model. Each point on the S-plot represents the exact mass retention time (t_R-*m/z*) pair. As a result, the white ginseng's differential variables (markers) associated with KWG and CWG are based on the threshold of variable importance in the projection (VIP) value (VIP > 1.0) from the S-plot [29]. The VIP value represents the importance of a variable in modeling both X (the projections) and Y (its correlation to all the responses). The VIP values of selected ions are enumerated in Table 3.

From the 1.0 μL injection test set, ions 1A, 1B, and 1C in Fig. 2E were the characteristics of KWG, and ions 2A–2G and 3A–3D were the characteristics of CWG. The fold values were obtained from dividing the mean value of mass intensity of KWG by the mean value of mass intensity of CWG. Ions 2A–2G, having fold values of 0.38–0.48 at t_R 9.06 min, imply that these ions originated from only one compound, which was identified as NG R2. This result is well matched with the fragmentation ion patterns of NG R2 in the MassFragment tool of MassLynx 4.1 (Waters, Manchester, UK) (Fig. 3A). It was found that ions 1A–1C, which were highly detected in KWG (fold values: 3.13–4.66) at t_R 9.05 min, were not from NG R2, although they had retention times similar to NG R2 (t_R 9.06 min). The structures of the ions could not be confirmed, but it was determined that the molecular weights were different from NG R2. Ions 3A–3D at t_R 11.36 min were assigned to chikusetsusaponin Iva, and were found by matching the molecular formula and fragment ion patterns [30]. Those ions were significant in CWG, with fold values of 0.30–0.37.

From the 0.2 μL injection test set, several ginsenoside ions were also detected in the S-plot (Fig. 2F). The fragment ion of 5A (765.4810 at t_R 8.86 min), which was assigned to ginsenoside Rf by matching the molecular formula and retention time with a standard compound, was postulated to be a tentative marker of KWG (VIP value >1.0). The ions 4A and 4B (985.5287 and 783.4919, respectively, at t_R 5.20 min) could be assigned to one of the NG R3 isomers, including 20-gluco-ginsenoside Rf, NG R6, NG M, or NG N. These isomers showed the same molecular ions and same fragmentation patterns at different retention times (peaks 1–4 in Table 2) [30,31]. From the results, ions 5A, 4A, and 4B can be postulated as tentative markers for KWG. Ions 6A–6F at t_R 10.28 min, which were assigned to ions derived from ginsenoside Ro (Fig. 3B), could be tentative markers for CWG by VIP value and fold values [32].

4. Conclusion

Two sample sets (0.2 μL and 1.0 μL) were applied in the UPLC-QTOF/MS with OPLS-DA and several ginsenosides were postulated for discriminating markers between the white ginseng sample sets

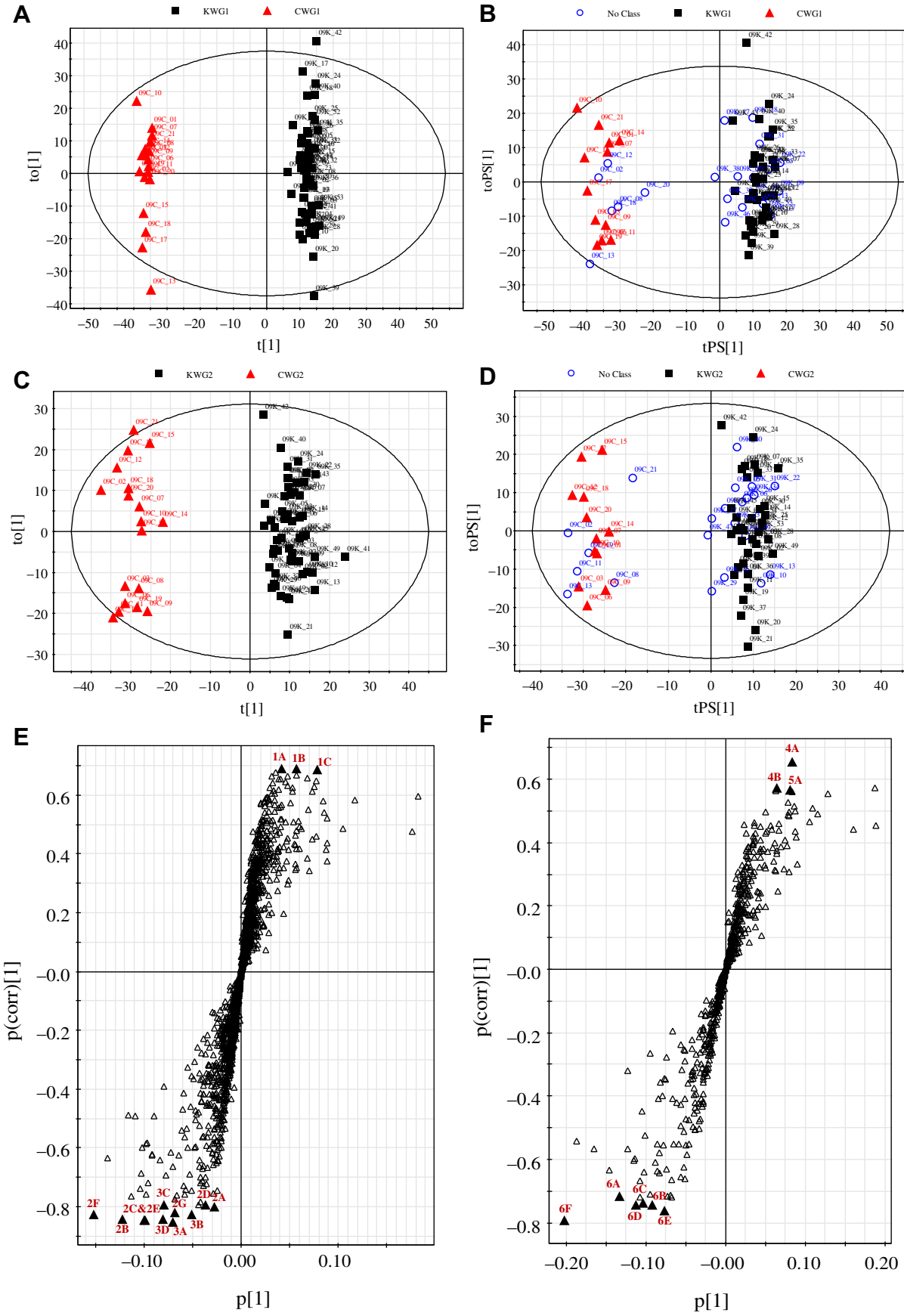

Fig. 2. Multivariate statistical analysis for Korean white ginseng (KWS) and Chinese white ginseng (CWG). (A) Orthogonal partial least squares discrimination analysis (OPLS-DA) score plots and (B) predicted score plot for the 1.0 μL injection data set, (C) OPLS-DA score plots and (D) predicted score plot for the 0.2 μL injection data set, (E) S-plot of OPLS-DA model for the 1.0 μL injection data set, and (F) S-plot of OPLS-DA model for the 0.2 μL injection data set.

Table 3
Characterization of differential variable ions from Korean white ginseng (KWG) and Chinese white ginseng (CWG)

Data set	Marker	*tR_m/z*	VIP[1)]	Formation of fragment ions	Parent compound	Average mass intensity		Fold[2)]
						KWG(53)	CWG(18)	
1.0 μL	1A	9.05_1379.6535	1.81	$[M+Na]^+$	Unknown	1.14	0.37	3.13
	1B	9.05_1357.6732	2.48	$[M+H]^+$		3.14	0.67	4.66
	1C	9.05_875.4757	3.39			6.15	1.85	3.33
	2A	9.06_771.4917	2.19	$[M+H]^+$	Notoginsenoside R2	1.70	4.52	0.38
	2B	9.06_753.4822	5.23	$[M+H-H_2O]^+$		11.08	27.15	0.41
	2C	9.06_735.4808	4.28	$[M+H-2H_2O]^+$		8.33	19.01	0.44
	2D	9.06_621.4376	1.57	$[M+H-Xyl]^+$		1.25	2.77	0.45
	2E	9.06_441.3727	4.26	$[M+H-Glc-Xyl-H_2O]^+$		8.63	19.13	0.45
	2F	9.06_423.3617	6.50	$[M+H-Glc-Xyl-2H_2O]^+$		23.24	48.32	0.48
	2G	9.06_405.3452	2.93	$[M+H-Glc-Xyl-3H_2O]^+$		4.54	9.66	0.47
	3A	11.36_817.4389	3.07	$[M+Na]^+$	Chikusetsusaponin Iva	3.07	8.31	0.37
	3B	11.36_812.4812	3.07	$[M+NH_4]^+$		1.00	2.90	0.35
	3C	11.36_633.4013	3.42	$[M+H-Glc]^+$		1.32	4.34	0.30
	3D	11.36_439.3546	3.56	$[M+H-GlcU-Glc-H_2O]^+$		3.56	10.83	0.33
0.2 μL	4A	5.20_985.5287	2.62	$[M+Na]^+$	Notoginsenoside R3 isomer	13.62	6.24	2.18
	4B	5.20_783.4919	1.62	$[M+H-Glc-H_2O]^+$		7.63	2.48	3.08
	5A	8.86_765.4810	2.06	$[M+H-2H_2O]^+$	Ginsenoside Rf	36.00	27.76	1.30
	6A	10.28_979.4910	3.52	$[M+Na]^+$	Ginsenoside Ro	65.06	83.65	0.78
	6B	10.28_974.5358	2.45	$[M+NH_4]^+$		28.86	37.52	0.77
	6C	10.28_957.6210	2.60	$[M+H]^+$		16.16	26.45	0.61
	6D	10.28_795.5720	2.90	$[M+H-Glc]^+$		22.58	35.02	0.64
	6E	10.28_633.5164	2.28	$[M+H-2Glc]^+$		9.00	14.72	0.61
	6F	10.28_439.3555	5.12	$[M+H-GlcU-2(Glc-H_2O)]^+$		84.38	121.26	0.70

[1)] Variable importance in the projection.
[2)] Fold value was calculated by dividing the mean value of ion mass intensity of KWG by that of CWG.

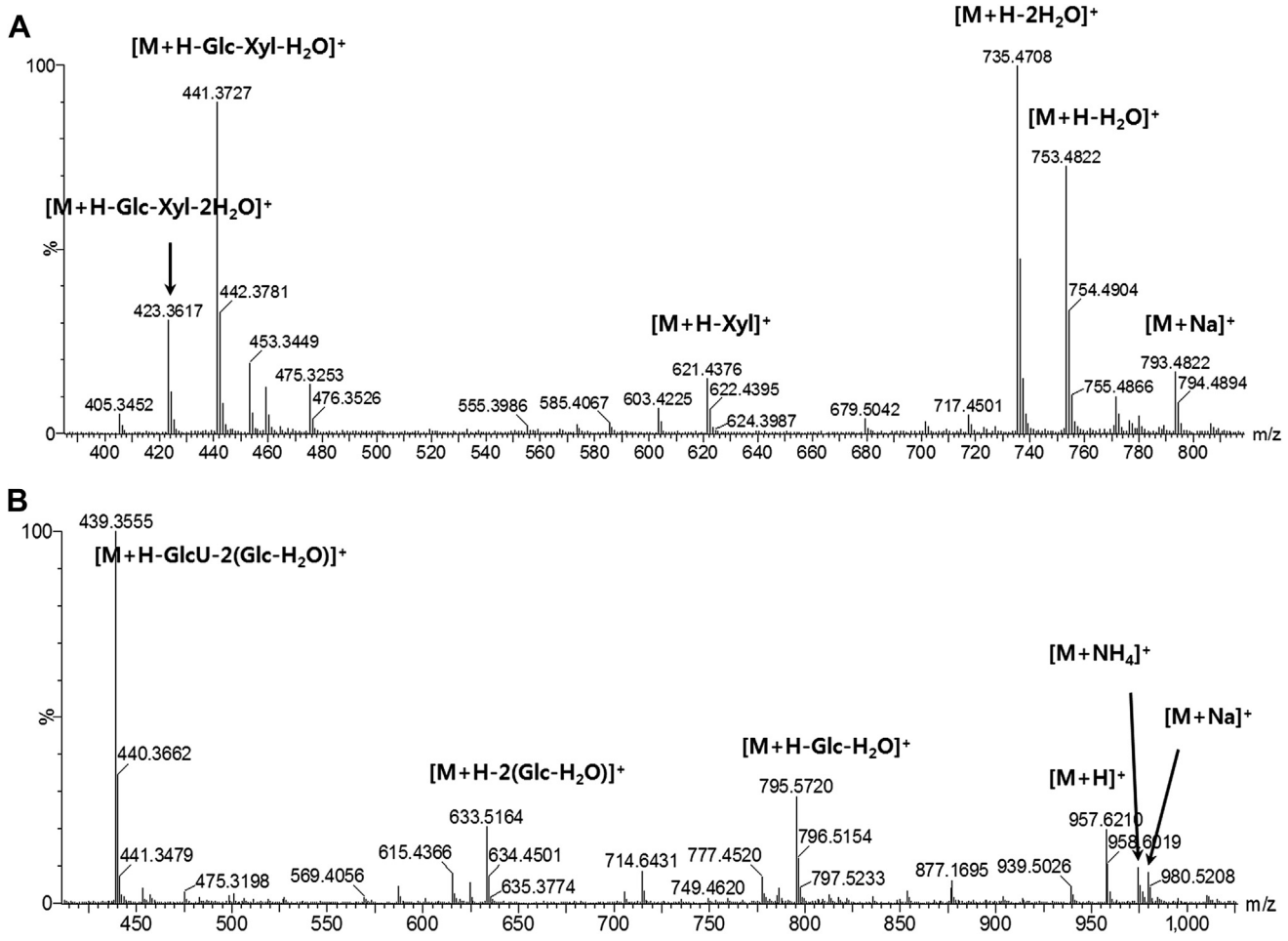

Fig. 3. Fragmentation ions patterns of tentative markers. (A) notoginsenoside R2 (793.4822; $[M+Na]^+$, t_R; 9.06 min) and (B) ginsenoside Ro (957.6210; $[M+H]^+$, t_R; 10.29 min).

originated from Korea and China. Blind tests with arbitrarily selected samples comprising one-third of the total were performed to validate the OPLS-DA model, and all of the samples were correctly assigned to their origins. Furthermore, profiling the details of the samples enabled the observation of the differences of ginsenosides between KWG and CWG. Our results suggest that the approach in the present study could be effectively applied to discriminate the geographical origins between KWG and CWG in the markets.

Conflicts of interest

All authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Korea Research Institute of Bioscience and Biotechnology Research Initiative Program (KGM1221413). This work was carried out with the support of the Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ008395), Rural Development Administration, Republic of Korea.

References

[1] Angelova N, Kong HW, van der Heijden R, Yang SY, Young HC, Kim HK, Wang M, Hankemeier T, vander Greef J, Xu G, Verpoorte R. Recent methodology in the phytochemical analysis of ginseng. Phytochem Anal 2008;19:2–16.
[2] Grandhi A, Mujumdar AM, Patwardhan B. A comparative pharmacological investigation of ashwagandha and ginseng. J Ethnopharmacol 1994;44:131–5.
[3] Rausch WD, Liu S, Gille G, Radad K. Neuroprotective effects of ginsenosides. Acta Neurobiol Exp 2006;66:369–75.
[4] Kim JS, Yoon KS, Lee YS. Antioxidant activity of main and fine roots of ginseng (*Panax ginseng* C.A. Meyer) extracted with various solvents. Food Sci Biotechnol 2008;17:46–51.
[5] Surh YJ, Na HK, Lee JY, Keum YS. Molecular mechanisms underlying anti-tumor romoting activities of heat-processed *Panax ginseng* C.A. Meyer. J Korean Med Sci 2011;16:38–41.
[6] Fuzzati N. Analysis methods of ginsenosides. J Chromatogr B 2004;812:119–33.
[7] Sengupta S, Toh SA, Sellers LA, Skepper JN, Koolwijk P, Leung HW, Yeung HW, Wong RN, Sasisekharan R, Fan FP. Modulating angiogenesis: the yin and the yang in ginseng. Circulation 2004;110:1219–25.
[8] Sievenpiper JL, Arnason JT, Leiter LA, Vuksan V. Decreasing, null and increasing effects of eight popular types of ginseng on acute postprandial glycemic indices in healthy humans: the role of ginsenosides. J Am Coll Nutr 2004;23:248–58.
[9] Watkins SM, German JB. Toward the implementation of metabolomic assessments of human health and nutrition. Curr Opin Biotech 2002;13:512–6.
[10] Weckwerth W. Metabolomics in systems biology. Annu Rev Plant Biol 2003;54:669–89.
[11] Robertson DG. Metabonomics in toxicology: a review. Toxicol Sci 2005;85:809–22.
[12] Shellie RA, Marriott PJ, Hule CW. Comprehensive two-dimensional gas chromatography (GC x GC) and GC x GC-quadrupole MS analysis of Asian and American ginseng. J Sep Sci 2003;26:1185–92.
[13] Xie GX, Ni Y, Su MM, Zhang YY, Zhao AH, Gao XF, Liu Z, Xiao PG, Jia W. Application of ultra-performance LC-TOF MS metabolite profiling techniques to the analysis of medicinal *Panax* herbs. Metabolomics 2008;4:248–60.
[14] Lee AR, Gautam M, Kim J, Shin WJ, Choi MS, Bang YS. A multianalytical approach for determining the geographical origin of ginseng using strontium isotopes, multielements, and ^{1}H NMR analysis. J Agric Food Chem 2011;59:8560–7.
[15] Liu Y, Li J, He J, Abliz Z, Qu J, Yu S. Identification of new trace triterpenoid samponins from the roots of *Panax notoginseng* by high-performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry. Rapid Commun Mass Spectrom 2009;23:667–79.
[16] Sun BS, Xu MY, Li Z, Wang YB, Sung CK. UPLC-Q-TOF-MS/MS analysis for steaming times-dependent profiling of steamed *Panax quiquefolius* and its ginsenosides transformations induced by repetitious steaming. J Ginseng Res 2012;36:277–90.
[17] Kim N, Kim K, Choi BY, Lee D, Shin YS, Bang KH, Cha SW, Lee JW, Choi HK, Jang DS, et al. Metabolomic approach for age discrimination of *Panax ginseng* using UPLC-Q-Tof MS. J Agric Food Chem 2011;59:10435–41.
[18] Chung YY, Lee MG, Chung CM, Jo JS. Comparison of plant growth and morphological characteristics among the Korean ginseng, the American ginseng and the Bamboo ginseng. J Ginseng Res 1998;22:147–53.
[19] Song HH, Kim DY, Woo S, Lee HK, Oh SR. An approach for simultaneous determination for geographical origins of Korean *Panax ginseng* by UPLC-QTOF/MS coupled with OPLS-DA models. J Ginseng Res 2013;37:341–8.
[20] Chan ECY, Yap SL, Lau AJ, Leow PC, Toh DF, Koh HL. Ultra-performance liquid chromatography/time-of-flight mass spectrometry based metabolomics of raw and steamed *Panax notoginseng*. Rapid Commun Mass Spectrom 2007;21:519–28.
[21] Chu C, Xu S, Li X, Yan J, Liu L. Profiling the ginsenosides of three ginseng products by Lc-Q-Tof/Ms. J Food Sci 2013;78:653–9.
[22] Qi LW, Wang HY, Zhang H, Wang CZ, Li P, Yuan CS. Diagnostic ion filtering to characterize ginseng saponins by rapid liquid chromatography with time-of-flight mass spectrometry. J Chromatogr A 2012;1230:93–9.
[23] Wilson ID, Nicholson JK, Castro-Perez J, Granger JH, Johnson KA, Smith BW, Plumb RS. High resolution "Ultra performance" liquid chromatography coupled to oa-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies. J Proteome Res 2005;4:591–8.
[24] Choi MY, Choi W, Park JH, Lim J, Kwon SW. Determination of coffee origins by integrated metabolomics approach of combining multiple analytical data. Food Chem 2010;121:1260–8.
[25] El-Abassy RM, Donfack P, Materny A. Discrimination between Arabica and Robusta green coffee using visible micro Raman spectroscopy and chemometric analysis. Food Chem 2011;126:1443–8.
[26] Trygg J, Wold S. Have you ever wondered why PLS sometimes needs more than one component for a single-y vector? J Chemometr 2002;16:119–28.
[27] Clayton TA, Lindon JC, Cloarec O, Antti H, Charuel C, Hanton G, Provost JP, Le Net JL, Baker D, Walley RJ, et al. Pharmaco-metabonomic phenotyping and personalized drug treatment. Nature 2006;440:1073–7.
[28] Kang J, Lee S, Kang S, Kwon HN, Park JH, Kwon SW, Park S. NMR-based metabolomics approach for the differentiation of ginseng (*Panax ginseng*) roots from different origins. Arch Pham Res 2008;31:330–6.
[29] Buscher JM, Czernik D, Ewald JC, Sauer U, Zamboni N. Cross-platform comparison of methods for quantitative metabolomics of primary metabolism. Anal Chem 2009;81:2135–43.
[30] Wan JB, Zhang QW, Hong SJ, Li P, Li SP, Wang YT. Chemical investigation of saponins in different parts of *Panax notoginseng* by pressurized liquid extraction and liquid chromatography-electrospray ionization-tandem mass spectrometry. Molecules 2012;17:5836–53.
[31] Yoshikawa M, Morikawa T, Yashiro K, Murakami T, Matsuda H. Bioactive saponins and glycosides. XIX. Notoginseng (3): immunological adjuvant activity of NGs and related saponins: structures of NGs-L, -M, and –N from the roots of *Panax notoginseng* (Burk.) F. H. Chen. Chem Pharm Bull 2001;49:1452–6.
[32] Kite GC, Howes MR, Leon CJ, Simmonds MSJ. Liquid chromatography/mass spectrometry of malonyl-ginsenosides in the authentication of ginseng. Rapid Commun Mass Spectrom 2003;17:238–44.