Research Article

Rapid Recognition of Geoherbalism and Authenticity of a Chinese Herb by Data Fusion of Near-Infrared Spectroscopy (NIR) and Mid-Infrared (MIR) Spectroscopy Combined with Chemometrics

Haiyan Fu,¹ Qiong Shi ,¹ Liuna Wei,¹ Lu Xu ,² Xiaoming Guo,¹ Ou Hu,¹ Wei Lan,¹ Shunping Xie,³ and Tianming Yang ¹

¹The Modernization Engineering Technology Research Center of Ethnic Minority Medicine of Hubei Province, College of Pharmacy, South-Central University for Nationalities, Wuhan 430074, China

²College of Material and Chemical Engineering, Tongren University, Tongren 554300, Guizhou, China

³State Key Laboratory Breeding Base of Green Chemistry-Synthesis Technology, College of Chemical Engineering, Zhejiang University of Technology, Hangzhou 310032, China

Correspondence should be addressed to Lu Xu; lxchemo@163.com and Tianming Yang; tmyang@mail.scuec.edu.cn

Received 30 May 2018; Revised 14 December 2018; Accepted 28 March 2019; Published 30 April 2019

Academic Editor: Massimo Lucarini

Copyright © 2019 Haiyan Fu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fourier transform near-infrared (NIR) spectroscopy and mid-infrared (MIR) spectroscopy play important roles in all fingerprint techniques because of their unique characteristics such as reliability, versatility, precision, and ease of measurement. In this paper, a supervised pattern recognition method based on the PLSDA algorithm by NIR and the NIR-MIR fusion spectra has been established to identify geoherbalism of *Angelica dahurica* from different regions and authenticity of *Corydalis yanhusuo* W. T. Wang. Comparing principle component analysis (PCA) cannot successfully identify geographical origins of *Angelica dahurica*. Linear discriminant analysis (LDA) also hardly distinguishes those origins. Furthermore, the PLSDA model based on the data fusion of NIR and IR was more accurate and efficient. But, the identification of authenticity of *Corydalis yanhusuo* W. T. Wang was still inaccurate in the PLSDA model. Consequently, data fusion of NIR-MIR original spectra combined with moving window partial least-squares discriminant analysis was firstly used and showed perfect properties on authenticity and adulteration discrimination of *Corydalis yanhusuo* W. T. Wang. It indicated that data fusion of NIR-MIR spectra combined with MWPLSDA could be considered as the promising tool for rapid discrimination of the geoherbalism and authenticity of more Chinese herbs in the future.

1. Introduction

Herbal medicines are of effective pharmacological functions, low toxicity, and less side effects to human body, so they have been widely used all over the world [1–3]. However, herbal medicines with different geographical origins have different chemical compositions and pharmacological activities [4, 5]. In addition, the processing of herbal medicines often removes morphological properties of species, and some herbal medicines at high cost are often the subject of fraudulent practices by replacing them with ones at low cost [6, 7], which may lead to an unfair competition in the pharmaceutical and harm the interest of consumers. Thus, the quality analysis method of

herbal medicines to distinguish the origins is an important concern for consumers [8–10]. Traditional methods such as high-performance liquid chromatography and mass spectroscopy are time-consuming, expensive, and laborious and have to be performed by highly trained technicians [11, 12]. Therefore, a rapid, more accurate, and sensitive identification method is required to determine herbal medicines.

Most studies focused on specific pharmacological ingredients in herbal medicines; however, the pharmacological activity of herbs is the result of the interaction of all ingredients rather than specific ingredients. Therefore, the specific ingredients could not be used as a proper criterion for characterization of the overall quality of the herbs

[13, 14]. Fourier transform mid-infrared (MIR) [15] and near-infrared (NIR) [16, 17] techniques are efficient tools for studying food and pharmaceutical quality control because of their fast and nondestructive analytical characteristics. For example, Zhu et al. used FT-IR and 2DCOS-IR methods to discriminate the cultivated *Codonopsis lanceolata* in different ages [18], and in the research done by Gayo and Hale, near-infrared spectroscopy was applied to detect and quantify the species authenticity in Crabmeat [19]. By studying the characteristic information of the spectra, different types of samples can be accurately distinguished. Nonetheless, the information obtained from by NIR spectra may be difficult to interpret directly because of the highly overlapped spectra. Although MIR spectra provide some significant differences of spectral peaks, they do not give abundant chemical and structural information of samples like NIR spectra. Therefore, establishing effective and robust chemometric methods has been extensively concerned [20, 21]. For example, Woo's team used Mahalanobis distance and discriminant PLS2 combined with NIR spectroscopy to discriminate herbal medicines according to geographical origins, but there are only two different classes from different geographical origins [22]. Frizon et al. used the PLS in determination of total phenolic compounds in yerba mate and predicted total phenolics with associated errors of 12% [23]. Liu et al. studied on the differentiation of the root of various ginseng by FT-IR and two-dimensional correlation IR spectroscopy, and the cluster analysis demonstrated that the three kinds of ginseng can be distinguished clearly from each other but with an exception [24]. PCA is a multivariate statistical technique that reduces the multidimensionality of data while minimizing information loss [25]. LDA can establish linear transformations to find the best boundary and achieve maximum separation between classes by constructing discriminant functions [26]. From another aspect, as a powerful pattern recognition method, PLSDA has successfully been applied to solve classification problems in many scientific fields [27, 28]. Furthermore, a global model with moving window partial least-squares (MWPLS) [29, 30] like other variable selection methods, MWPLSDA was successfully applied to spectra interval selection for calibration problems, and desirable results were obtained [31]. A subset of the whole wavelengths to develop the calibration model, the wavelengths carrying serious heteroscedastic noises, and especially the spectral ranges contaminated by external factors are excluded from the model, and wavelength ranges sensitive only to the chemical compositions of the samples are selected to develop a simplified yet stable calibration model.

Sometimes, it is difficult to discriminate the origins of herbal medicines only through the pattern recognition method by single NIR or MIR spectra [32] combined with chemometrics, and it is necessary to extract from the data fusion of NIR and MIR spectroscopy [33]. There is abundant information related to combinatory MIR and NIR spectroscopy coupled with chemometrics for quality control of herbal medicines.

In this study, different supervised pattern recognition algorithms including principal component analysis

(PCA), linear discriminant analysis (LDA), and partial least squares discriminant analysis (PLSDA) with raw NIR spectra were used to discriminate five different geographical origins of *Angelica dahurica*. Moreover, moving window partial least-squares discriminant analysis (MWPLSDA) and the fusion spectra variables evaluate authenticity and adulteration of *Corydalis yanhusuo* W. T. Wang. The result shows that PLSDA model is of great performance than PCA and LDA in identifying geographical origins of herbal medicines. In addition, the full spectra information fused by NIR and MIR combined with MWPLSDA showed the best ability in determination of authenticity of herbal medicines. This method provides pattern recognition models that can be applied in geographical origin discrimination or authenticity and adulteration recognition at the same time and can further be widely used in various herbal medicines.

2. Material and Methods

2.1. Collection of Raw Materials. A total of 50 *Angelica dahurica* samples from five geographical origins (Hebei, Anhui, Yunnan, Zhejiang, and Sichuan) were purchased from the Derentang pharmacy, and each region included 10 batches. Besides, two kinds of authentic *Corydalis yanhusuo* W. T. Wang (Zhejiang) were purchased from the Derentang pharmacy and the Kangderuiqi flagship store, while three kinds of adulterations *Corydalis decumbens* (Thunb.) Pers., *Typhonium flagelliforme* (Lodd.) Blume, and *Dioscorea opposita* (Thunb.) were, respectively, collected from Anhui, Jiangsu, and Fujian, and the aforementioned five samples for identification of adulteration were collected in 10 batches.

2.2. Apparatus. The following apparatuses were used: Nicolet 6700 FT-IR, OMNIC 8.2 spectral collecting software (Thermo Fisher Scientific Inc., USA); Antaris II FT-NIR spectrometer, RESULT 3.0 spectral collecting software (Thermo Electron Co., USA); DZF-6021 vacuum oven (Shanghai YIHENG Technical Co., Ltd); and FW135 herbal grinder (Tianjin Taisite Instrument Co., LTD).

2.3. Methods of Sample Measurement and Data Preprocessing by NIR and MIR. All samples used in NIR were crushed with the grinder, sieved into fine powders by a 200 mesh sieve, then vacuum-dried at 60°C for 24 hours, and stored in a dryer spare. The sample powder was placed directly into the quartz cup, and the air background was subtracted. Spectra were collected by integrating sphere diffuse reflectance with the collecting region at 4000–10000 cm^{-1} and a resolution of 8 cm^{-1} . Data processing was performed using the average of the five measured spectra for each sample. In total, 250 spectra from different geographic origins (5 samples $\times$ 10 batched $\times$ 5 measurements) were obtained. And 250 spectra were discriminated for the authenticity and adulteration of *Corydalis yanhusuo* W. T. Wang.

2.4. Method of Chemometrics. PCA, LDA, PLSDA, and MWPLSDA methods were written and performed through a Matlab 2010a (MathWorks, Natick, MA, USA). All pre-processing in those chemometrics only used the original spectra. PLSDA is based on the simultaneous decomposition response matrix and the class matrix extraction factor. By arranging the extraction factors in order of their correlation, the virtual vectors are encoded to represent different classes, wherein the virtual vector f_j encoded for the j th element is 1. The other elements are 0 for the j th class, and then each column of the response matrix is associated with the class matrix. The principle of moving window partial least-squares discriminant analysis (MWPLSDA) is that a suitable window moves along the full spectral interval according to our past study [34, 35]. In MWPLSDA, a suitable window of width H is constructed and moved along the entire spectrum to select useful wavelength intervals, and then the selected spectral spacing is used to construct the PLSDA model. The principle of MWPLSDA is based on the virtual setting of a window, which contains the number of variables from the first wavelength to the end of $(i + H - 1)$ wavelength. A series of submatrices are obtained continuously by moving the window. According to the variables in the moving window, a series of PLS submodels are constructed. Then, according to the principle of least residual square (SSR), the interval of measurement matrix with smaller classification error and latent variable is selected as the final MWPLSDA model.

3. Results and Discussion

3.1. Geographical Origin Discrimination of *Angelica dahurica* by NIR. In order to analyze the five different samples more effectively, the classical quick data analysis, and non-destructive analytical technique, NIR was used in the measurement. The average NIR spectra of each group are displayed to reflect the overlay in Figure 1. The peaks located at 8319 cm^{-1} might be associated with the second overtone of C-H, O-H, and N-H stretching modes and those around 6780 cm^{-1} were caused by the C-H deformation vibration of CH_3 . Due to the second overtone of the C=O stretching vibration, bands at 5164 cm^{-1} emerge and the C-H combination and second overtone can be seen at $4200\text{--}4300\text{ cm}^{-1}$. However, owing to the overlaps and the systematic noise in NIR spectra, chemometric methods were required to extract useful information for the recognition of *Angelica dahurica* samples. Herein, three classical chemical pattern recognition methods using principal component analysis (PCA), linear discriminant analysis (LDA), and partial least squares discriminant analysis (PLSDA) models were associated with virtual coding of original NIR spectral variables of different sample sets. The 250 sample spectra of five different *Angelica dahurica* samples were randomly divided into a training set and a prediction set (Table 1). The model was built using the training set, the number of latent variables (LVs) was determined to be 5 by eightfold cross-validation using the prediction set, and the discrimination results were analyzed for comparison.

Firstly, as a common method in the chemical pattern recognition which is mainly used for classification in the analytical processes of Chinese herbal identification, principal component analysis (PCA) is one of the most classic high-dimensional methods, which reduces the high-dimensional data of FTNIR and converts 1557 raw variables into fewer new principal components. PCA used fewer principal component features to represent the original features of the sample by decomposing the sample matrix in the training set and prediction. Based on the PCA technology, the vector scores of the training and prediction sets of the aforementioned samples are reflected in Figure 1(b), and all samples from five different geographic origins in the training and prediction sets could not be clearly distinguished, but these samples were with same shape. This phenomenon could be attributed to small differences in the chemical properties reflected in its geographical origin. The results demonstrated that the PCA method can effectively reduce and extract fewer new variables from the original high-dimensional data, but the restoration process also leads to loss some information useful for sample differentiation.

Other than looking for the vector space that can best describe the original data like PCA, linear discriminant analysis (LDA) is a linear discriminant function based on input response variables for searching linear transformations and dimensionality reduction. The axes of interest for LDA can maximize the distinction between classes, projecting feature spaces (multidimensional samples in the dataset) into smaller dimensional k -dimensional subspaces while maintaining information that distinguishes categories. Figure 1(c) shows the vector scores of the first two latent variables based on the LDA model for the training and prediction sets of samples. It clearly distinguished samples from different geographical origins in the training set, while those in the prediction set were not clearly distinguished. The result may be due to some special requirements of the LDA model, of which at least one of the needs to be nonsingular. In addition, when the so-called outlier class dominates in estimating the scattering matrix, the LDA model cannot guarantee that the optimal subspace is found [36]. Furthermore, PLSDA can reduce the effects of multicollinearity between variables, and it can simultaneously decompose the extraction factors of the prediction measurement matrix and the class matrix and arrange them according to the correlation between them. Five different geographical sources of *Angelica dahurica* are identified based on the maximum virtual coding position of the NIR spectral data. In order to optimize the predictive power of the PLSDA model and simplify the complexity of the PLSDA model, we selected the number of latent variables (LV) as 5 by 8-fold cross-validation. Figure 1(d) shows the plots of dummy codes of the training and prediction sets for five group samples of different geographic origins. Table 1 shows the virtual code attribution maps for the training and prediction sets of the original spectra in the PLSDA model. We encode five sets of samples into f_1 (1, 0, 0, 0, 0), f_2 (0, 1, 0, 0, 0), f_3 (0, 0, 1, 0, 0), f_4 (0, 0, 0, 1, 0), and f_5 (0, 0, 0, 0, 1), respectively, according to the position of the largest virtual code. As shown in Figure 1(d), all training and prediction samples belonging to

FIGURE 1: The raw NIR spectra of five different origins of *Angelica dahurica* (a) and the results by PCA (b), LDA (c), and PLSDA (d).

TABLE 1: A detailed list for the training set and the prediction set of five different kinds of *Angelica dahurica* samples.

Groups code	f1	f2	f3	f4	f5
Sample group	1	2	3	4	5
Symbol	.	○	×	+	*
PCA, LDA, or PLSDA training set	1st–37th	38th–73rd	74th–105th	106th–138th	139th–171st
PCA, LDA, or PLSDA prediction set	1st–13th	14th–27th	28th–45th	46th–62nd	63rd–79th

all groups of *Angelica dahurica* by original NIR spectra combined with PLSDA were identified accurately with a perfect recognition rate of 100%. This demonstrated that the PLSDA model successfully discriminates *Angelica dahurica* samples of different geographic origins. This further revealed that NIR spectroscopy combined with PLSDA method can be used to identify herbal medicines more rapidly, effectively, and reliably than the traditional ones.

3.2. Authenticity and Adulteration Discrimination of *Corydalis yanhusuo* W. T. Wang by NIR and Combinatory of NIR. Herbal medicine processing often removes morphological properties of species, which leads to failure of distinguishing

one type from another. For this reason, NIR spectra were used to discriminate the authenticity and adulteration of *Corydalis yanhusuo* W. T. Wang. As is shown Figure 2(a), the peaks around 6826 cm^{-1} were due to the C-H deformation vibration of CH_3 . Due to the C-H first overtone of $-\text{CH}_2-$ groups, bands at 5800 and 5600 cm^{-1} were observed and bands at 5172 cm^{-1} were the second overtone of the C=O stretching vibration. Furthermore, the C-H combination and second overtone can be seen at $4200\text{--}4300\text{ cm}^{-1}$. The seriously overlapped raw spectra hardly reflect the differences between samples. Thus, PCA technology and LDA and PLSDA models were used to relate the dummy code for the full original and preprocessing spectral variables. 250 sample spectra of two kinds of authenticity,

FIGURE 2: Raw NIR (a) and MIR (b) spectra of two kinds of *Corydalis yanhusuo* W. T. Wang, *Corydalis decumbens* (Thunb.) Pers., *Typhonium flagelliforme* (Lodd.) Blume, and *Dioscorea opposita* (Thunb.).

Corydalis yanhusuo W. T. Wang 1 and 2, and three kinds of adulteration, *Corydalis decumbens* (Thunb.) Pers. (3), *Typhonium flagelliforme* (Lodd.) Blume (4), and *Dioscorea opposita* (Thunb.) (5) were randomly divided into the training set and the prediction set (Table 2). However, both PCA technology and LDA model failed to show the correct results in prediction sets for five different groups by NIR (not shown here). Thus, PLSDA was adopted for the identification of authentic *Corydalis yanhusuo* W. T. Wang.

In our work, all training and prediction samples were correctly identified except for the two samples in the training set (34th and 88th) and the two samples in the prediction set (35th and 82nd). The 34th sample in the training set of f2 is incorrectly discriminated as f1, and the 84th sample of f5 is erroneously classified as f2. Furthermore, the 35th sample in the prediction set of f2 is incorrectly assigned as f3, and the 82nd sample in the prediction set belonging to f5 is incorrectly classified as f2. It may account for the useless information of some spectral variables. The total correction rate was 97.94% on the test set in PLSDA models. On the other hand, MIR spectroscopy provides more specific and distinct absorption bands than NIR spectroscopy. As is shown in Figure 2(b), the band centered at 2931 cm⁻¹ is due to a stretching vibration of aliphatic C-H in terminal CH₃ groups. The strong single peak of the C=O stretching vibration of ketone groups is observed at about 1635 cm⁻¹,

whereas the band centered at 1250 cm⁻¹ is due to the antisymmetric stretching vibrations of =C-O-C.

In order to better identify the origin of Chinese herbal medicines, we combined the mid-infrared spectrum with the near-infrared spectrum to obtain fusion spectra with more abundant sample information (Figure 3). The PLSDA was also applied to relate the dummy code for the full fused spectral variables.

As is shown in Figure 3(a), only the 17th sample in the prediction set of f1 was misclassified as f2 in fusion spectra (Table 3). It suggested that fusion spectra of NIR and MIR spectroscopy combined with PLSDA has better use in authenticity and adulteration discrimination of *Corydalis yanhusuo* W. T. Wang. But, it also failed to get 100% predictive accuracy.

In MWPLSDA, the appropriate window with H width is constructed, and the useful wavelength range is selected by moving the whole spectrum. Then, all the selected windows are constructed into the PLSDA model. Finally, according to the minimum SSR principle of the MWPLSDA algorithm, the feature differences among the five samples are extracted. As shown in Figure 3(b), when the window size is 20, the optimum number of potential variables in the MWPLSDA model is 12. Figure 3(b) shows that when the number of variables in the fusion spectrum of NIR and MIR is 140-200, 750-930, and 1250-1380, the SSR is the smallest. At this time, MWPLS-DA

TABLE 2: A detailed list of two kinds of *Corydalis yanhusuo* W. T. Wang, *Corydalis decumbens* (Thunb.) Pers., *Typhonium flagelliforme* (Lodd.) Blume, and *Dioscorea opposita* (Thunb.)

Groups code	f1	f2	f3	f4	f5
Sample group	1	2	3	4	5
NIR training set	1st–33rd	34th–66th	67th–89th	90th–121st	122nd–154th
NIR prediction set	1st–17th	18th–34th	35th–61st	62nd–79th	80th–96th
Fusion training set	1st–33rd	34th–66th	67th–89th	90th–121st	122nd–154th
Fusion prediction set	1st–17th	18th–34th	35th–61st	62nd–79th	80th–96th

FIGURE 3: Data fusion of NIR and MIR spectroscopy for authenticity or adulteration discrimination of *Corydalis yanhusuo* W. T. Wang (a) and the residue line obtained by MWPLSDA for the training sets (b).

TABLE 3: Forecasted results of NIR-MIR fusion spectra of two kinds of *Corydalis yanhusuo* W. T. Wang, *Corydalis decumbens* (Thunb.) Pers., *Typhonium flagelliforme* (Lodd.) Blume, and *Dioscorea opposita* (Thunb.) by PLSDA and MWPLSDA.

Sample sets	Discrimination	Combinatory spectra	
		PLSDA	MWPLSDA
Training	Accuracy (%)	100	100
	Error number	0	0
Prediction	Accuracy (%)	98.97	100
	Error number	1	0

FIGURE 4: Continued.

FIGURE 4: Assigned plots of dummy codes of the training set and prediction set for the raw fusion MIR-NIR spectra in the PLSDA model (a) and the MWPLSDA (b) model.

was benefited in the selection of combined informative fusion spectral regions of $660\text{--}950\text{ cm}^{-1}$, $3550\text{--}4400\text{ cm}^{-1}$, and $5900\text{--}6600\text{ cm}^{-1}$ for classification modeling of all samples and yielded the results better than that obtained from a partial least squares-discrimination analysis (PLS-DA) model built by using the whole NIR spectral region. As shown in Figure 4(b) and Table 3, all training and prediction set samples were correctly predicted by 100%. MWPLSDA can improve sample classification accuracy by eliminating useless information variables and noncomponent-related factors.

4. Conclusions

Supervised pattern recognition methods based on PLSDA and MWPLSDA algorithms by NIR and the data fusion of both NIR and MIR has been established to study *Angelica dahurica* and to identify the authenticity of *Corydalis yanhusuo* W. T. Wang. In addition, it was clarified from the results that other than PCA and LDA that can merely have well learning performance and do well in the training sets, the PLSDA model shows good performance in the area of identification of *Angelica dahurica* or *Corydalis yanhusuo* W. T. Wang and can be employed in the analysis of the geographical origins of *Angelica dahurica* and the authenticity or adulteration of *Corydalis yanhusuo* W. T. Wang. Furthermore, the full spectrum information of NIR and MIR spectroscopy combined with MWPLSDA performed much better than the single NIR spectra or PLSDA model and demonstrated an unparalleled ability of herbal medicine discrimination. This new recognition method provided a promising approach for the identification of herbal medicines widely.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (21576297, 21776321, and 21665022), the Talented Youth Cultivation Program from the “Fundamental Research Funds for the Central University,” South-Central University for Nationalities (No. CRZ18002), and the Key Research Program (Nos. 2015ZD001 and 2015ZD002) from the Modernization Engineering Technology Research Center of Ethnic Minority Medicine of Hubei Province. Lu Xu is financially supported by Guizhou Provincial Department of Science and Technology (Nos. QKHJC[2017]1186 and QKHZC[2019]2816) and the Talented Researcher Program from Guizhou Provincial Department of Education (QJHKYZ[2018]073). We also gratefully acknowledge the help of Yao Fan, Ji Yang, Li Liu, Hanyue Lan, Chuang Ni, and Yuan-Bin She.

Supplementary Materials

A figure depicting rapid recognition of geoherbals and authenticity of herbal medicines using near-infrared, mid-infrared, and data fusion spectroscopy combined with chemometric methods. (*Supplementary Materials*)

References

- [1] F.-T. Chau, H.-Y. Chan, C.-Y. Cheung, C.-J. Xu, Y. Liang, and O. M. Kvalheim, “Recipe for uncovering the bioactive components in herbal medicine,” *Analytical Chemistry*, vol. 81, no. 17, pp. 7217–7225, 2009.
- [2] M.-Y. Lee, C.-S. Seo, J.-A. Lee et al., “Anti-asthmatic effects of *Angelica dahurica* against ovalbumin-induced airway

- inflammation via upregulation of heme oxygenase-1," *Food and Chemical Toxicology*, vol. 49, no. 4, pp. 829–837, 2011.
- [3] K.-M. Ng, Z. Liang, W. Lu et al., "In vivo analysis and spatial profiling of phytochemicals in herbal tissue by matrix-assisted laser desorption/ionization mass spectrometry," *Analytical Chemistry*, vol. 79, no. 7, pp. 2745–2755, 2007.
  - [4] X. R. Liang, M. L. Ma, and W. K. Su, "Fingerprint analysis of *Hibiscus mutabilis* L. leaves based on ultra performance liquid chromatography with photodiode array detector combined with similarity analysis and hierarchical clustering analysis methods," *Pharmacognosy Magazine*, vol. 9, no. 35, pp. 238–243, 2013.
  - [5] H. Zhu, Y. Wang, H. Liang, Q. Chen, P. Zhao, and J. Tao, "Identification of *Portulaca oleracea* L. from different sources using GC-MS and FT-IR spectroscopy," *Talanta*, vol. 81, no. 1–2, pp. 129–135, 2010.
  - [6] Q. Jiang, W. Gao, X. Li, and J. Zhang, "Characteristics of native and enzymatically hydrolyzed *Zea mays* L., *Fritillaria ussuriensis* Maxim. and *Dioscorea opposita* Thunb. Starches," *Food Hydrocolloids*, vol. 25, no. 3, pp. 521–528, 2011.
  - [7] M. S. M. Preto, M. I. B. Tavares, P. J. O. Sebastião, and R. B. V. Azeredo, "Determination of herb authenticity by low-field NMR," *Food Chemistry*, vol. 136, no. 3–4, pp. 1272–1276, 2013.
  - [8] S. Y. K. Fong, M. Liu, H. Wei et al., "Establishing the pharmaceutical quality of Chinese herbal medicine: a provisional BCS classification," *Molecular Pharmaceutics*, vol. 10, no. 5, pp. 1623–1643, 2013.
  - [9] S. Anubala, R. Sekar, and K. Nagaiah, "Development and validation of an analytical method for the separation and determination of major bioactive curcuminoids in *Curcuma longa* rhizomes and herbal products using non-aqueous capillary electrophoresis," *Talanta*, vol. 123, pp. 10–17, 2014.
  - [10] J. Kang, L. Zhou, J. Sun, J. Han, and D.-A. Guo, "Chromatographic fingerprint analysis and characterization of furocoumarins in the roots of *Angelica dahurica* by HPLC/DAD/ESI-MSn technique," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 47, no. 4–5, pp. 778–785, 2008.
  - [11] M. Z. Končić, D. Kremer, W. Schühly, A. Brantner, K. Karlović, and Z. Kaloder, "Chemical differentiation of *Berberis croatica* and *B. Vulgaris* using HPLC fingerprinting," *Croatia Chemica Acta*, vol. 83, pp. 451–456, 2010.
  - [12] S. Su, W. Cui, W. Zhou, J.-A. Duan, E. Shang, and Y. Tang, "Chemical fingerprinting and quantitative constituent analysis of Siwu decoction categorized formulae by UPLC-QTOF/MS/MS and HPLC-DAD," *Chinese Medicine*, vol. 8, no. 1, pp. 5–20, 2013.
  - [13] D. Kumar, H. Kumar, J. R. Vedasiromoni, and B. C. Pal, "Bioassay guided isolation of α -glucosidase inhibitory constituents from *Hibiscus mutabilis* leaves," *Phytochemical Analysis*, vol. 23, no. 5, pp. 421–425, 2012.
  - [14] C. Wang, S. Wang, G. Fan, and H. Zou, "Screening of antinociceptive components in *Corydalis yanhusu* W. T. Wang by comprehensive two-dimensional liquid chromatography/tandem mass spectrometry," *Analytical and Bioanalytical Chemistry*, vol. 396, no. 5, pp. 1731–1740, 2010.
  - [15] Y. Wang, P. Wang, C. Xu et al., "Macro-fingerprint analysis-through-separation of licorice based on FT-IR and 2DCOS-IR," *Journal of Molecular Structure*, vol. 1070, pp. 1–9, 2014.
  - [16] J. Lu, B. Xiang, H. Liu, S. Xiang, S. Xie, and H. Deng, "Application of two-dimensional near-infrared correlation spectroscopy to the discrimination of Chinese herbal medicine of different geographic regions," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 69, no. 2, pp. 580–586, 2008.
  - [17] A. P. Craig, A. S. Franca, L. S. Oliveira, J. Irudayaraj, and K. Iileji, "Application of elastic net and infrared spectroscopy in the discrimination between defective and non-defective roasted coffees," *Talanta*, vol. 128, pp. 393–400, 2014.
  - [18] Y. Zhu, C.-H. Xu, J. Huang et al., "Rapid discrimination of cultivated *Codonopsis lanceolata* in different ages by FT-IR and 2DCOS-IR," *Journal of Molecular Structure*, vol. 1069, pp. 272–279, 2014.
  - [19] J. Gayo and S. A. Hale, "Detection and quantification of species authenticity and adulteration in crabmeat using visible and near-infrared spectroscopy," *Journal of Agricultural and Food Chemistry*, vol. 55, no. 3, pp. 585–592, 2007.
  - [20] L. M. Dale, A. Thewis, C. Boudry et al., "Discrimination of grassland species and their classification in botanical families by laboratory scale NIR hyperspectral imaging: preliminary results," *Talanta*, vol. 116, pp. 149–154, 2013.
  - [21] H. Yang, J. Irudayaraj, and M. Paradkar, "Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy," *Food Chemistry*, vol. 93, no. 1, pp. 25–32, 2005.
  - [22] Y.-A. Woo, H.-J. Kim, J. Cho, and H. Chung, "Discrimination of herbal medicines according to geographical origin with near infrared reflectance spectroscopy and pattern recognition techniques," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 21, no. 2, pp. 407–413, 1999.
  - [23] C. N. T. Frizon, G. A. Oliveira, C. A. Perussello et al., "Determination of total phenolic compounds in yerba mate (*Ilex paraguariensis*) combining near infrared spectroscopy (NIR) and multivariate analysis," *LWT—Food Science and Technology*, vol. 60, no. 2, pp. 795–801, 2015.
  - [24] D. Liu, Y.-G. Li, H. Xu, S.-Q. Sun, and Z.-T. Wang, "Differentiation of the root of cultivated ginseng, mountain cultivated ginseng and mountain wild ginseng using FT-IR and two-dimensional correlation IR spectroscopy," *Journal of Molecular Structure*, vol. 883–884, pp. 228–235, 2008.
  - [25] E. K. Kemsley, "Discriminant analysis of high-dimensional data: a comparison of principal components analysis and partial least squares data reduction methods," *Chemometrics and intelligent laboratory systems*, vol. 33, no. 1, pp. 47–61, 1996.
  - [26] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.
  - [27] J. Peerapattana, H. Shinzawa, K. Otsuka, Y. Hattori, M. Otsuka, and J. Peerapattana, "Partial least square discriminant analysis of mangosteen pericarp powder by near infrared spectroscopy," *Journal of Near Infrared Spectroscopy*, vol. 21, no. 3, pp. 195–202, 2013.
  - [28] H.-Y. Fu, D.-C. Huang, T.-M. Yang, Y.-B. She, and H. Zhang, "Rapid recognition of Chinese herbal pieces of *Areca catechu* by different concocted processes using Fourier transform mid-infrared and near-infrared spectroscopy combined with partial least-squares discriminant analysis," *Chinese Chemical Letters*, vol. 24, no. 7, pp. 639–642, 2013.
  - [29] Y. P. Du, Y. Z. Liang, J. H. Jiang, R. J. Berry, and Y. Ozaki, "Spectral regions selection to improve prediction ability of PLS models by changeable size moving window partial least squares and searching combination moving window partial least squares," *Analytica Chimica Acta*, vol. 501, no. 2, pp. 183–191, 2004.
  - [30] S. Fang, M. Q. Zhu, and C. H. He, "Moving window as a variable selection method in potentiometric titration multivariate calibration and its application to the simultaneous determination of ions in Raschig synthesis mixtures," *Journal of Chemometrics*, vol. 23, no. 3, pp. 117–123, 2010.

- [31] L. Xu, H. Y. Fu, C. B. Cai, and Y. B. She, "Quality degradation of Chinese white lotus seeds caused by dampening during processing and storage: rapid and nondestructive discrimination using near-infrared spectroscopy," *Journal of Analytical Methods in Chemistry*, vol. 2015, no. 12, p. 345352, 2015.
- [32] M. Casale, P. Oliveri, C. Casolino et al., "Characterisation of PDO olive oil chianti classico by non-selective (UV-visible, NIR and MIR spectroscopy) and selective (fatty acid composition) analytical techniques," *Analytica Chimica Acta*, vol. 712, pp. 56–63, 2012.
- [33] Y. Wang, M. Mei, Y. Ni, and S. Kokot, "Combined NIR/MIR analysis: a novel method for the classification of complex substances such as *Illicium verum* Hook. F. and its adulterants," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 130, pp. 539–545, 2014.
- [34] W. Luo, S. Huan, H. Fu et al., "Preliminary study on the application of near infrared spectroscopy and pattern recognition methods to classify different types of apple samples," *Food Chemistry*, vol. 128, no. 2, pp. 555–561, 2011.
- [35] H.-Y. Fu, S.-Y. Huan, L. Xu et al., "Construction of an efficacious model for a nondestructive identification of traditional Chinese medicines liuwe dihuang pills from different manufacturers using near-infrared spectroscopy and moving window partial least-squares discriminant analysis," *Analytical Sciences*, vol. 25, no. 9, pp. 1143–1148, 2009.
- [36] H. Zhao, B. Guo, Y. Wei, and B. Zhang, "Near infrared reflectance spectroscopy for determination of the geographical origin of wheat," *Food Chemistry*, vol. 138, no. 2-3, pp. 1902–1907, 2013.