JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Multicomponent Spectral Correlative Chromatography Applied to Complex Herbal Medicines

YUN HU, YI-ZENG LIANG,* BO-YAN LI, AND XIAO-NING LI

College of Chemistry and Chemical Engineering, Research Center of Modernization of Chinese Herbal Medicines, Central South University, Changsha 410083, People's Republic of China

YI-PING DU

College of Chemical Engineering, Shandong University of Technology, Zibo 255000, People's Republic of China

In this study, a novel chemometric algorithm is presented to facilitate the comparison of relevant chemical components from different herbal samples. This so-called multicomponent spectral correlative chromatography (MSCC) is developed to detect and decide whether two chromatographic clusters are correlated spectrally with each other. The target chromatographic cluster is first partitioned from one herbal spectrochromatogram obtained by hyphenated chromatography. Then, a projection operator is constructed with the principal spectral features extracted from the target to judge the presence or absence of a spectral correlative chromatographic cluster within another herbal spectrochromatogram. For this judgment, congruence coefficient between the original spectral vector and its projected residual is proposed to eliminate the influences from background and noises, especially heteroscedastic noises in the original data. The performance of the MSCC algorithm is demonstrated on both simulated data and real data, and its advantages and disadvantages are also discussed in some detail.

KEYWORDS: Multicomponent spectral correlative chromatography; spectrochromatogram; fingerprinting; chemometrics; orthogonal projection; *Schisandra chinensis* (Turcz.) Baill.; *Houttuynia cordata* Thunb.

INTRODUCTION

With the development of traditional herbal medicines and desirable progress of relevant researches, both the compilations of pharmaceutical attributes and the comparison of componential distributions about herbal samples have become more and more important, in particular for the medicinal quality assessment to support their uses. Conventional analysis of herbal medicines concentrated mainly on investigating single or few marker components for their pharmaceutical efficacy and quality control. Nevertheless, it is not enough to reveal the complexity and synergic effect of the phytochemicals in herbal medicines and sometimes leads to a certain biased assessment of the investigated systems. As pointed out in ref *1*, "The quantity and quality of the safety and efficacy data on traditional medicine are far from sufficient to meet the criteria needed to support its use worldwide. The reasons for the lack of research data are due to not only to health care policies, but also to a lack of adequate or accepted research methodology for evaluating traditional medicine." It is well-known that, unlike a chemically synthetic drug with much purity, an herbal medicine may consist of many complex phytochemicals. As a result, it becomes very difficult or impossible in most cases to identify most of these phytochemicals by means of common approaches (*2*−*4*). Thus, the chromatographic fingerprinting (*5*−*7*) is proposed as a more meaningful alternative to the conventional analysis. It emphasizes a systemic characterization of herbal samples. However, when it comes to the comparison and evaluation of various chromatographic fingerprints, it seems inevitable to identify the presence or absence of the same compositions in them. Additionally, there are many factors influencing the chromatographic fingerprints of herbal medicines, such as different collected seasons, different cultured places, different chemical extraction methods, and even the measurement devices. Therefore, the efficient and reasonable evaluation of the sameness and difference among the fingerprints from different sources becomes very important.

Fortunately, with the rapid development of analytical instruments, a vast array of instruments, such as GC−MS, GC−IR, LC−DAD, CE−MS, and LC−NMR, have appeared, which will give us a good opportunity to fulfill the previously mentioned task with the help of chemometric methods (*8*−*14*). The spectrochromatograms can be obtained from such kinds of instruments. These spectrochromatographic fingerprints contain abundant information about the compositional spectra and

* To whom correspondence should be addressed. Tel.: 86-731-8822841. Fax: 86-731-8825637. E-mail: yizeng_liang@263.net.

10.1021/jf0489318 CCC: $27.50
Published on Web 12/02/2004

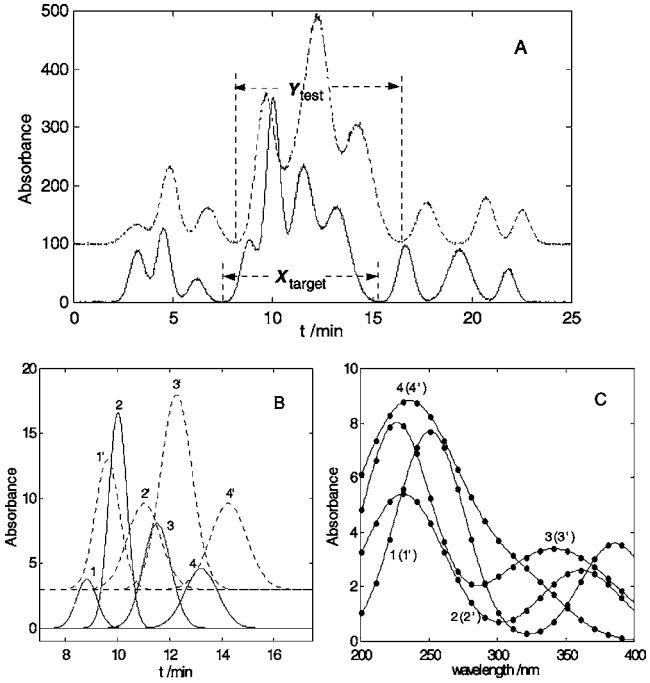

Figure 1. (**A**) Chromatograms of two simulated spectrochromatographic fingerprints at some wavelength. (**B**) Chromatographic profiles of four-component clusters $\boldsymbol{X}_{target}$ (solid lines) and $\boldsymbol{Y}_{test}$ (dotted lines). (**C**) Spectra of clusters $\boldsymbol{X}_{target}$ (solid lines) and $\boldsymbol{Y}_{test}$ (dotted lines).

chromatograms. Thus, the efficient usage of such information to evaluate reasonably the sameness or difference among the fingerprints becomes possible (*15*−*19*). In ref *18*, a spectral correlative chromatography (SCC) algorithm is presented to detect same components from different spectrochromatograms, on the assumption that the target cluster is not disturbed by any other component, in other words, the target cluster must be pure. However, in a real complex herbal system, some relevant components often coelute within a same chromatographic cluster, and it is beyond the detection ability of SCC. Accordingly, multicomponent spectral correlative chromatography (MSCC) (*19*) was developed to deal with such multicomponent−coeluted chromatographic clusters. It utilizes the spectral features extracted from the target chromatographic cluster by principal component analysis (PCA) to construct a projection operator, which is able to annihilate the spectral correlative information by projecting other chromatographic clusters or spectrochromatogrms onto this operator. With this approach, one can easily compare and reveal the sameness or difference among different spectrochromatograms without the necessity to acquire the pure spectra and chromatograms of some relevant components or chemical standards existing in different herbal samples, especially in the case that active constituents or characteristic ingredients cannot be separated completely in the chromatography. Extensive theory of the MSCC algorithm (*19*) has recently been proposed and applied successfully to data analysis. This work focuses on some practical application of MSCC to complex herbal medicine data sets. The results obtained are quite satisfactory.

MATERIALS AND METHODS

Simulated Data. Two spectrochromatographic fingerprints were simulated to demonstrate the procedure of MSCC method, as shown in **Figure 1**. The chromatographic profiles and spectra of 10 components were generated by means of the Gaussian function, respectively. Two four-component peak clusters, named $\boldsymbol{X}_{target}$ and $\boldsymbol{Y}_{test}$, were selected to test this method. The chromatographic profiles and spectra of the individual components are marked with 1&1′, 2&2′, 3&3′, and 4&4′, respectively. Both homoscedastic noises with 0.2% standard deviation of maximum intensity and heteroscedastic noises with the standard deviations in the range of 0.2 to ∼1% proportional to the corresponding responses were superimposed to these two simulated data.

Materials. *Schisandra chinensis* (Turcz.) Baill. was purchased from an authoritative pharmaceutical store and identified at the Institute of Materia Medica, Hunan Academy of Traditional Chinese Medicine and Materia Medica, Changsha, Hunan, China. *Houttuynia cordata* Thunb. samples were provided by Xingzhong Pharmaceutical Corp. Ltd. and Yusi Pharmaceutical Corp. Ltd. of Yunnan province, People's Republic of China.

Extraction of the Volatile Fractions. Plant material of *S. chinensis* was dried for about 60 min at 40 °C and then crushed (screen size 20 mesh) in a pharmaceutical disintegrator. Two extraction techniques were involved: (1) according to the standard extraction method of *Chinese Pharmacopoeia* (*25*), extracting the essential oil from traditional medicines by use of a standard essential oil extractor and (2) extracting the essential oil via steam distillation. About 500 g of preweighed fresh *H. cordata* was cut into fragments and soaked for 2 h with 1000 mL of distilled water in a standard extractor at ambient temperature. Then, the essential oil was prepared according to the standard extraction method of *Chinese Pharmacopoeia* (*25*), as mentioned in the first technique.

Analysis of the Volatile Fractions. A GC-17A gas chromatograph and QP-5000 mass spectrometer from Shimadzu (Tokyo, Japan) was employed in this study. In the analysis of the volatile fractions from *S. chinensis*, an OV-1 capillary column (30 m × 0.25 mm i.d.) was initially held at 40 °C and then linearly heated to 230 °C at a rate of 5 °C min^{-1}. The inlet temperature was kept at 210 °C, and the helium carrier gas was employed at a flow rate of 0.7 mL min^{-1}. In the mass spectrometer, electron impact (EI^+) mass spectra were recorded at 70 eV ionization energy in full scan mode in the mass range of 20 to ∼350 amu with a 0.2 s $scan^{-1}$ velocity. The ionization source temperature was set at 230 °C. In the analysis of the volatile fractions from *H. cordata*, an OV-1 quartz capillary column (30 m × 0.25 mm i.d.) was used with a split ratio of 30:1. The column temperature was initially maintained at 50 °C for 6 min and then programmed from 50 to 230 °C at the rate of 10 °C min^{-1}. The inlet temperature was kept at 280 °C. Helium carrier gas was also used at a flow rate of 0.7 mL min^{-1}. In the mass spectrometer, the electron impact (EI^+) mass spectra were recorded at 70 eV ionization energy in full scan mode in the 30 to ∼350 amu mass ranges with 0.2 s $scan^{-1}$ velocity. The ionization source temperature was set at 250 °C.

Data Analysis. All the involved programs were coded in MATLAB 5.3 environment and were executed on a Pentium III 850 (Intel) personal computer with 256MB RAM under Microsoft Windows 98 operating system. The library search and spectral match for chemical species were conducted on the National Institute of Standards and Technology (NIST) MS database containing about 107 000 compounds.

RESULTS AND DISCUSSION

Multicomponent Spectral Correlative Chromatography for Data Analysis. Suppose that data matrix $\boldsymbol{X}_{target}$ of size $m_1 \times n$ is a target chromatographic peak cluster from spectrochromatographic fingerprint $\boldsymbol{X}$ of one herbal sample, and $\boldsymbol{Y}$ of size $m \times n$ is another spectrochromatographic fingerprint, as shown in **Figure 1A**. Moreover, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are measured at n same spectral positions. In this figure, two peak clusters, named $\boldsymbol{X}_{target}$ (solid lines) and $\boldsymbol{Y}_{test}$ (dotted lines), are also marked. **Figure 1B**,**C** depicts the chromatographic profiles and spectra of four components in them, marked 1&1′, 2&2′, 3&3′, and 4&4′, respectively.

As can be seen from them, the four components in $\boldsymbol{X}_{target}$ continually elute and partially overlap, and $\boldsymbol{Y}_{test}$ contains the same spectra but different chromatographic eluting profiles. The purpose of MSCC algorithm is to detect the cluster within $\boldsymbol{Y}$, spectrally correlating to the cluster $\boldsymbol{X}_{target}$. To fulfill the

requirement, the spectral features have first to be extracted from $\boldsymbol{X}_{\text{target}}$ by means of PCA (*20*, *21*). The target cluster $\boldsymbol{X}_{\text{target}}$ can be decomposed by singular value decomposition (SVD)

$$\mathbf{X}_{\text{target}} = \boldsymbol{USV}^{\text{T}} \quad (1)$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the score and loading matrices, respectively, and $\boldsymbol{S}$ is a diagonal matrix, which collects the square-root values of all eigenvalues of $\boldsymbol{X}_{\text{target}}$. The superscript T denotes the transposition of matrix or vector. It is worth noting that the loading matrix $\boldsymbol{V}$ collects all the spectral information of $\boldsymbol{X}_{\text{target}}$. If the number of p principal components (PCs) included in the cluster $\boldsymbol{X}_{\text{target}}$ is estimated correctly, eq 1 can be rewritten

$$\mathbf{X}_{\text{target}} = \boldsymbol{U}_p\boldsymbol{S}_p\boldsymbol{V}_p^{\text{T}} + \boldsymbol{E} \quad (2)$$

Here, the matrix $\boldsymbol{E}$ contains instrumental noises, experimental errors, and the part of the data $\boldsymbol{X}_{\text{target}}$ not explained by p PCs. As a result, the loading matrix $\boldsymbol{V}$ in eq 1 can be reduced to matrix $\boldsymbol{V}_p$ that consists of p orthonormal spectral features (i.e., p principal components). Subsequently, these p extracted spectral features are utilized to construct an orthogonal projection operator $\boldsymbol{P}$

$$\mathbf{P} = (\boldsymbol{I} - \boldsymbol{V}_p\boldsymbol{V}_p^{\text{T}}) \quad (3)$$

$\boldsymbol{I}$ denotes an identity matrix of the same size as matrix $(\boldsymbol{V}_p\boldsymbol{V}_p^{\text{T}})$. Then, we project the original spectrum $\boldsymbol{y}_j$ ($j = 1, ..., m$) recorded at every chromatographic point of the investigated spectrochromatogram $\boldsymbol{Y}$ onto this operator $\boldsymbol{P}$, and a series of corresponding residual vectors are obtained

$$\boldsymbol{y}_j^* = \boldsymbol{P}\boldsymbol{y}_j = (\boldsymbol{I} - \boldsymbol{V}_p\boldsymbol{V}_p^{\text{T}})\boldsymbol{y}_j \ (j = 1, ..., m) \quad (4)$$

$\boldsymbol{y}_j^*$ represents the residual of recorded spectrum $\boldsymbol{y}_j$ at the chromatographic retention time j after orthogonal projection operation.

If the spectrum $\boldsymbol{y}_j$ embodies the spectral information correlating to p feature spectra of $\boldsymbol{V}_p$, the correlative information would be removed from the original $\boldsymbol{y}_j$, and the Euclidean norm (re_j) of its projected residual $\mathbf{y}_j^*$ should theoretically be zero, that is

$$re_j^2 = ||\boldsymbol{y}_j^*||^2 = \boldsymbol{y}_j^{*\text{T}}\boldsymbol{y}_j^* = (\boldsymbol{P}\,\boldsymbol{y}_j)^{\text{T}}\boldsymbol{P}\,\boldsymbol{y}_j = \boldsymbol{y}_j^{\text{T}}\boldsymbol{P}\,\boldsymbol{y}_j = 0 \quad (5)$$

Otherwise, the Euclidean norm (re_j) of its projected residual $\mathbf{y}_j^*$ will be extremely large. In fact, the projected residual norm (re_j) can only be close to zero instead of zero because there exist instrumental and experimental noises or errors during the measurements. If the residuals in the chromatographic region $\boldsymbol{Y}_{\text{test}}$ are merely homoscedastic noises, a minimum projection flat where the values of re_j are close to zero should be obtained by use of eq 5, for example, the projection result from the simulated data in **Figure 2A**. However, the occurrence of heteroscedastic noises (*22*, *23*) in real data often results in very large Euclidean norm values (re_j) at some peak positions, as shown in the marked region in **Figure 2B**. Such a projection result maybe lead to an erroneous determination that clusters $\boldsymbol{X}_{\text{target}}$ and $\boldsymbol{Y}_{\text{test}}$ consist of different compositions at all. To eliminate the influence of heteroscedastic noises, the congruence coefficient between the original spectral vector and its projected residual is adopted as a better alternative to eq 5

$$r_{\text{j}} = \frac{\boldsymbol{y}_j^{\text{T}}\cdot\boldsymbol{y}_j^*}{||\boldsymbol{y}_j||\cdot||\boldsymbol{y}_j^*||} \ (j = 1, ..., m) \quad (6)$$

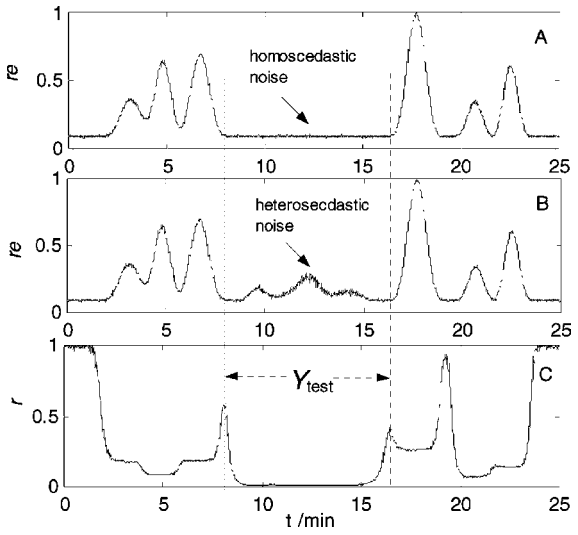

Figure 2. (**A**) Projected residual Euclidean norm curve (*re*) from every spectrum of spectrochromatogram $\boldsymbol{Y}$ into which homoscedastic noises were only superimposed. (**B**) The projected residual Euclidean norm curve (*re*) from every spectrum of $\boldsymbol{Y}$ into which homoscedastic and heteroscedastic noises were both superimposed. (**C**) The spectral correlative chromatogram (*r*) from every spectrum of $\boldsymbol{Y}$ into which homoscedastic and heteroscedastic noises were both superimposed.

The values r_j ($j = 1, ..., m$) are in the range of $0 \le r \le 1$. The smaller the value of r_j is, the higher the correlation between $\boldsymbol{y}_j$ and the feature spectra within matrix $\boldsymbol{V}_p$ is. When $\boldsymbol{y}_j$ falls completely into the subspace spanned by feature spectra within matrix $\boldsymbol{V}_p$, the value of r_j should be zero. The congruence efficient between the original $\boldsymbol{y}_j$ and its projected residual $\mathbf{y}_j^*$ takes mainly the morphologic distribution of the projected residuals into account: if the residual $\mathbf{y}_j^*$ is of noise property, it will be orthogonal to both the spectral signal of $\boldsymbol{y}_j$ and the spanned subspace. Thus, a projection curve about r_j can be obtained in the chromatographic direction of spectrochromatogram $\boldsymbol{Y}$ (**Figure 2C**). Moreover, this curve has the same length with spectrochromatogram $\boldsymbol{Y}$ and contains the spectral correlative information about $\boldsymbol{Y}$. With this correlative curve, one may easily pick up the spectral correlative information in $\boldsymbol{Y}$ to $\boldsymbol{X}_{\text{target}}$, and one can easily compare the sameness or difference between two spectrochromatographic fingerprints without necessity to obtain the pure spectra and chromatograms of each component or reference standards in herbal samples. Consequently, this method is named multicomponent spectral correlative chromatography (MSCC) in this work.

GC−MS Data of *S. chinensis*. With the increasing demand of quality control for herbal medicines, comparing the same and different compositions among the spectrochromatograms from different extraction methods plays an important role in experimental researches of herbal chromatographic fingerprints. For simple systems, this comparison can be done easily. For complicated systems, however, seriously overlapping peaks are often encountered. It is common that the overlapping peaks have to be first resolved into pure chromatographic and spectral profiles by employing some sophisticated chemometric methods (*12*−*14*), and then this comparison becomes possible. These methods not only undertake heavy calculation but also need considerable experiences so as to obtain good resolution results (*26*) when they are employed. Thus, the simpler method is obviously preferable for such a quick comparison. Here, the MSCC approach is introduced to detect the common components in herbal samples from different extraction methods or materials measured by chromatography, such as GC−MS and HPLC−

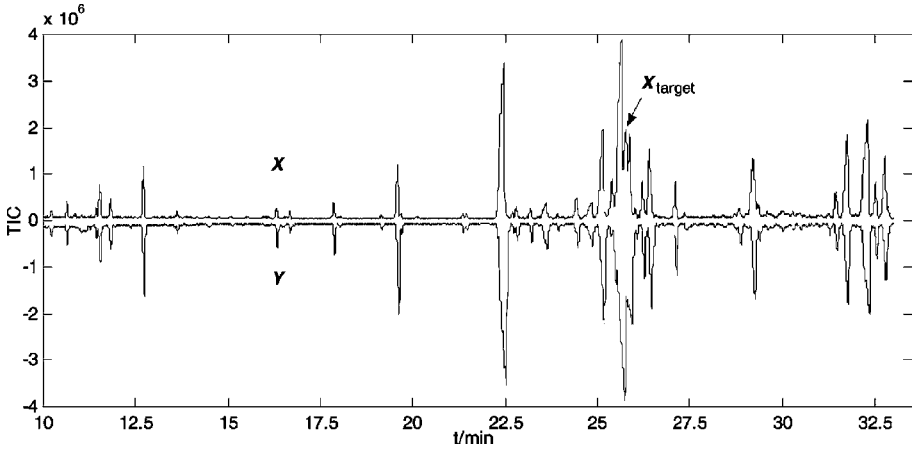

Figure 3. TIC chromatograms of the volatile fractions of *S. chinensis* from different extraction methods: fingerprints ***X*** corresponding to method 1 and ***Y*** to method 2, respectively.

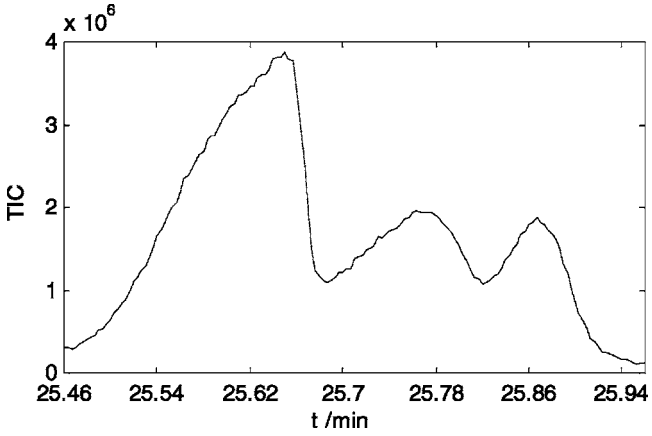

Figure 4. TIC chromatogram of target cluster $\boldsymbol{X}_{\text{target}}$ in the region of 25.46 to ~25.94 min of ***X***.

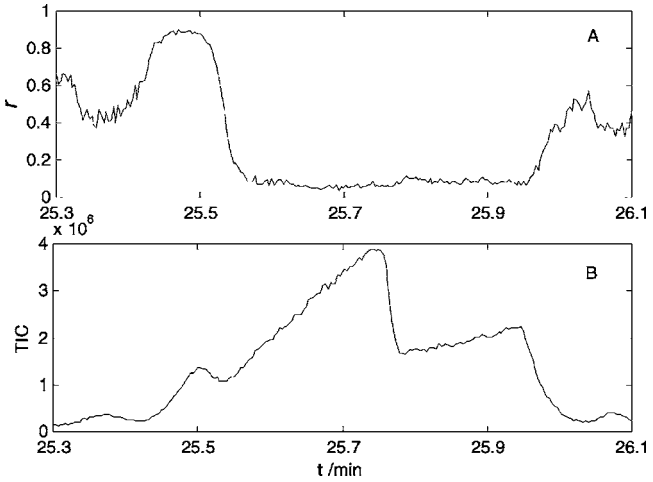

Figure 5. (**A**) Amplified segment of the whole spectral correlative chromatogram of fingerprint ***Y*** after being projected onto the spectral feature operator of $\boldsymbol{X}_{\text{target}}$, where the values of *r* reach lowest. (**B**) TIC chromatogram of the tested cluster in ***Y*** correlating spectrally to $\boldsymbol{X}_{\text{target}}$.

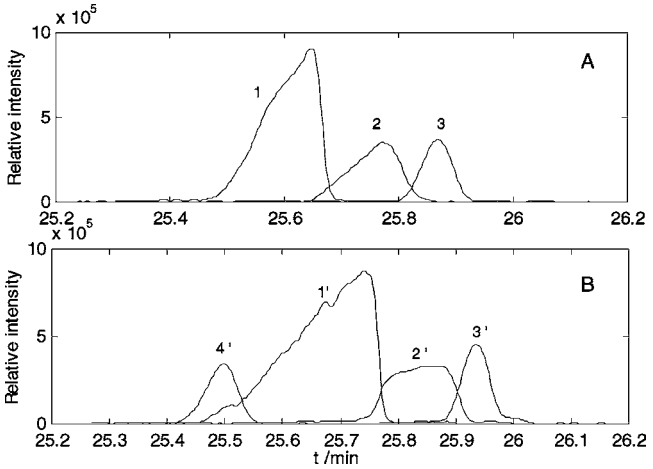

Figure 6. Resolved chromatographic profiles of target cluster $\boldsymbol{X}_{\text{target}}$ (**A**) and the tested spectral correlative cluster (**B**), respectively.

DAD; furthermore, it is not necessary to resolve individual components into pure spectra and chromatograms.

Figure 3 shows two total ionic current (TIC) chromatograms of *S. chinensis* volatile fractions derived from different extraction methods but measured at the same GC−MS procedure. The general features of the two fingerprints are quite similar even though there are some chromatographic shifts. However, seen from them, it is difficult to conclude that the same compounds are present in some overlapping chromatographic peak clusters or not without any resolution, for instance, the peak $\boldsymbol{X}_{\text{target}}$ derived from extraction method 1 and corresponding chromatographic cluster derived from extraction method 2.

Now, we extract the spectral features from $\boldsymbol{X}_{\text{target}}$ in the region of 25.46 to ~25.94 min (**Figure 4**) with PCA. It is found that this cluster consists of three PCs, and a projection operator ***P*** is constructed with these three spectral features, say $\boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{V}_3\boldsymbol{V}_3^{\mathrm{T}})$. Then, every spectral vector $\boldsymbol{y}_j$ of spectrochromatogram ***Y*** extracted by method 2 is projected onto ***P***, and a spectral correlative chromatogram is obtained, consequently. **Figure 5A** just depicts a segment of this chromatogram. A minimum flat illustrates that the last two peaks of ***Y*** in the region of 25.55 to ~26.00 min are spectrally correlative to $\boldsymbol{X}_{\text{target}}$, whereas the congruence coefficients in the region of 25.45 to ~25.55 min are very large, and the small peak in this region is excluded from spectral correlative cluster (**Figure 5B**). To confirm this, we resolved these two peak clusters. The results are shown in **Figure 6**. The peak marked with 4′ (see **Figure 6B**) is really not the common component of these two clusters. More details on them can be found in ref *17*. From this result, it can be concluded that the compositions are not consistent in their individual retention time even though the same herbal sample is extracted by similar methods, and the MSCC algorithm allows us to rapidly decide whether some relevant chemical components

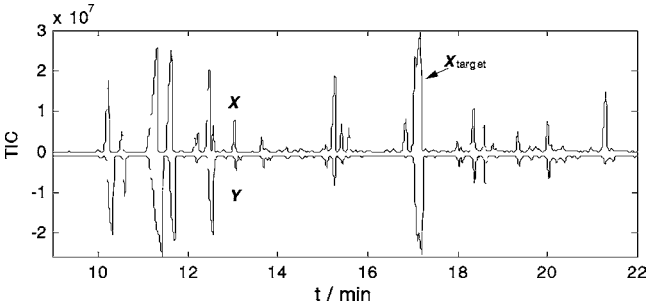

Figure 7. TIC chromatograms of *H. cordata* volatile fractions from two Pharmaceutical Corp. Ltds.

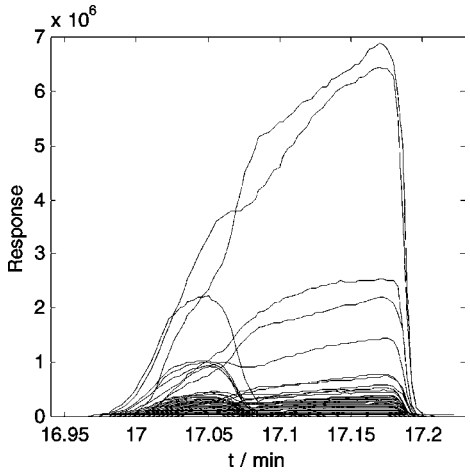

Figure 8. Amplified spectrochromatogram of target cluster $\boldsymbol{X}_{\text{target}}$.

are same or not. It is possible to further evaluate which extraction method is more advisable.

GC−MS Data of *H. cordata*. Besides being from the same herbal medicine extracted by different methods, the MSCC method can also detect spectral correlative substances in herbal chromatographic fingerprints from different resources. **Figure 7** displays two total ionic current (TIC) chromatograms of *H. cordata* volatile fractions from different resources, the top $\boldsymbol{X}$ is from Xingzhong Pharmaceutical Corp. Ltd., and the bottom $\boldsymbol{Y}$ is from Yusi Pharmaceutical Corp. Ltd. Because bornyl acetate is one of critical active and characteristic constituents for quality control of *H. cordata* samples, and it is very difficult to separate this compound completely, the cluster $\boldsymbol{X}_{\text{target}}$ where bornyl acetate exists is selected as the investigated target. It is amplified in **Figure 8**. Seen intuitively from it, this cluster consists of two components at least. However, if only two feature spectra extracted from $\boldsymbol{X}_{\text{target}}$ are used to construct $\boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{V}_2\boldsymbol{V}_2^{\mathrm{T}})$ and to project every spectral vector $\boldsymbol{y}_j$ of $\boldsymbol{Y}$ onto it, we achieve an overdetermined result (see the minimum projection values of curve 1 in **Figure 9A**). It can be speculated that two spectral features seem to be not enough to characterize the whole spectral contributions in $\boldsymbol{X}_{\text{target}}$, and there needs to be more PCs to construct the operator $\boldsymbol{P}$.

More should be discussed about the estimate of p principal feature spectra. If p is too small, the $\boldsymbol{V}_p$ subspace may not adequately represent the subspace of all components in cluster $\boldsymbol{X}_{\text{target}}$. The value of p should be large enough to summarize all of the significant components within the few spectral features but not so large that it incorporates a substantial amount of noise. In principle, p principal feature spectra in $\boldsymbol{X}_{\text{target}}$ could be estimated, but it is difficult to correctly estimate p in practice due to great variations in the concentration level of each component and noise. In this study, the method named noise

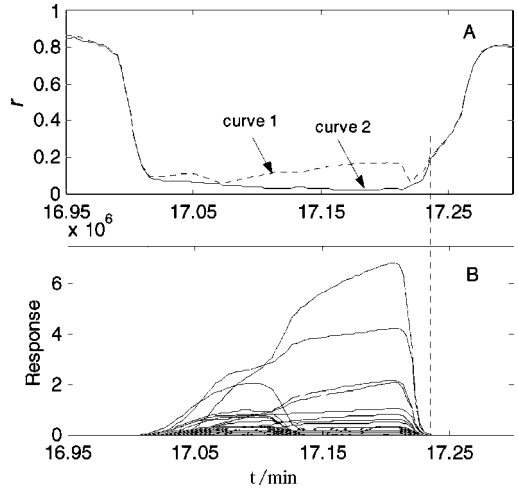

Figure 9. Amplified tested spectral correlative chromatogram with the lowest values of r (**A**) and its corresponding spectrochromatographic cluster (**B**).

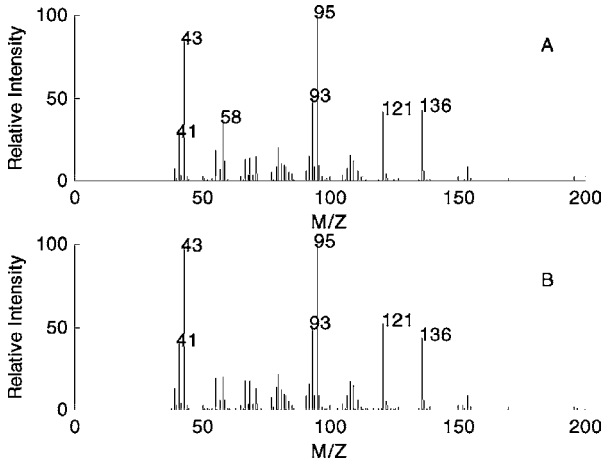

Figure 10. Resolved mass spectra of the first component (Bornyl acetate, $C_{12}H_{20}O_2$) that the target cluster $\boldsymbol{X}_{\text{target}}$ (**A**) and the tested spectral correlative cluster (**B**) have in common.

perturbation in functional principal component analysis (NPFPCA) is employed to estimate the spectral feature number. Factually, the number of principal feature spectra in $\boldsymbol{X}_{\text{target}}$ was estimated to be 4 with NPFPCA, say $p = 4$. Since an explicit description about the NPFPCA method was documented in ref *21*, here it is not elaborated for the sake of brevity.

As a result of the MSCC method, a minimum flat of the correlative chromatogram is obtained in the region 17.02 to ~17.23 min (see curve 2 in **Figure 9A**), when $\boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{V}_4\boldsymbol{V}_4^{\mathrm{T}})$. The spectral congruence coefficients in this region are close to zero. This flat corresponds to a cluster in the fingerprint $\boldsymbol{Y}$ shown in **Figure 9B**. The result shows that this tested cluster is spectrally correlative to the cluster $\boldsymbol{X}_{\text{target}}$ and that they encompass common chemical species. Because of the existence of selective and zero regions of the left first component of the target cluster $\boldsymbol{X}_{\text{target}}$ and the tested cluster, one may resolve its pure mass spectrum (**Figure 10**) from these two clusters with the help of chemometric resolution methods (*13*, *14*). The similarity match for this resolved component is conducted in the NIST mass library. Compared with the standard mass spectra, it is identified as bornyl acetate ($C_{12}H_{20}O_2$). However, the other three common components could not be resolved properly. MSCC method facilitates the comparison of critical active or characteristic constituents from herbal samples and further quality assessment of samples.

Conclusion. On the basis of spectral correlation, MSCC has attempted to detect and determine same components and to facilitate comparison of the sameness and difference of relevant compositions in various chromatographic fingerprints from herbal medicine samples, especially in the case where two or more components coelute in some chromatographic peak clusters. The introduction of the congruence coefficient between the original spectral vector and its projected residual can eliminate effectively the influences from heteroscedastic noises in original data. In addition, it is necessary to note that the comparison of the compositions in small contents might be beyond the ability of MSCC when the signal-to-noise ratio is very low.

ACKNOWLEDGMENT

We thank Dr. Cheng-Jian Xu for helpful discussions and technical assistance. We gratefully acknowledge anonymous reviewers for their many criticisms and suggestions, which greatly improved the paper.

LITERATURE CITED

(1) World Health Organization. *General Guidelines for Methodologies on Research and Evaluation of Traditional Medicines*; WHO: Geneva, 2000; pp 1−10.
(2) Yan, X. J.; Zhou, J. J.; Xie, G. R.; Milne, G. W. A. *Traditional Chinese Medicines: Molecular Structures, Natural Sources, and Applications*; Aldershot: Ashgate, 1999.
(3) Raven, P. H.; Evert, R. F.; Eichhorn, S. E. *Biology of Plants*, 6th ed.; W. H. Freeman: New York, 1999.
(4) Bauer, R. Quality criteria and standardization of phytopharmaceuticals: can acceptable drug standards be achieved? *Drug Inform. J.* **1998**, *32*, 101−110.
(5) *FDA Guidance for Industry—Botanical Drug Products (Draft Guidance)*; U.S. Food and Drug Administration: Washington, DC, 2000; Vol. VIII B, pp 18−22.
(6) Philipsom, J. D. *British Herbal Pharmacopoeia*; British Herbal Medicine Association: Bornemouth, 1996.
(7) Xie, P. S. A feasible strategy for applying chromatography fingerprint to assess quality of Chinese herbal medicine. *Trad. Chin. Drug Res. Clin. Pharm.* **2001**, *12*, 141−151.
(8) Lawton, W. E.; Sylvester, E. A. Self-modeling curve resolution. *Technometrics* **1971**, *13*, 617−633.
(9) Malinowski, E. R. *Factor Analysis in Chemistry*, 3rd ed.; Wiley: New York, 2002.
(10) Karjalainen, E. J. The spectrum reconstruction problem, use of alternating regression for unexpected spectral components in two-dimensional spectroscopies. *Chemom. Intell. Lab. Syst.* **1989**, *6*, 31−38.
(11) Malinowski, E. R. Window factor analysis: theoretical derivation and application to flow injection analysis data. *J. Chemom.* **1992**, *6*, 29−40.
(12) Kvalheim, O. M.; Liang, Y. Z. Heuristic evolving latent projections: resolving two-way multicomponent data. 1. Selectivity, latent-projective graph, datascope, local Rank, and unique resolution. *Anal. Chem.* **1992**, *64*, 936−946.
(13) Liang, Y. Z.; Kvalheim, O. M.; Keller, H. R.; Massart, D. L.; Kiechle, P.; Erni, F. Heuristic evolving latent projections: resolving two-way multicomponent data. 2. Detection and resolution of minor constituents. *Anal. Chem.* **1992**, *64*, 946−953.
(14) Manne, R.; Shen, H. L.; Liang, Y. Z. Subwindow factor analysis. *Chemom. Intell. Lab. Syst.* **1999**, *45*, 171−176.
(15) Lohnes, M. T.; Guy, R. D.; Wentzell, P. D. Window target-testing factor analysis: theory and application to the chromatographic analysis of complex mixtures with multiwavelength fluorescence detection. *Anal. Chim. Acta* **1999**, *389*, 95−113.
(16) Brown, C. D.; Wentzell, P. D. A modification to window target-testing factor analysis using a Gaussian window. *Chemom. Intell. Lab. Syst.* **2000**, *51*, 3−7.
(17) Li, X. N.; Cui, H.; Song, Y. Q.; Liang, Y. Z.; Chau, F. T. Analysis of volatile fractions of *Schisandra Chinensis* (Turcz.) Baill. using GC-MS and chemometric resolution. *Phyochem. Anal.* **2003**, *14*, 23−33.
(18) Li, B. Y.; Hu, Y.; Liang, Y. Z.; Huang, L. F.; Xie, P. S. Spectral correlative chromatography and its application to analysis of chromatographic fingerprints of herbal medicines. *J. Sep. Sci.* **2004**, *27*, 581−588.
(19) Hu, Y.; Liang, Y. Z.; Li, B. Y.; Zeng, Z. D. Multicomponent spectral correlative chromatography applied to analysis of chemical fingerprints of traditional Chinese medicines. *Acta Chim. Sinica.* **2003**, *61*, 1466−1470.
(20) Maeder, M. Evolving factor analysis for the resolution of overlapping chromatographic peaks. *Anal. Chem.* **1987**, *59*, 527−530.
(21) Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37−52.
(22) Xu, C. J.; Liang, Y. Z.; Li, Y.; Du, Y. P. Chemical rank estimation by noise perturbation in functional principal component analysis. *Analyst* **2003**, *128*, 75−81.
(23) Keller, H. R.; Massart, D. L.; Liang, Y. Z.; Kvalheim, O. M. Evolving factor analysis in the presence of heteroscedastic noise. *Anal. Chim. Acta* **1992**, *263*, 29−36.
(24) Kvalheim, O. M.; Brakstad, F.; Liang, Y. Z. Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise. *Anal. Chem.* **1994**, *66*, 43−51.
(25) Chinese Pharmacopoeia Committee. *Chinese Pharmacopoeia*; Chemical Industry Press: Beijing, China, 2000; Appendix 64.
(26) Gong, F.; Liang, Y. Z.; Cui, H.; Chau, F. T.; Chan, B. T. P. Determination of volatile components in peptic powder by gas chromatography−mass spectrometry and chemometric resolution. *J. Chromatogr. A* **2001**, *909*, 237−247.

Received for review June 30, 2004. Revised manuscript received October 3, 2004. Accepted October 5, 2004. This research work was financially supported by National Natural Science Foundation of China (Grant 20175036 and 20235020).

JF0489318