DOI 10.5012/bkcs.2010.31.04.825

Differentiation of Roots of *Glycyrrhiza* Species by ^{1}H Nuclear Magnetic Resonance Spectroscopy and Multivariate Statistical Analysis

Seung-Ok Yang, Sun-Hee Hyun, So-Hyun Kim, Hee-su Kim, Jaehwi Lee, Wan Kyun Whang, Min-Won Lee, and Hyung-Kyoon Choi*

*College of Pharmacy (WCU), Chung-Ang University, Seoul 156-756, Korea. *E-mail: hykychoi@cau.ac.kr*
Received December 2, 2009, Accepted January 22, 2010

To classify *Glycyrrhiza* species, samples of different species were analyzed by ^{1}H NMR-based metabolomics technique. Partial least squares discriminant analysis (PLS-DA) was used as the multivariate statistical analysis of the ^{1}H NMR data sets. There was a clear separation between various *Glycyrrhiza* species in the PLS-DA derived score plots. The PLS-DA model was validated, and the key metabolites contributing to the separation in the score plots of various *Glycyrrhiza* species were lactic acid, alanine, arginine, proline, malic acid, asparagine, choline, glycine, glucose, sucrose, 4-hydroxyphenylacetic acid, and formic acid. The compounds present at relatively high levels were glucose, and 4-hydroxyphenylacetic acid in *G. glabra*; lactic acid, alanine, and proline in *G. inflata*; and arginine, malic acid, and sucrose in *G. uralensis*. This is the first study to perform the global metabolomic profiling and differentiation of *Glycyrrhiza* species using ^{1}H NMR and multivariate statistical analysis.

Key Words: *Glycyrrhiza* species, Metabolomics, ^{1}H NMR, Partial least squares- discriminant analysis

Introduction

Dried roots of *Glycyrrhiza* species (licorice), which are used as traditional medicines and food resources in China, Japan, and Korea, exhibit various beneficial bioactivities, including antiinflammatory,[1] anticancer,[2] and antiestrogenic,[3] the inhibiting of the absorption of dietary lipids,[4] and the reducing of blood pressure.[5] Analyses of the chemical components of samples of *Glycyrrhiza* species have revealed that they contain various biochemical components, including triterpene saponins,[6] flavonoids,[7] phenolic compounds,[8] isoflavones,[9] coumarins, and stilbenoids.[10] The United States Pharmacopeia and the Japan Pharmacopeia, described that the dried roots of *G. glabra* and *G. uralensis* should contain no less than 2.5% glycyrrhizic acid calculated on a dried basis using HPLC analysis.[11,12] Also, it is indicated in the Korean Pharmacopeia that the glycyrrhizic acid and liquiritin contents of dried roots of *Glycyrrhiza* species have to be more than 2.5% and 1.0%, respectively, calculated on a dried basis.[13] The cut roots of *G. glabra* have a relatively compact texture and are yellow, while those of *G. inflata* have many pores. The cut roots of *G. uralensis* have a compact texture, yellowish-white color, a distinct cambium ring, and are thin (Figure 1).[14] However, no classification based on global metabolomic profiling using ^{1}H NMR and multivariate statistical analysis of *Glycyrrhiza* species has been reported, and their commercial value or the identification of a particular sample is generally determined by subjective observations of smell, taste, and appearance. Therefore, a practical and standardized method for identifying and assessing the quality of *Glycyrrhiza* species is needed. The metabolome refers to the total metabolites present in cells, tissues, or biofluids,[15,16] and chemometric techniques combining ^{1}H NMR and multivariate statistical analysis have been applied for the metabolomic profiling and characterization of various types of medicinal plants and foods.[17-20] Partial least squares-discriminant analysis (PLS-DA) is the most frequently used method for classification in the metabolomics field. PLS-DA extends a regression of principal component analysis (PCA) and it uses class information to maximize the separation between groups of observations. Frequently used classification method is a categorical one (categories described with dummy variables) which expresses the class membership of the statistical units.[21-23] In the present study, we performed a global metabolomic profil-

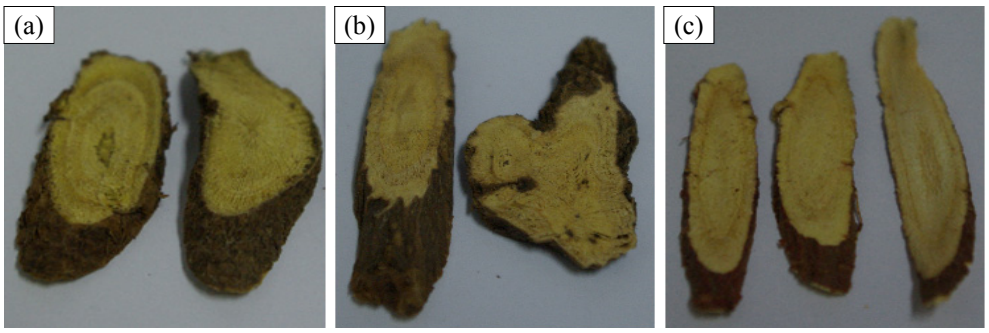

Figure 1. Representative appearances of cut roots of *Glycyrrhiza* species: (**a**) *G. glabra*, (**b**) *G. inflata*, (**c**) *G. uralensis*.

ing and differentiation of *Glycyrrhiza* species based on data from ^{1}H NMR coupled with PLS-DA.

Experimental

Sample preparation. Root samples of *Glycyrrhiza* species. In 2006, 30 samples (50 - 70 slices in each sample) of 3 types of roots of *Glycyrrhiza* species (*G. glabra* Linne, *G. inflata* Batalin, and *G. uralensis* Fisher) were collected from local markets in China and Uzbekistan, and they were identified based on the shape and color of cut surfaces by a veteran expert in this field. The voucher specimens were deposited in the College of Pharmacy of Chung-Ang University. To investigate the major metabolomic differences of each species, identified samples of each species were pooled and analyzed with five replications to reflect the commercial situation that the *Glycyrrhiza* species were circulated in slice form rather than whole non-cut root and to reduce the effects of variability of individual slices. Those were ground into powder by a blender (Hanil, Seoul, Korea) and passed through a 20-mesh (0.9-mm) sieve. The samples were stored at −70 °C until further analysis. For NMR analysis, 100 mg of ground-root samples of *Glycyrrhiza* species were transferred into a centrifuge tube. Five milliliters of a 50% water-methanol mixture were added to the ground-root samples of *Glycyrrhiza* species in the tube and vortexed for 1 min and sonicated for 1 min. The materials were then centrifuged at 500 × *g* for 10 min. The aqueous fractions (50% methanol extracts) were transferred separately into a 50-mL round-bottomed flask and dried in a rotary vacuum evaporator. KH_2PO_4 was added to D_2O to a concentration of 0.1 M as a buffering agent and the pH of the D_2O was adjusted to 6.0 by the addition of 1 N NaOD solution. One milliliter of D_2O at the adjusted pH was added to the 50-mL round-bottomed flask to dissolve the dried aqueous extract. The dissolved solutions were transferred to an NMR tube (Norell, Landisville, NJ, USA) for NMR measurements. Each experiment was performed five times.

Solvents and chemicals. First-grade methanol, and D_2O [99.9%, containing 0.05% 3-(trimethylsilyl)-propionic-2,2,3,3-d_4 acid sodium salt (TSP) as an internal standard] were purchased from Sigma (St. Louis, MO, USA), and NaOD was purchased from Cortec (Paris, France).

NMR measurements. All spectra were obtained by a NMR spectrometer (Avance 600 FT-NMR, Bruker, Germany) operating at a proton NMR frequency of 600.13 MHz. For each sample, 128 scans were recorded with the following parameters: 0.15 Hz/point, pulse width of 4.0 μs (30°), spectral width of 18.94 ppm (11,363.64 Hz), and a relaxation delay of 1.0 s. TSP (0.05%, w/v) was used as the internal standard for D_2O.

Data analysis. The spectral ^{1}H NMR region from δ = 0.52 to δ = 10.00 was segmented into regions with widths of 0.04 ppm (giving 237 integrated regions) using AMIX software (version 3.7, Biospin, Bruker). The region from δ = 4.60 to δ = 4.90 was excluded from the analysis due to the presence of the signal from residual water in aqueous extracts. All spectral data were mean centered and scaled to unit variance. PLS-DA was performed with SIMCA-P software (version 12.0, Umetrics, Umeå, Sweden). The statistical significances of the mean intensities of selected peaks of lactic acid, alanine, proline, glucose, 4-hydroxyphenylacetic acid, arginine, sucrose and malic acid were tested by Tukey's multiple t-test of one-way ANOVA using SPSS software (version 12.0, SPSS, Chicago, IL).

Results and Discussion

^{1}H NMR analysis and assignment of nonvolatile metabolites in samples of *Glycyrrhiza* species. Figure 2 shows a representative NMR spectrum of the 50% methanol extracts, and the following peaks were assigned based on comparisons with the chemical shifts of standard compounds and by using Chenomx NMR suite software (version 5.1, Chenomx, Edmonton, Canada): lactic acid at δ 1.30 (d, J = 6.7 Hz); alanine at δ 1.46 (d, J = 7.3 Hz); arginine at δ 1.66 (m), and 1.90 (m); proline at δ 1.98 (m), 2.34 (m), 3.34 (m), and 4.10 (m); malic acid at δ 2.42 (dd, J = 15.6, 9.6 Hz), 2.66 (dd, J = 15.6, 3.0 Hz), and 4.30 (dd, J = 9.4, 3.2 Hz); asparagine at δ 2.86 (dd, J = 16.8, 7.7 Hz), 2.94 (dd, J = 17.0, 4.1 Hz), and 4.02 (dd, J = 7.8, 4.4 Hz); choline at δ 3.18 (s); glycine at δ 3.58 (s); glucose at δ 3.21- 3.90 (sugar region), and 5.22 (d, J = 3.7 Hz); sucrose at δ 3.44 - 4.21 (sugar region), and 5.42 (d, J = 3.8 Hz); 4-hydroxyphenylacetic acid at δ 6.82 (m), and 7.14 (m); formic acid at δ 8.46 (s).

PLS-DA. Figure 3 shows that each score plot derived from PLS-DA model was clearly separated according to various *Glycyrrhiza* species using ^{1}H NMR data of 50% methanol extracts. The two PLS components (PLS Component 1 and PLS Component 2 together) accounted for 78.9% of the total variance. It suggested that the metabolic discrimination of various *Glycyrrhiza* species was possible using ^{1}H NMR coupled with PLS-DA.

The PLS-DA model was validated by a permutation method. The idea of this validation is to compare the goodness of fit (R^2Y) and the predictability (Q^2Y) of the original model with the goodness of fit and the predictability of several models based on data where the order of the Y-observations has been ran-

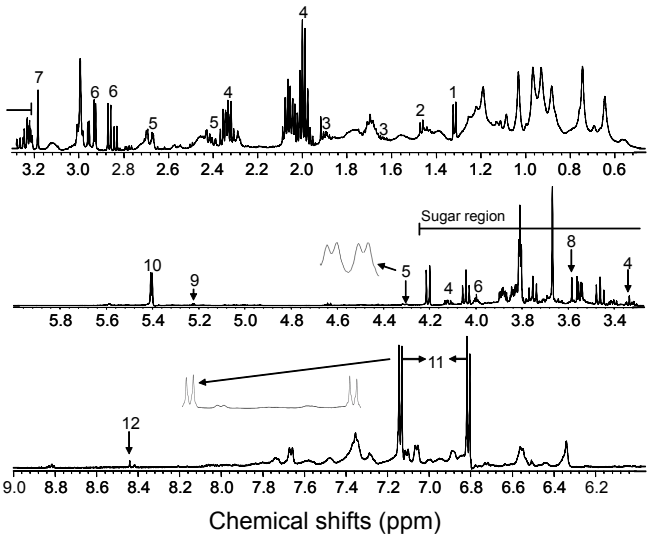

Figure 2. ^{1}H NMR spectrum of the total region of 50% methanol extracts of roots of *Glycyrrhiza glabra*. (1, lactic acid, 2, alanine, 3, arginine, 4, proline, 5, malic acid, 6, asparagine, 7, choline, 8, glycine, 9, glucose, 10, sucrose, 11, 4-hydroxyphenylacetic acid, 12, formic acid)

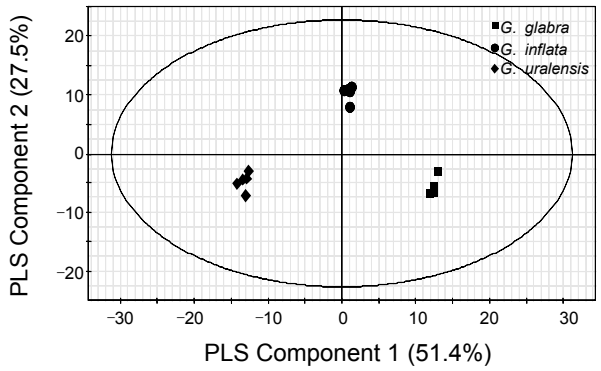

Figure 3. PLS-DA derived score plots obtained using ^1H NMR data of 50% methanol extracts of *Glycyrrhiza* species. ■, *G. glabra*, ●, *G. inflata*, ◆, *G. uralensis*.

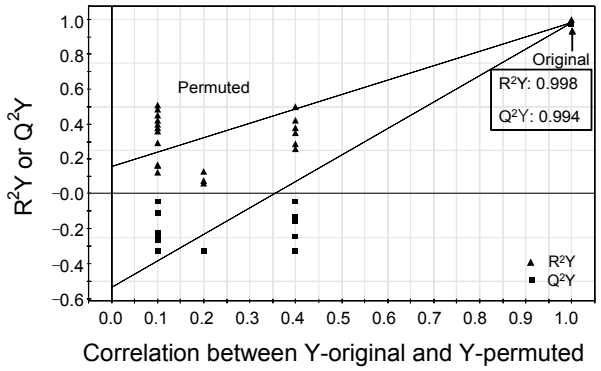

Figure 4. Validation plots of the PLS-DA model using ^1H NMR data of 50% methanol extracts of *Glycyrrhiza* species.

domly permuted, while the X-matrix has been kept intact. The plot shows, for a selected Y-variable, on the vertical axis the values of R^2Y and Q^2Y for the original model (far to the right) and of the Y-permuted models further to the left. The horizontal axis shows the correlation between the permuted Y-vectors and the original Y-vectors for the selected Y. Generally, R^2Y - which describes how well the data in the training set are mathematically reproduced - varies between 0 and 1, where 1 indicates a model with a perfect fit. Q^2Y values of > 0.5 and > 0.9 are considered indicative of good and excellent predictive abilities, respectively.[24] Figure 4 shows the validation plots by the permutation method through 20 applications in which all Q^2Y values of permuted Y vectors were lower than original ones and the regression of Q^2Y lines intersect at below zero. Q^2Y-intercepts below 0.05 indicate valid models.[22] In addition, R^2Y and Q^2Y values were 0.998 and 0.994 respectively in this PLS-DA model, which meant that the PLS-DA model in this study has excellent fitness and predictive abilities.

Variable importance in the projection (VIP) value is an weighted sum of squares of the PLS-DA weights, both with respect to Y as the correlation to all the responses and X as its projection, picking components that play important roles in the separation. It has been indicated that cutting-off for VIP around 0.7 - 0.8 worked well for variable selection even though the variables with larger than 1 were most influential for model.[24-26]

Table 1. The VIP values of the major contributing compounds for the separation in the score plots derived from PLS-DA model

Chemical shifts (ppm)	Compounds	VIP values
1.30	lactic acid	1.31
1.46	alanine	1.22
1.98	proline	1.15
2.94	asparagine	1.12
6.82	4-hydroxyphenylacetic acid	1.11
5.22	glucose	1.10
3.58	glycine	1.02
1.90	arginine	0.99
5.42	sucrose	0.97
2.42	malic acid	0.95
8.46	formic acid	0.74
3.18	choline	0.73

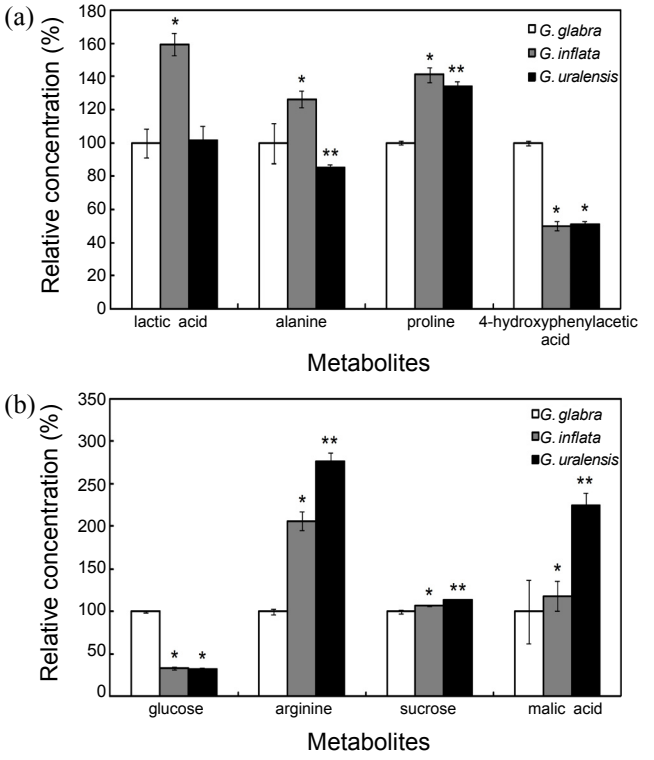

Figure 5. Relative concentrations of lactic acid, alanine, proline and glucose (**a**), and 4-hydroxyphenylacetic acid, arginine, sucrose, and malic acid (**b**) in 50% methanol extracts of *Glycyrrhiza* species. The intensities of each peak corresponding to various compounds were scaled to 100% for those of *G. glabra* samples. The error bars are expressed as the standard deviation. *p < 0.05 *vs G. glabra*, **p < 0.05 *vs G. inflata*

As shown in Table 1, the VIP values of the major contributing compounds for the separation in the score plots derived from PLS-DA were like follows; lactic acid: 1.31, alanine: 1.22, proline: 1.15, asparagine: 1.12, 4-hydroxyphenyl acetic acid: 1.11, glucose: 1.10, glycine: 1.02, arginine: 0.99, sucrose: 0.97, malic

acid: 0.95, formic acid: 0.74, and choline: 0.73. Among those compounds, lactic acid, alanine and proline were the most influential for the separation of each sample in the PLS-DA model in this study.

To obtain clearer information on the relative levels of each compounds from the VIP analysis, we performed ANOVA test as shown in Figure 5. There were no significant differences of the levels of asparagine, glycine, formic acid, and choline in various *Glycyrrhiza* samples (data not shown). However, the levels of glucose, and 4-hydroxyphenylacetic acid were significantly ($p < 0.05$ in all cases) higher in *G. glabra*, those of lactic acid, alanine, and proline were significantly higher in *G. inflata*, and those of arginine, malic acid, and sucrose were significantly higher in *G. uralensis*. Even though the glycyrrhizic acid and liquiritin were not detected in the NMR spectra in this study, the obtained results showed that the metabolomic profiles differ within the species. Further experiment will be performed using LC-MS to detect metabolites such as glycyrrhizic acid and liquiritin and to integrate the data from NMR and LC-MS. It can be suggested that the combined use of NMR and PLS-DA is an efficient technique for the classification of dried roots of *Glycyrrhiza* species, and would be suitable for discriminating samples in commercial applications.

Conclusion

Dried roots of *Glycyrrhiza* species are consumed widely as traditional medicines and food resources in China, Japan, Korea, and their consumption is likely to increase in the future, due to its various beneficial bioactivities. In this study we used a metabolomic approach based on ^{1}H NMR spectroscopic analysis to reveal how the characteristic metabolic profiles of dried roots vary between *Glycyrrhiza* species. The obtained results show that the metabolomic profiles differ within the species, making further investigations of the bioactivities of the different species necessary. It can be suggested that the combined use of NMR and PLS-DA is an efficient technique for the classification of dried roots of *Glycyrrhiza* species, and would be suitable for discriminating samples in commercial applications.

Acknowledgments. This research was supported by the Chung-Ang University Research Scholarship Grants in 2009.

References

1. Mauricio, I.; Francischetti, B.; Monteiro, R. Q.; Guimarães, J. A. *Biochem. Biophys. Res. Commun*. **1997**, *235*, 259.
2. Dong, S.; Inoue, A.; Zhu, Y.; Tanji, M.; Kiyama, R. *Food Chem. Toxicol*. **2007**, *45*, 2470.
3. Kim, I. G.; Kang, S. C.; Kim, K. C.; Choung, E. S.; Zee, O. P. *Environ. Toxicol. Pharmacol*. **2008**, *25*, 75.
4. Won, S. R.; Kim, S. K.; Kim, Y. M.; Lee, P. H.; Ryu, J. H.; Kim, J. W.; Rhee, H. I. *Food Res. Int*. **2007**, *40*, 1046.
5. Schambelan, M. *Steroid* **1994**, *59*, 127.
6. Fiore, C.; Eisenhut, M.; Krausse, R.; Ragazzi, E.; Pellati, D.; Armanini, D.; Bielenberg, J. *Phytother. Res*. **2008**, *22*, 141.
7. Hayashi, H.; Hiraoka, N.; Ikeshiro, Y.; Yamamoto, H. *Plant Sci*. **1996**, *116*, 233.
8. Nomura, T.; Fukai, T.; Akiyama, T. *Pure Appl. Chem*. **2002**, *74*, 1199.
9. Kinoshita, T.; Tamura, Y.; Mizutani, K. *Chem. Pharm. Bull*. **2005**, *53*, 847.
10. Asl, M. N.; Hosseinzadeh, H. *Phytother. Res*. **2008**, *22*, 709.
11. *The United States Pharmacopeia*; The United States Pharmacopeial Convention: Rockville, U. S. A., 2007; p 955.
12. *The Japanese Pharmacopeia*; Yakuji Nippo Ltd.: Tokyo, Japan, 2006; p 932.
13. *The Korean Pharmacopeia*; Shinil Books Company: Seoul, Korea, 2008; p 903.
14. *Pharmacopoeia of the People's Republic of China*; People's Medical Publishing House: Beijing, China, 2005; p 207.
15. Ott, K. H.; Aranibar, N.; Singh, B.; Stockton, G. W. *Phytochemistry* **2003**, *62*, 971.
16. Tweeddale, H.; Notley-McRobb, L.; Ferenci, T. *J. Bacteriol*. **1998**, *180*, 5109.
17. Consonni, R.; Cagliani, L. R.; Benevelli, F.; Spraul, M.; Humpfer, E.; Stocchero, M. *Anal. Chim. Acta* **2008**, *611*, 31.
18. Tarachiwin, L.; Ute, K.; Kobayashi, A.; Fukusaki, E. *J. Agric. Food Chem*. **2007**, *55*, 9330.
19. Tarachiwin, L.; Katoh, A.; Ute, K.; Fukusaki, E. *J. Pharm. Biomed. Anal*. **2008**, *48*, 42.
20. Whang, W. K.; Lee, M. W.; Choi, H. K. *Bull. Korean Chem. Soc*. **2007**, *28*, 557.
21. Barker, M.; Rayens, W. J. *Chemometr*. **2003**, *17*, 166.
22. Wishart, D. S. *Trends Food Sci. Tech*. **2008**, *19*, 482.
23. Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; Van Velzen, E. J. J.; Van Duijnhoven, J. P. M.; Van Dorsten, F. A. *Metabolomics* **2008**, *4*, 81.
24. Eriksson, L.; Johansson, E.; Kettaneh-Wold, N.; Trygg, J.; Wikstrom, C.; Wold, S. *Multi- and Megavariate Data Analysis Part I: Basic Principles and Applications*; Umetrics, Inc.: Umeå, Sweden, 2006; p 86.
25. Lu, X.; Zhao, X.; Bai, C.; Zhao, C.; Guo, L.; Xu, G. *J. Chromatogr. B* **2008**, *866*, 64.
26. Woo, H. M.; Kim, K. M.; Choi, M. H.; Jung, B. H.; Lee, J. G.; Kong, G.; Nam, S. J.; Kim, S. H.; Bai, S. W.; Chung, B. C. *Clin. Chim. Acta* **2009**, *400*, 63.