Biomolecules & Therapeutics, **17(2)**, 133-137 (2009)
www.biomolther.org

ISSN: 1976-9148(print)/2005-4483(online)
DOI: 10.4062/biomolther.2009.17.2.133

Fingerprinting Differentiation of *Astragalus membranaceus* Roots According to Ages Using ^{1}H-NMR Spectroscopy and Multivariate Statistical Analysis

Yoo-Soo Shin[1],, Kyong-Hwan Bang[1],**, Dong-Su In[2], Jung-Sook Sung[3], Seon-Young Kim[4], Bon Cho Ku[5], Suk-Weon Kim[5], Dongho Lee[6], and Hyung-Kyoon Choi[7],***

[1]National Institute of Horticultural & Herbal Science, Rural Development Administration, Eumseong 369-873, [2]Kongju National University, Gongju 314-701, [3]National Academy of Agricultural Science, Rural Development Administration, Suwon 441-707, [4]College of Agriculture & Life Science, Seoul National University, Seoul 151-742, [5]Biological Resource Center, KRIBB, Daejeon 305-806, [6]School of Life Sciences and Biotechnology, Korea University, Seoul 136-701, [7]College of Pharmacy, Chung-Ang University, Seoul 156-756, Republic of Korea

(Received October 13, 2008; Revised January 8, 2009; Accepted January 10, 2009)

Abstract − The root of *Astragalus membranaceus* is a traditional folk medicine that has been used for many therapeutic purposes in Asia. It reportedly acts as an immunostimulant, tonic, hepatoprotective, diuretic, antidiabetic, analgesic, expectorant, sedative, and anticancer drug. In this study, metabolomic profiling was applied to the roots of *A. membranaceus* of different ages using NMR coupled with two multivariate statistical analysis methods: such as principal components analysis (PCA) and canonical discriminant analysis (CDA). This allowed various metabolites to be assigned in NMR spectra, including γ-aminobutyric acid (GABA), aspartic acid, succinic acid, glutamic acid, glutamine, N-acetyl aspartic acid, acetic acid, arginine, alanine, threonine, lactic acid, and valine. The score plot from PCA and also CDA allowed a clear separation between samples according to age.

Keywords: Fingerprinting analysis, *Astragalus membranaceus*, ^{1}H NMR spectroscopy, Principal component analysis, Canonical discriminant analysis

INTRODUCTION

The root of *Astragalus membranaceus*, which belongs to the Leguminosae family, is a traditional folk medicine that has been used for many therapeutic purposes in the Asia including in China, Japan, Korea, Mongolia and Siberia (Chang *et al.,* 1987). It has been shown to be an immunostimulant, tonic, hepatoprotective, diuretic, antidiabetic, analgesic, expectorant, sedative, and anticancer drug (Wu *et al.,* 2005). The main constituents of *A. membranaceus* root are polysaccharides, saponins, flavonoids, amino acids, and γ-aminobutyric acid (GABA) (Ma *et al.,* 2002; Shao *et al.,* 2004). Astragaloside IV is used as a marker for quality control and is reportedly an important component responsible for the proliferation of T and B lymphocytes and antibody production both in vivo and in vitro. Also, isoflavonoids are considered useful as marker compounds for standardizing *A. membranaceus* (Wu *et al.,* 2005).

However, *A. membranaceus* roots of different ages cannot be reliably discriminated using classical methods such as capillary electrophoresis and near-infrared spectroscopy. Ma *et al.* reported that the total isoflavonoids content of *A. membranaceus* roots cultivated in China was higher in 3-year-old samples than in 4-year-old samples, whereas the average content of astragaloside I was higher in the older samples (Ma *et al.,* 2002). However, these results were obtained by targeted analysis using high performance liquid chromatography, and the reliable differentiation of *A. membranaceus* roots requires a systemic method that includes consideration of a variety of metabolites. The term "metabolome" has been used to describe the observable chemical profile or fingerprint of the metabolites in whole tissues (Ott *et al.,* 2003). To obtain the most complete metabolomic profile, it is desirable to use a wide spectrum of analytical tech-

*Corresponding author
Tel: +82-2-820-5605 Fax: +82-2-812-3921
E-mail: hykychoi@cau.ac.kr
**These authors equally contributed to this work.

niques that are rapid, reproducible and temporally stable, and require only a very simple sample preparation. NMR is a technique that meets these requirements (Pauli, 2001), and has recently been combined with principal components analysis (PCA) for the metabolomic profiling of several types of plants and plant functional genomics (Belton *et al.,* 1998; Fiehn *et al.,* 2000; Raamsdonk *et al.,* 2001; Choi *et al.,* 2004; Krishnan *et al.,* 2004). The combined NMR and PCA method has also been used as a fingerprinting tool for the interpretation and quality assessment of both industrial and natural products (Bailey *et al.,* 2000; Sobolev *et al.,* 2003; Le Gall *et al.,* 2004). Moreover, canonical discriminant analysis (CDA) has been used to classify various natural resources such as honey, ginseng, and olive oils (Nozal *et al.,* 2005; Shin *et al.,* 2007; Petrakis *et al.,* 2008). Here we report the fingerprinting analysis of *A. membranaceus* roots of various ages using a ^{1}H NMR spectroscopic method coupled with PCA and CDA.

MATERIALS AND METHODS

Plant materials

The roots of *A. membranaceus* at different ages (1-, 2-, 3- and 4-year-old), which were harvested in Jeongseon, Gangwon-do, Korea, were purchased from Jeongseon Medicine Herbal Association at March of 2006 for this experiment. The root samples were confirmed by Dr. Jung-Sook Sung of National Institute of Horticultural & Herbal Science, and the voucher specimens were deposited at the herbarium of College of Pharmacy, Chung-Ang University.

Reagents

Methanol-d4 (99.8%) was purchased from Cambridge Isotope Laboratories (Miami, FL, USA) and D_2O (99.0%) was purchased from Sigma (St. Louis, MO, USA).

Extraction

To investigate the individual variability of root samples, 20 different root samples per each age of *A. membranaceus* were freeze-dried and grinded to a fine powder in a pestle and mortar, respectively, and then stored at -80^oC until analyzed. Freeze-dried materials (50 mg) were weighed into an autoclaved 1.5 ml eppendorf tube. D_2O:CD_3OD (1 ml, 80:20) containing 0.05 % (w/v) TSP-*d*4 (sodium salt of trimethysilypropionic acid) added to each sample. The contents of the tube were mixed thoroughly and sonicated at 40^oC for 10 min. After cooling, the samples were centrifuged at 1,000×g in a micro-centrifuge for 5 min. The supernatant was transferred to a 5-mm NMR tube. For verifying recovery of metabolites the first extracts were compared with the second and the third extracts using the same amount of solvents. The first extracts showed more than 92.9 (± 2.6%) of recovery of metabolites based on total ^{1}H-NMR intensity.

NMR measurements and data analysis

^{1}H NMR spectra were recorded at 25^oC on a 500 MHz Varian Unity 500 spectrometer operating at a ^{1}H NMR frequency of 500.202 MHz. Prior to Fourier transformation, exponential multiplication (LB with 0.5 Hz) and linear prediction were performed to each FID and each spectrum was auto-phased and spectrum averaging baseline corrected (box width: 16, noise factor: 5). TSP was used as a reference for chemical shift (0.00 ppm). The processing of each spectrum was carried out using ACD NMR processor (ver. 7.0, Advanced Chemistry Development, Inc., Canada) and saved as ASCII files. Spectral intensities were scaled to TSP-d_4, and the spectral region δ=0.52-10.00 was segmented into regions of 0.04 ppm width. Peak alignment was performed using SpecAlign (1.2.1). The four regions $\delta-0.52$ -δ0.7 and δ4.80 - δ5.10 were excluded from the analysis because of the residual signal of TSP and water, respectively. The remaining regions were normalized to the whole spectrum for PCA. All spectral data were mean centered with scaled to unit variance, then analyzed by PCA based on correlation matrix. PCA and CDA were performed with SPSS program (Version 3.13, Kovach Computing Services, United Kingdom).

RESULTS AND DISCUSSION

As shown in the representative ^{1}H NMR spectrum of D_2O:CD_3OD (80:20) extracts of *A. membranaceus* root in Fig. 1, various peaks were detected in the aliphatic, carbohydrate, and phenolic regions of ^{1}H NMR spectra without any chromatographic separation. Mono- or disaccharides such as sucrose, maltose, and glucose were identified in the carbohydrate region (δ=3.1-6.0) of the ^{1}H NMR spectra. The signals at δ=5.42, δ=5.35, δ=5.20, δ=4.61, δ=4.17, and δ=4.13 were assigned as anomeric protons of sucrose, maltose, α-glucose, β-glucose, and fructofuran moiety of sucrose, respectively, in accordance with spectral data in the literature (Choi *et al.,* 2004). The signals of several amino acids observed in this region were assigned to N-acetyl aspartic acid (δ=4.30), aspartic acid (δ=3.97), glutamic acid (δ=3.74), glutamine (δ=3.72), and taurine (δ=3.51) (Fig. 1B). In the phenolic region (δ=6.0-10.0), fumaric acid was detected at δ=6.56 at a low level (Fig. 1C). Several compounds in the aliphatic region including GABA, aspartic acid, succinic acid, glutamic acid, glutamine, N-acetyl aspartic acid, acetic acid, arginine, alanine, threonine, and va-

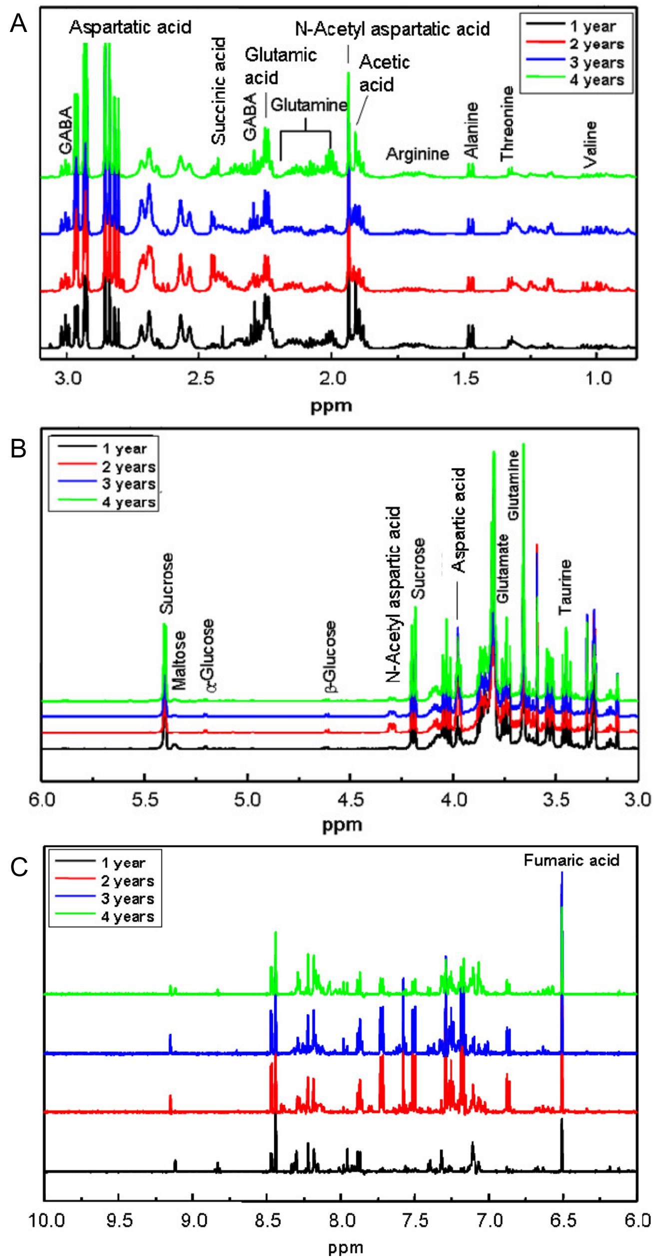

Fig. 1. ^{1}H NMR spectra of the aliphatic region (0.7-3.2 ppm, A), sugar region (3.2-6.0 ppm, B), and aromatic region (6.0-10.0 ppm, C) of D_2O:CD_3OD (80:20) extracts of *A. membranaceus* roots of different ages.

line were also identified with signals at δ=3.01 (*t*), δ= 2.82-2.96 (*dd*), δ=2.46 (*s*), δ=2.28 (*m*), δ=1.94-2.24 (*m*), δ= 1.93 (*s*), δ=1.91 (*s*), δ=1.66-1.74 (*m*), δ=1.46 (*d*), δ=1.32 (*d*), and δ=1.04 (*m*), respectively, based on spectral data in the literature (Le Gall *et al.,* 2004).

To ensure the objective interpretation of the results, the samples were analyzed using PCA and CDA. PCA is an unsupervised clustering method that does not require any

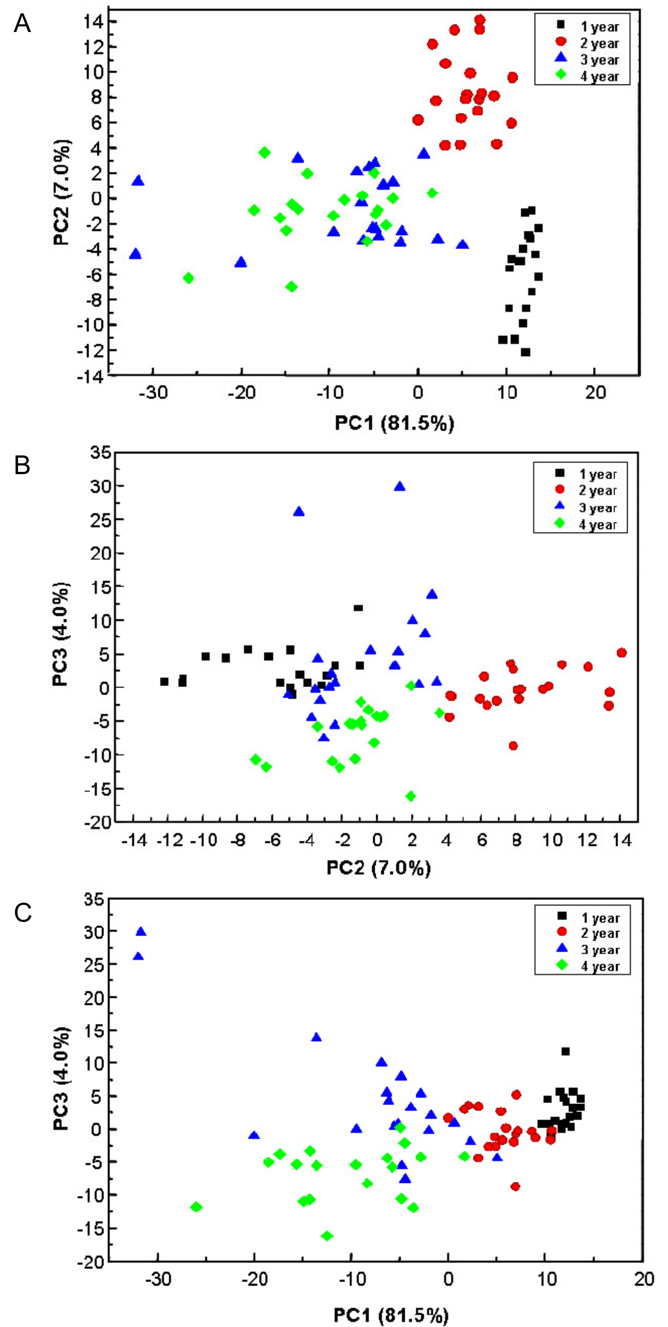

Fig. 2. PCA-derived score plots of all regions of the 1H NMR spectra of D_2O:CD_3OD (80:20) extracts of the *A. membranaceus* roots of different ages. (A) PC1 vs PC2, (B) PC2 vs PC3, (C) PC1 vs PC3.

knowledge of the data set and reduces the dimensionality of multivariate data while preserving most of the variance therein (Eriksson *et al.,* 2001). For reliable and easy comparison of the metabolomic profiling, a large ^{1}H NMR data set obtained in the presence of numerous metabolites can be reduced to PC1, PC2, or PC3 using PCA. Application of

PCA to the ^{1}H NMR spectra resulted in a clear separation between 1-year and 2-year samples, but the separation between 3-year and 4-year samples was not clear in score plots of PC1 versus PC2, which accounted for 88.5% of the total variance. However, the 3-year and 4-year samples were separated more clearly in the score plots of PC1 versus PC3, accounting for 85.5% of the total variance (Fig. 2). This indicates that PCA is an ideal method for differentiating *A. membranaceus* roots of different ages based on NMR fingerprinting.

For a data set comprising *n* observations on each of *N* variables, CDA takes the *N* variables in the original data set and creates a new set of variables. Each new variable is a linear function (called the canonical discriminant function) of the original variables, and the discriminant function can be used to calculate a set of *n* discriminant scores for each discriminant function. For any particular discriminant function, the correlations between the *n* observations for each of the *N* original variables and the *n* discriminant scores (often termed the canonical structure) can be used to interpret the features of the data set described by a particular discriminant function (Mattew *et al.,* 1994). The essential difference between CDA and PCA is that PCA - unless it is specifically manipulated to do so - does not partition the data matrix to account for the design structure of a particular experiment. Instead, PCA maximizes the proportion of the total variance of the data set expressed by successive principal components. That is, each set of n principal component scores exhibits the maximum possible variation among the *n* observations, for a linear function of the *N* variables in the data set. As a supervised method, CDA identified the 1-, 2-, 3-, and 4-year-old *A. membranaceus* roots with 100% discrimination accuracy based on NMR data (Table I).

Table I. Posterior accuracy of classification by CDA of NMR data of D_2O:CD_3OD (80:20) extracts of *A. membranaceus* roots showing that 100% of cross-validation grouped cases correctly classified (CV: cross validation)

Group	Predicted group membership				
	1 year	2 years	3 years	4 years	Total
CV					
Frequency					
1 year	20	0	0	0	20
2 year	0	20	0	0	20
3 year	0	0	20	0	20
4 year	0	0	0	20	20
Percentage					
1 year	100	0	0	0	100
2 year	0	100	0	0	100
3 year	0	0	100	0	100
4 year	0	0	0	100	100

The presented results indicate that combining ^{1}H NMR with PCA and CDA is a very promising tool for the authentication and quality control of *A. membranaceus* roots of different ages. We plan to investigate the specific metabolite markers of *A. membranaceus* roots at different ages by applying loading plot analysis to NMR data or using GC-MS and LC-MS techniques coupled with multivariate statistical analysis methods.

REFERENCES

Bailey, N. J. C., Cooper, P., Hadfield, S. T., Lenz, E. M., Lindon, J. C., Nicholson, J. K., Stanley, P. D., Wilson, I. D., Wright, B. and Taylor, S. D. (2000). Application of directly coupled HPLC-NMR-MS/MS to the identification of metabolites of 5-trifluoromethylpyridone(2-hydroxy-5-trifluoromethylpyridin e) in hydroponically grown plants. *J. Agric. Food Chem.* **48**, 42-46.

Belton, P. S., Colquhoun, I. J. , Kemsley, E. K., Delgadillo, I., Roma, P., Dennis, M. J., Sharman, M., Holmes, E., Nicholson, J. K. and Spraul, M. (1998). Application of chemometircs to the 1H NMR spectra of apple juices: discrimination between apple varieties. *Food Chem.* **61**, 207-213.

Chang, H. and But, P. (1987). *Pharmacology and applications of Chinese material medica*. World Scientific., Singapore.

Choi, H. K., Choi, Y. H., Verberne, M., Lefeber, A. W. M., Erkelens, C. and Verpoorte, R. (2004). Metabolic fingerprinting of wild type and transgenic tobacco plants by ^{1}H NMR and multivariate analysis technique. *Phytochemistry* **65**, 857-864.

Eriksson, L., Johansson, E., Kettaneh-Wold, N. and Wold, S. (2001). *Multi- and Megavariate data analysis. Principles and applications.* Umetrics Academy., Umea, Sweden.

Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R. N. and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. *Nature Biotechnol.* **18**, 1157-1161.

Krishnan, P., Kruger, N. J. and Ratcliffe, R. G. (2004). Metabolite fingerprinting and profiling in plants using NMR. *J. Exp. Bot.* **56**, 255-265.

Le Gall, G., Coloquhoun, I. J. and Defernez, M. (2004). Metabolite profiling using ^{1}H NMR spectroscopy for quality assessment of green tea, *Camellia sinensis* (L.). *J. Agric. Food Chem.* **52**, 692-700.

Ma, X. Q., Shi, Q., Duan, J. A., Dong, T. T. X. and Tsim, K. W. K. (2002). Chemical analysis of radix astragali (Huangqi) in china: a comparison with its adulterants and seasonal varations. *J. Agric. Food Chem.* **50**, 4861-4866.

Matthew, C., Lawoko, C. R. O, Korte, C. J. and Smith, D. (1994). Application of canonical discriminant analysis, principal component analysis, and canonical correlation analysis as tools for evaluating differences in pasture botanical composition. *New Zealand J. Agric. Res.* **37**, 509-520.

Nozal, M. J., Bernal, J. L., Toribio, L., Alamo, M. and Diego, J. C. (2005). The use of carbohydrate profiles and chemometrics in the characterization of natural honeys of identical geographical origin. *J. Agric. Food Chem.* **53**, 3095-3100.

Ott, K. H., Aranibar, N., Singh, B. and Stockton, Q. W. (2003).

Metabolomic classifies pathway affected by bioactive compounds. Artificial neural network classification of NMR spectra of plant extracts. *Phytochemistry* **62**, 971-985.

Pauli, G. F. (2001). qNMR-a versatile concept for the validation of natural product reference compounds. *Phytochem. Anal.* **12**, 28-42.

Petrakis, P. V., Agiomyrgianaki, A., Christophoridou, S., Spyros, A. and Dais, P. (2008). Geographical characterization of greek virgin olive oils (cv. Koroneiki) using ^{1}H and ^{31}P NMR fingerprinting with canonical discriminant analysis and classification binary trees. *J. Agric. Food Chem.* **56**, 3200-3207.

Raamsonk, L. M., Teusink, B. and Broadhurst, D. (2001). A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnol.* **19**, 45-50.

Shao, B. M., Xu, W., Dai, H., Yu, P., Li, Z. and Gao, X. M. (2004). A study on the immune receptors for polysaccharides from the roots of *Astragalus membranaceus*, a chinese medicinal herb. *Biochem. Biophys. Res. Comm.* **320**, 1103-1111.

Shin, Y. S., Bang, K. H., In, D. S., Kim, O. T., Hyun, D. Y., Ahn, I. O., Ku, B. C., Kim, S. W., Seong, N. S., Cha, S. W., Lee, D. and Choi, H. K. (2007). Fingerprinting analysis of fresh ginseng roots of different ages using ^{1}H-NMR spectroscopy and principal components analysis. *Arch. Pharm. Res.* **12**, 1625-1628.

Sobolev, L. J., Last, R. L. and Fernie, A. R. (2003). Proton high-field NMR study of tomato juice. *Mag. Res. Chem.* **41**, 237-245.

Wu, T., Annie Bligh, S. W., Gu, L. H., Wang, Z. T., Liu, H. P., Cheng, X. M., Branford-White, C. J. and Hu, Z.B. (2005). Simultaneous determination of six isoflavonoids in commercial radix Astragali by HPLC-UV. *Fitoterapia* **76**, 157-165.