Clustering Analysis of Keishibukuryogan Formulas by Use of Self-Organizing Maps

Hisanori SATOMI,*[,a,b] Yoshikazu MORI,[c] Bunsho MAKINO,[d] Youichiro NAKAI,[d] Shuichi TAKEDA,[b] Masaki ABURADA,[e] and Ken-ichi MIYAMOTO[a]

[a] *Department of Clinical Pharmacy, Graduate School of Natural Science and Technology, Kanazawa University; 13–1 Takaramachi, Kanazawa, Ishikawa 920–8641, Japan:* [b] *Analytical Technology Center, Tsumura & Co.;* [c] *Production Division, Tsumura & Co.;* [d] *Botanical Raw Materials Division, Tsumura & Co.; 3686 Yoshiwara, Ami-machi, Inashiki-gun, Ibaraki 300–1192, Japan: and* [e] *Faculty of Pharmacy, Musashino University; 1–1–20 Shinmachi, Nishi-Tokyo, Tokyo 202–8585, Japan.* Received June 15, 2010; accepted August 26, 2010

Kampo medicines, traditional herbal medicines in Japan, are comprised of multiple botanical raw materials that contain a number of pharmacologically active substances. Conventionally, the quality control of kampo medicines has been performed by analyzing the contents of two or three representative components. However, it is not sufficient to check quality only with a limited number of specific components. We performed HPLC of 287 lots of keishibukuryogan formulas, calculated the areas of 11 components on chromatograms as feature values and made a cluster analysis using self-organizing maps (SOMs). We verified the precision (repeatability and intermediate precision) of clustering results when using the same samples and successfully established an clustering method using SOMs that is capable of precisely clustering differences in HPLC-fingerprints among pharmaceutical manufacturers, differences in HPLC-fingerprints due to the combination ratios of botanical raw materials, and differences in HPLC-fingerprints due to changes in component contents caused by time-course deterioration. Consequently, we could confirm that this method is useful for controlling the quality of multiple component drugs and analyzing quality differences.

Key words self-organizing map; high-performance liquid chromatography; fingerprint; kampo medicine; keishibukuryogan; quality control

Kampo medicines consisting of multiple botanical raw materials are the drugs that contain a number of pharmacologically active substances. Conventionally, quality control has been performed by analyzing the contents of two or three representative components that are contained in botanical raw materials.

The keishibukuryogan formula, whose botanical raw materials are Cinnamomi Cortex, Poria, Moutan Cortex, Persicae Semen, and Paeoniae Radix, possesses pharmacological actions, *e.g.*, blood flow-increasing activity,[1,2] anti-inflammatory activity,[3] and estrogen-like activity,[4] and has been reported to be effective for sorts of disorders, *e.g.*, endometritis,[5,6] menstrual irregularity,[5] menopausal disorders,[7,8] and rheumatoid arthritis.[9]

The keishibukuryogan extract listed on the Japanese Pharmacopoeia[10] is specified only for the contents of its three components—cinnamic acid derived from Cinnamomi Cortex, paeoniflorin derived from Moutan Cortex and Paeoniae Radix, and amygdalin derived from Persicae Semen. However, keishibukuryogan contains components apart from the above three components, *e.g.*, cinnamic aldehyde, benzoic acid, and coumarin derived from Cinnamomi Cortex, albiflorin and pentagalloylglucose derived from Paeoniae Radix, paeonol derived from Moutan Cortex, as well as oxypaeoniflorin and benzoylpaeoniflorin derived from Moutan Cortex/Paeoniae Radix. Among them, not a few components are involved in the pharmacological actions of the drug. Furthermore, many researches to elucidate the pharmacological actions of keishibukuryogan have been made strenuously.[11,12] However, only a part of the pharmacologically active components has been elucidated.

Nowadays when analytical technologies are advanced, it is not sufficient to check quality only with a limited number of specific components. The construction of techniques to evaluate multiple component drugs is being required. In fact, the guidelines on botanical drug products that were issued by the Food and Drug Administration (FDA)[13] have specified that fingerprints using spectroengineering procedures to evaluate quality should be provided. In general, multivariate analysis has frequently been used to evaluate chromatography-fingerprints of multiple component drugs. Applied procedures, *e.g.*, principle component analysis (PCA),[14—17] hierarchical clustering method (HCM),[18] multivariate curve resolution-alternating least squares analysis (MCR-ALS),[19] correlation optimized warping (COW),[20] and Mahalanobis–Taguchi system method (MTS),[21] have been reported.

We performed this study with an objective to perform the clustering of pharmaceutical manufacturers and botanical raw materials by using HPLC-fingerprints of keishibukuryogan formulas.

HPLC analysis was made with the samples of the keishibukuryogan formulas (287 lots) principally consisting of TJ-25 (keishibukuryogan extract granules manufactured by Tsumura & Co.). The peak areas of the above 11 major components (Table 1) obtained from chromatograms were used as feature values, and self-organizing maps (SOMs)—a class of the neural network—were applied to make a cluster analysis.

Analytical Conditions for HPLC Solvent systems, solvent ratios, and acid concentrations were compared to examine conditions for extraction. Consequently, a mixture of methanol and 0.1% phosphoric acid (3 : 1) was selected as an extracting solvent that is capable of efficaciously recovering multiple components and ensuring the stability of components in the extracting solvent. Regarding analytical conditions for HPLC, furthermore, solvent systems, solvent ratios,

* To whom correspondence should be addressed. e-mail: satomi_hisanori@mail.tsumura.co.jp

Table 1. Components Analyzed by High-Performance Liquid Chromatography and with Self-Organizing Maps

Peak no.	Substance	Origin
1	Oxypaeoniflorin	Moutan Cortex/Paeoniae Radix
2	Amygdalin	Persicae Semen
3	Albiflorin	Paeoniae Radix
4	Paeoniflorin	Moutan Cortex/Paeoniae Radix
5	Pentagalloylglucose	Paeoniae Radix
6	Benzoic acid	Cinnamomi Cortex
7	Coumarin	Cinnamomi Cortex
8	Cinnamic acid	Cinnamomi Cortex
9	Cinnamaldehyde	Cinnamomi Cortex
10	Benzoylpaeoniflorin	Moutan Cortex/Paeoniae Radix
11	Paeonol	Moutan Cortex

the presence or absence of adding phosphoric acid, and gradient conditions were examined to find HPLC conditions that excel in resolution among multiple components and in their precision and linearity (Fig. 1).

On the UV spectra of the 11 major components, many components exhibited a maximum at not more than 200 nm, while other components also exhibited end absorption at the wavelength range between 200 nm and 210 nm (Fig. 2). Therefore, the use of wavelengths between 200 nm and 210 nm was considered to allow the encompassing detection of all peak components. The wavelength of 210 nm was employed in consideration of increases in baseline associated with increased concentrations of phosphoric acid in the mobile phase and of decreases in detector's precision at the wavelength of about 200 nm.

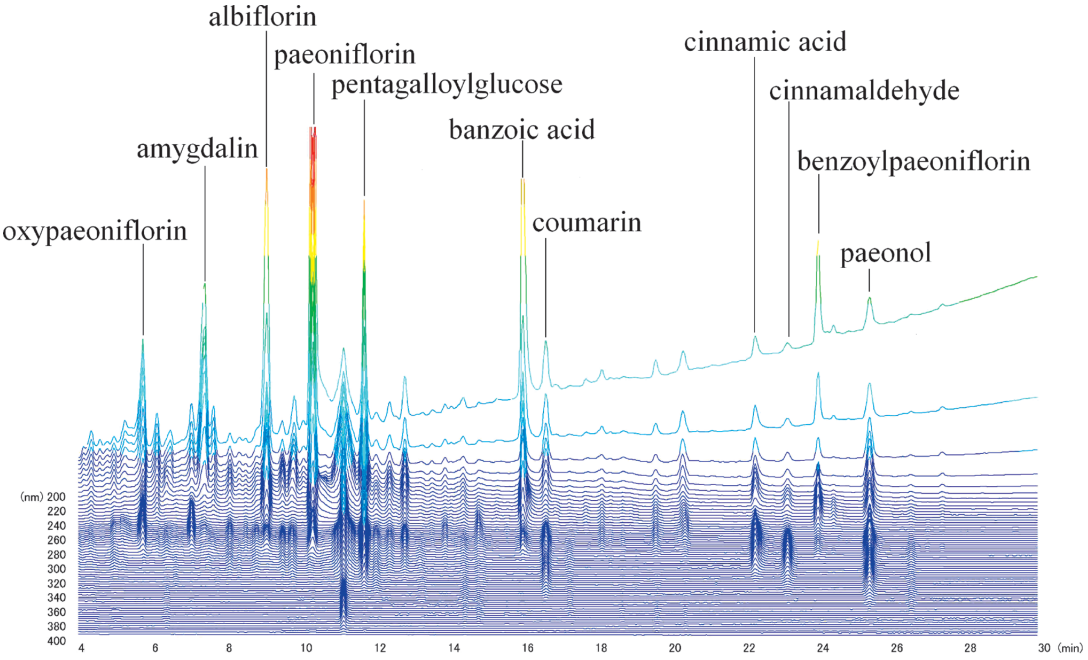

Fig. 1. Typical 3D-HPLC Chromatogram of the Keishibukuryogan Formula

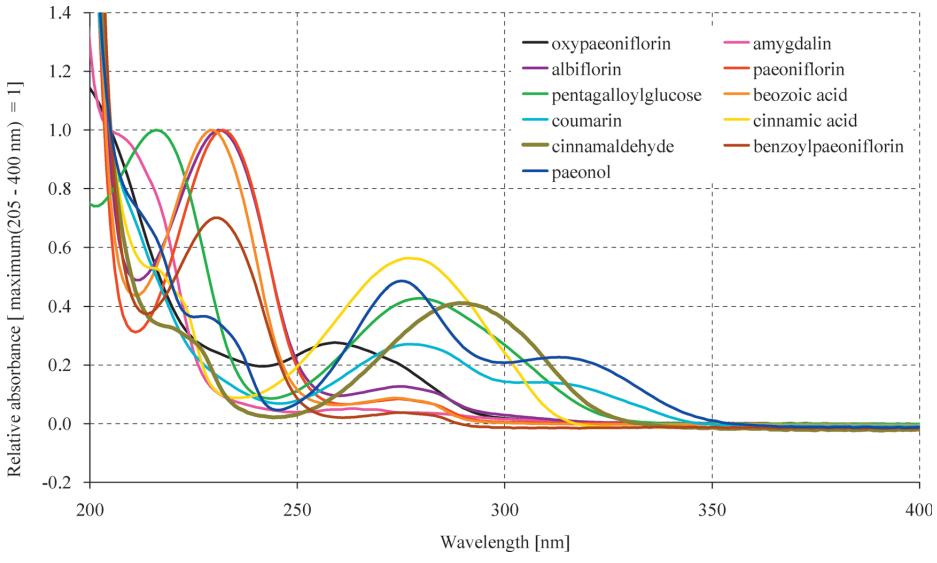

Fig. 2. UV Spectra of the Components in the Keishibukuryogan Formula

Table 2. Manufacturers of the Keishibukuryogan Formulas and Numbers of Lots

Product	Number of lots
TJ-25	267
Company A	2
Company B	2
Company C	1
Company D	2
Company E	2
Company F	6
Company G	1
Company H	1
Company I	3

Samples for HPLC Analysis A total of 287 lots, 267 lots of TJ-25 and 20 lots manufactured by other nine drug manufacturers, were used as samples (Table 2).

For the purpose of verifying precision, furthermore, one given lot among the TJ-25 lots was additionally analyzed 21 times to verify intermediate precision and 8 times to verify repeatability. A total of 316 analytical data were obtained.

The content of the keishibukuryogan extract differs among the formulas manufactured by different drug manufacturers. Therefore, samples were prepared in such a manner to contain a given amount (about 115 mg) of the extract based on the content of the keishibukuryogan extract contained in each manufacturer's formulas.

Clustering Using SOMs Multivariate analyses, which evaluate HPLC-fingerprints and express them visually or numerically, involve sorts of procedures. However, each of them has drawbacks and advantages. In this study, SOMs were used as a clustering procedure because not less than 300 records and 11 feature values were handled and there were numerous feature classifications that should be clustered.

In consideration of the number of records, 25 (5×5) classes of SOMs were used to cluster 316 records and 11 feature values which had been obtained.

Furthermore, an add-in software of Excel, Neural Network Predict version 3.11 (Neuralware, Carnegie, PA, U.S.A.), was used to make a cluster analysis using SOMs.

Results and Discussion

Results of the clustering of other manufacturers' keishibukuryogan formulas and records to verify precision are shown in Table 3.

Precision All records of precision, 21 records of intermediate precision, and 8 records of repeatability were assigned to the cluster (3, 3). Therefore, we could verify the precision of analytical conditions for HPLC and of clustering results when using the same samples.

Differences in HPLC-Fingerprints among Drug Manufacturers The results of the TJ-25 and other manufacturers' keishibukuryogan formula clustering are shown in Fig. 3. Records of TJ-25 and other manufacturers' keishibukuryogan formulas are present concurrently in cluster (1, 4), cluster (2, 4), and cluster (5, 4). However, TJ-25 and other manufacturers' keishibukuryogan formulas were classified nearly completely.

Regarding two [cluster (1, 5) and cluster (2, 5)] of 3 clusters to which only other manufacturers' keishibukuryogan formulas had been assigned, Euclidean distance among adjacent clusters was not less than 1.5 (Fig. 4). Therefore, other manufacturers' keishibukuryogan formulas assigned to these

Table 3. Results of the Clustering of Other Manufacturers' Keishibukuryogan Formulas and TJ-25 for Precision Verification

Product	Number of lots	Cluster X	Cluster Y
Company A	2	2	1
Company B	2	1	5
Company C	1	3	5
Company D	2	2	5
Company E	2	2	4
Company F	6	2	5
Company G	1	5	4
Company H	1	4	4
Company I	3	1	5
TJ-25 (intermediate precision)	21	3	3
TJ-25 (repeatability)	8		

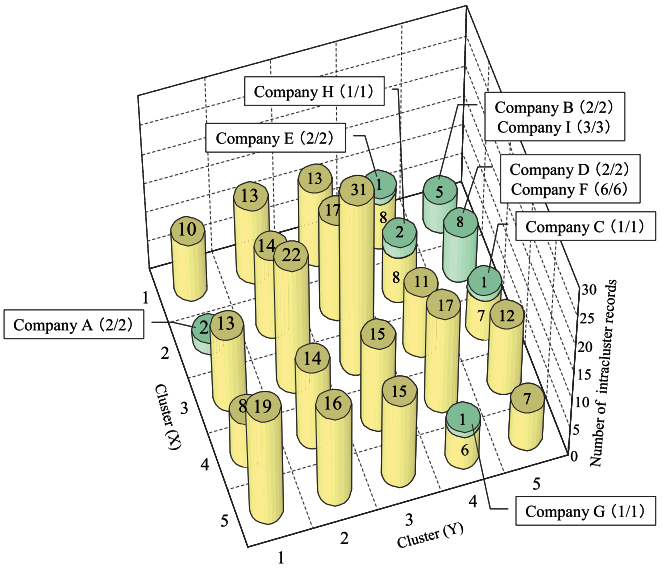

Fig. 3. 3-D Diagram of the Clusters of TJ-25 and Other Manufacturers' Keishibukuryogan Formulas

Yellow: TJ-25. Green: other manufacturers' keishibukuryogan formulas.

Cluster (X) \ Cluster (Y)	1		2		3		4		5
1	10	0.64	13	0.60	13	1.26	9	1.71	5
	1.35		0.56		0.75		0.92		2.48
2	2	1.38	14	0.83	17	1.46	10	1.92	8
	1.46		0.58		0.72		1.20		1.90
3	13	0.83	22	1.31	31	1.17	11	1.11	8
	0.68		0.95		0.36		1.24		1.15
4	8	1.27	14	0.50	15	0.83	17	1.07	12
	0.72		0.82		0.69		1.27		1.14
5	19	0.47	16	1.16	15	1.37	7	0.99	7

Fig. 4. Numbers of Records in Clusters and Euclidean Distances among Adjacent Clusters

Blue: intercluster distance is not more than 1.0. Yellow: intercluster distance is more than 1.0 and not more than 1.5. Pink: intercluster distance is more than 1.5. Green: cluster to which only other manufacturers' keishibukuryogan formulas are assigned.

clusters were disclosed to have greatly different amounts of contained components.

Regarding all TJ-25-assigned clusters, on the other hand, Euclidean distance among adjacent clusters was les than 1.5. Furthermore, the same manufacturer's lot(s) were assigned to individual clusters with respect to other manufacturer's keishibukuryogan formulas (Table 3). This result revealed that intramanufacturer consistency in amounts of contained components in the same manufacturer's formulas is higher than intermanufacturer consistency.

Comparisons of the components of 3 clusters, in which only other manufacturers' keishibukuryogan formulas had been clustered, revealed the following differences in HPLC-fingerprints among manufacturers: the cluster (1, 5) has high contents of paeoniflorin, benzoic acid, and cinnamic acid and low contents of oxypaeoniflorin, pentagalloylglucose, and paeonol; the cluster (2, 1) has low contents overall, especially low contents of paeoniflorin and benzoic acid; and the cluster (2, 5) has high contents of amygdalin, benzoic acid, and cinnamic aldehyde and low contents of cinnamic acid and coumarin (Fig. 5).

The same amounts of Cinnamomi Cortex, Poria, Moutan Cortex, Persicae Semen, and Paeoniae Radix are mixed to make the keishibukuryogan extract. Therefore, there is no difference in the combination ratios of botanical raw materials among drug manufacturers. However, conditions for the manufacturing process of the kampo medicine—extraction, concentration, desiccation, and granulation—differ among drug manufacturers. Furthermore, there are great differences in suppliers of botanical raw materials and in cultivation sites that affect the quality of the kampo medicine extracts. The successful evaluation of differences among drug manufacturers resulted from the facts that differences in manufacturing conditions among manufactures and in quality of botanical raw materials are manifested as differences in the feature values of peak areas on high-performance liquid chromatograms and that the cluster analysis using SOMs could precisely classify these differences.

Differences in HPLC-Fingerprints Due to Differences in Botanical Raw Materials Every drug manufacturer produces their formulas under established conditions. Therefore, the major factor to affect the quality of products is the quality of botanical raw materials. On the other hand, botanical raw materials are naturally-derived products. It is hence difficult to continually obtain botanical raw materials of given quality. There is a few bias in the quality of botanical raw materials, and the bias is considered to affect product quality.

Multiple lots, which had been manufactured at the same lot combination ratios of botanical raw materials (more than 5 records), were extracted, and the results of their clustering are shown in Table 4.

The records of 2 lot compositions (lot compositions h and j) were assigned to 2 clusters. However, the records of other lot compositions were assigned into an individual cluster per lot composition.

The above results clarified that the clustering using SOMs discerns slight differences in formula quality due to differences in botanical raw material lots of the keishibukuryogan formulas of the same manufacturer and allows the clustering almost exactly by botanical raw material lot composition.

Differences in HPLC-Fingerprints Due to Differences in Storage Period Drugs inevitably suffer changes in qual-

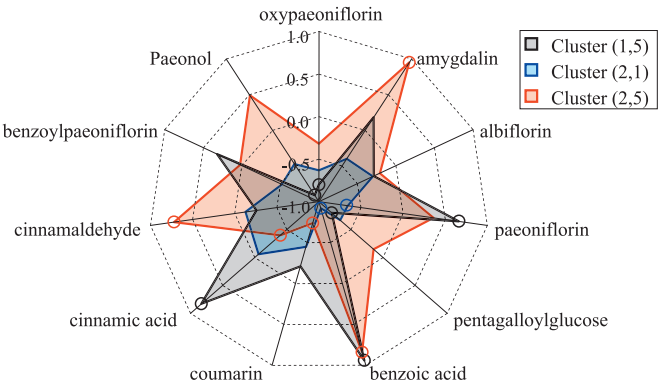

Fig. 5. Feature Values (Normalized Peak Areas) of Clusters into Which Only Other Manufacturers' Keishibukuryogan Formulas Were Classified

Table 4. Results of the Clustering of TJ-25 with the Same Lot Compositions of Botanical Raw Materials

Compositions of botanical raw materials	Number of lots	Cluster X	Cluster Y	Number of intracluster records
Lot composition a	8	1	2	8
Lot composition b	7	3	1	7
Lot composition c	7	5	2	7
Lot composition d	7	5	1	7
Lot composition e	7	4	1	7
Lot composition f	7	3	2	7
Lot composition g	7	5	2	7
Lot composition h	7	2	2	6
		5	1	1
Lot composition i	6	1	3	6
Lot composition j	6	2	3	1
		5	1	5
Lot composition k	5	1	1	5
Lot composition l	5	3	1	5

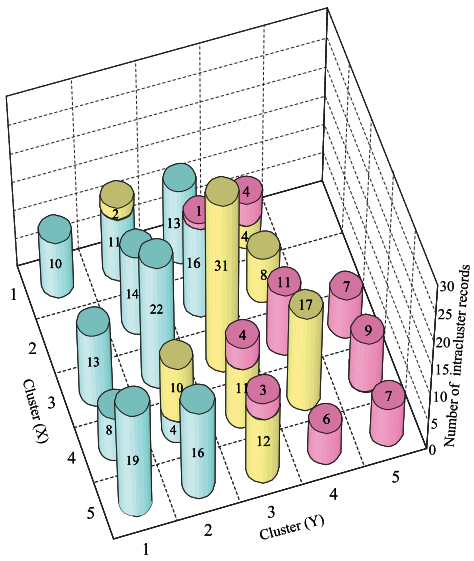

Fig. 6. 3-D Diagram of the Clusters of TJ-25 (Differences Due to Storage Period)

Blue: group A (storage period: ≤5 years). Yellow: group B (storage period: 6—8 years). Pink: group C (storage period: 9—10 years).

ity during storage period after manufacture. To assess differences in HPLC-finger prints during storage period, a total of 267 lots of TJ-25 were classified to 3 groups (group A: ≤5 years after manufacture; group B: 6—8 years after manufacture; and group C: 9—10 years after manufacture); the results of their clustering were compared (Fig. 6).

Although a slight overlap was found, these groups were classified nearly completely according to storage period. Therefore, classified clusters were verified to change along with storage period. Furthermore, a comparison of the components in each group verified that coumarin and cinnamic aldehyde contents decrease and benzoic acid contents increase along with a longer storage period.

Changes in the quality of botanical raw materials according to the year of manufacture and the minor effects of manufacturing conditions cannot be denied completely. Nevertheless, we consider that the cluster analysis using SOMs successfully detected differences in formula quality (amount of contained component) due to time-course changes.

Conclusion

There are many cases of using multivariate analyses to evaluate the quality of multiple component drugs, and cases of classifying interfactory differences by cluster analyses have been reported.[18] However, we used SOMs to evaluate the quality of multiple component drugs, ensured the precision of the same samples, and could develop a clustering procedure capable of discerning not only differences among drug manufacturers but also differences due to botanical raw materials of the same manufacturers.

In the keishibukuryogan formulas of the same drug manufacturers that show no large differences in manufacturing conditions and in procuration routes for botanical raw materials, we could evaluate, by cluster analyses, 1) differences in HPLC-fingerprints due to differences in lots of botanical raw materials and 2) differences in HPLC-fingerprints due to long-term storage-induced changes in the amounts of contained components.

Regarding differences in HPLC-fingerprints due to differences among drug manufacturers, furthermore, we verified that the analysis of feature values that were normalized after the SOM analysis can readily specify clustering factors (*i.e.*, feature values of component contents among manufacturers).

Therefore, we could demonstrate the usefulness of SOMs in evaluating the quality of multiple component drugs by HPLC fingerprints.

Experimental

HPLC Analysis. Preparation of Samples A sample was pulverized with a porcelaneous mortar and then passed through a sieve with wire opening of 300 μm. The pulverized sample equivalent to 115 mg of the extract was weighed and transferred into a 50-ml glass-stoppered centrifuge tube.

Add 30 ml of a mixture of methanol and 0.1% phosphoric acid (3 : 1), and use the reciprocating shaker to perform extraction by shaking for 30 min. After centrifugation (3000 rpm, 5 min), separate the supernatant liquid. Repeat the procedure with the residue using 15 ml of the mixture of methanol and 0.1% phosphoric acid (3 : 1), combine the supernatant liquids, and add the mixture of methanol and 0.1% phosphoric acid (3 : 1) to make exactly 50 ml. Centrifuge (12000 rpm, 10 min), and use the supernatant liquid as the sample solution.

Analytical Conditions Column: XBridge C18 (5 μm, 250×4.6 mm i.d., Waters).

Mobile Phase: Solution A, a mixture of 0.1% phosphoric acid, methanol, and acetonitrile (85 : 7.5 : 7.5). Solution B, a mixture of acetonitrile, 0.1 % phosphoric acid, and methanol (85 : 7.5 : 7.5). For 35 min after injection of the sample, deliver an increasing volume of the solution B in the mobile phase by the lineal ingredient from the solution A 100% in such a manner to obtain a mixture of solution A and solution B (1 : 1).

Flow Rate: 1.0 ml/min.
Column Temperature: 40 °C.
Injection Volume: 10 μl.
Detection Wavelength: 210 nm.

Equipment

Reciprocating Shaker:	SR-1 (TAITEC).
Centrifuge:	LC-121 (TOMY), 5415C (eppendorf).
HPLC:	1200 series (Agilent Technologies)
• Quaternary pump	#G1311,
• Degasser	#G1379A,
• Autosampler	#G1313A,
• Column compartment	#G1316A,
• Diode-array detector	#G1315B,
• ChemStation LC 3D System	Rev. A.10.01J.

References

1) Iwaoka E., Oku H., Ishiguro K., *J. Tradit. Med.*, **26**, 97—103 (2009).
2) Goto H., Kiga C., Nakagawa T., Koizumi K., Sakurai H., Shibagaki Y., Ogawa K., Shibahara Y., Shimada Y. Saiki I., *J. Tradit. Med.*, **22**, 237—243 (2005).
3) Li T. J., Qiu Y., Mao J. Q., Yang P. Y., Rui Y. C., Chen W. S., *J. Pharmacol. Sci.*, **105**, 34—40 (2007).
4) Negishi E., Takeo C., Nakajima A., Hisamitsu K., Amano K., Hirai A., Ueno K., *Jpn. J. Pharm. Health Care Sci.*, **32**, 21—26 (2006).
5) Tanaka T., Mizuno K., Umesaki N., Ogita S., *J. Tradit. Med.*, **14**, 442—443 (1997).
6) Tanaka T., Mizuno K., Umesaki N., Ogita S., *Osaka City Med. J.*, **44**, 117—124 (1998).
7) Chen J. T., Shiraki M., *Maturitas*, **45**, 199—204 (2003).
8) Koyama T., *J. Jpn. Menopause Soc.*, **1**, 75—79 (1993).
9) Shimada Y., Kogure T., Goto H., Sakai S., Sekiya N., Mantani N., Terasawa K., *J. Tradit. Med.*, **18**, 122—126 (2001).
10) "Supplement I to The Japanese Pharmacopoeia," Fifteenth Edition 140 (2007).
11) Chen L., Wang D., Wu J, Yu B., Zhu D., *J. Pharm. Biomed. Anal.*, **49**, 267—275 (2009).
12) Tani T., *Journal of Traditional Sino-Japanese Medicine*, **5**, 53—56 (1984).
13) Center for Drug Evaluation and Research, "Guidance for Industry Botanical Drug Products," Department of Health and Human Services, Food and Drug Administration, U.S.A., 2004.
14) Lu G. H., Chan K., Liang Y. Z., Leung K., Chan C. L., Jiang Z. H., Zhao Z. Z., *J. Chromatogr. A*, **1073**, 383—392 (2005).
15) Gan F., Ye R., *J. Chromatogr. A*, **1104**, 100—105 (2006).
16) Xie P., Chen S., Liang Y., Wang X., Tian R. Upton R., *J. Chromatogr. A*, **1112**, 171—180 (2006).
17) Xu C. J., Liang T. Z., Chau F. T., Heyden Y. V., *J. Chromatogr. A*, **1134**, 253—259 (2006).
18) Lu H. M., Liang Y. Z., Wu X. J., Qiu P., *Chem. Pharm. Bull.*, **54**, 725—730 (2006).
19) Wiberg K., *J. Chromatogr. A*, **1108**, 50—67 (2006).
20) Nederkassel A. M., Daszykowski M., Ellers P. H. C., Heyden Y. V., *J. Chromatogr. A*, **1118**, 199—210 (2006).
21) Nakai Y., Yano K., Shiba M., Kondo K., Takeda O., Sakakibara I., Terebayashi S., Takeda S., Okada M., *J. Jpn. Botany*, **81**, 63—74 (2006).