



Manufacturer identification and storage time determination of “Dong’e Ejiao” using near infrared spectroscopy and chemometrics*

Wen-long LI¹, Hai-fan HAN¹, Lu ZHANG², Yan ZHANG², Hai-bin QU^{†‡1}

(¹Pharmaceutical Informatics Institute, Zhejiang University, Hangzhou 310058, China)

(²Shandong Dong’e Ejiao Co., Ltd., Liaocheng 252299, China)

[†]E-mail: quhb@zju.edu.cn

Received Aug. 3, 2015; Revision accepted Dec. 27, 2015; Crosschecked Apr. 15, 2016

Abstract: We have developed a set of chemometric methods to address two critical issues in quality control of a precious traditional Chinese medicine (TCM), Dong’e Ejiao (DEEJ). Based on near infrared (NIR) spectra of multiple samples, the genuine manufacturer of DEEJ, e.g. Dong’e Ejiao Co., Ltd., was accurately identified among 21 suppliers by the fingerprint method using Hotelling T^2 , distance to Model X (DModX), and similarity match value (SMV) as discriminate criteria. Soft independent modeling of the class analogy algorithm led to a misjudgment ratio of 6.2%, suggesting that the fingerprint method is more suitable for manufacturer identification. For another important feature related to clinical efficacy of DEEJ, storage time, the partial least squares-discriminant analysis (PLS-DA) method was applied with a satisfactory misjudgment ratio (15.6%) and individual prediction error around 1 year. Our results demonstrate that NIR spectra comprehensively reflect the essential quality information of DEEJ, and with the aid of proper chemometric algorithms, it is able to identify genuine manufacturer and determine accurate storage time. The overall results indicate the promising potential of NIR spectroscopy as an effective quality control tool for DEEJ and other precious TCM products.

Key words: Dong’e Ejiao, Near infrared spectroscopy, Manufacturer identification, Storage time identification, Quality control

<http://dx.doi.org/10.1631/jzus.B1500186>

CLC number: R917

1 Introduction

Colla corii asini (Ejiao in Chinese), which is made from the hide of *Equus animus* L., is a precious traditional Chinese medicine (TCM) widely used in China for thousands of years. As recorded in many classic monographs on TCM and in ancient poetry, Ejiao displays a great efficacy in enriching blood and staunching bleeding, being mainly used for the treatment of gynecological diseases, such as menoxenia and post-partum uterine bleeding (Gao and Zhong,

2012). It is no surprise that multiple pharmaceutical factories manufacture Ejiao under different brand names in China; however, it is well accepted that Dong’e donkey-hide glue (Dong’e Ejiao, DEEJ), with a history of use spanning more than 2500 years, is one of the brands with the highest quality. DEEJ is made from the hide of Dezhou donkey and groundwater from Dong’e county, which is located in the west of Shandong Province, China; it is currently the largest brand of Ejiao in China.

Two problems occur in the quality control of DEEJ: manufacturer identification and storage time determination. There are many manufacturers of Ejiao using different raw materials and production processes, and the quality of their products varies wildly, affecting medical efficacy and also price. However, there is no consensus on which brand is

[‡] Corresponding author

* Project supported by the National Science and Technology Major Project (No. 2011ZX09201-201-10), China

ORCID: Wen-long LI, <http://orcid.org/0000-0002-7736-9055>

© Zhejiang University and Springer-Verlag Berlin Heidelberg 2016

best, because of the complexity of the material basis of Ejiao and the relative lack of regulatory standards in the Chinese Pharmacopoeia (Ministry of Public Health of the People's Republic of China, 2015). Vendors urgently need a way of discriminating Ejiao for the purpose of brand protection and competitive advantage. There is a proverb in TCM that "new ginseng is better than old, while old Ejiao is better than new," implying that the medicinal properties of Ejiao vary with storage time, which is confirmed by clinical experience. However, storage time is often difficult to determine from product labels, so a rapid way of determining storage time is also needed.

Near infrared (NIR) spectroscopy is a rapid, information-rich, non-destructive analytical technology, which is admirably suitable for the quality control of natural products. In recent years, NIR spectroscopy has been increasingly applied in TCMs, which are used as both food and medicine. Several articles have reported the use of NIR spectroscopy for the brand-protection of agricultural products (Wang *et al.*, 2008; Latorre *et al.*, 2013; Teye *et al.*, 2014) and the determination of storage time of wine (Yu *et al.*, 2008; Fernández-Navales *et al.*, 2009; Ghasemi-Varnamkhasti and Forina, 2014).

We proposed combining NIR spectroscopy with chemometrics to address the two aforementioned issues. For the discriminant analysis of DEEJ, NIR spectroscopy-based fingerprint method was established, and three statistics, Hotelling T^2 , distance to Model X (DmodX), and similarity match value (SMV), were used for identification. For the storage time determination, partial least squares-discriminant analysis (PLS-DA) method was developed and DEEJ with different storage time was accurately classified. The methodology was found to be helpful for the quality control of DEEJ, and shows promise of addressing similar needs in other precious TCM products in future.

2 Materials and methods

2.1 Samples

The collection of representative samples is a key step in the development of the manufacturer discriminant model. In the present study, 152 Ejiao samples were collected through various channels from almost all the Ejiao manufacturers in China. To expand the scope of the established models, 36 samples of other animal glues, made from antlers (*Colla Cornus Cervi*), tortoise beetles (*Colla Carapacis et Plastris Testudinis*), and cowhide (*Oxhide Gelatin*), produced by the Shandong Dong'e Ejiao Co., Ltd. were also collected. Detailed information on all the samples is listed in Table 1.

To determine DEEJ storage time, 240 samples produced from 2005 to 2012 (30 samples for each year) were collected.

2.2 NIR spectral collection

After optimization of NIR spectral collecting conditions (detailed information can be found in the Results and Discussion section), an Antaris II Fourier transform (FT)-NIR analyzer (Thermo Scientific, Madison, WI, USA) equipped with integrating-sphere attachments, was used to collect diffuse reflectance spectra. Instrument resolution was specified at 8 cm^{-1} between 10000 cm^{-1} and 4000 cm^{-1} and each spectrum was collected by adding a total of 128 scans, using air as the reference. During the acquisition of the spectra of all samples, the environment temperature was kept at $(25\pm 0.5)\text{ }^{\circ}\text{C}$. Result workflow-based software was used to collect the NIR spectra.

2.3 Chemometrics methods

To discriminate between manufacturers, two kinds of fingerprint analysis were conducted: SMV

Table 1 Detailed information on the samples used in the research

Sample	Abbreviation	Manufacturer	Sample No.
<i>Colla Cornus Cervi</i>	LJJ	Shandong Dong'e Ejiao Co., Ltd.	1–10
<i>Colla Carapacis et Plastris Testudinis</i>	GJJ	Shandong Dong'e Ejiao Co., Ltd.	11–18
<i>Oxhide Gelatin</i>	HMJ	Shandong Dong'e Ejiao Co., Ltd.	19–36
Dong'e Ejiao	DEEJ	Shandong Dong'e Ejiao Co., Ltd.	37–98
Ejiao (<i>Colla Corii Asini</i>)	OEJ	Other 20 different manufacturers	99–188

obtained using the Gram-Schmidt algorithm and principal component analysis (PCA) with Hotelling T^2 and DModX as criteria. The thresholds for these three statistics were determined according to the 3σ principle. For storage time determination, the PLS-DA algorithm was adopted.

2.3.1 Similarity match algorithm

Consider a spectral data matrix X , in which there are m standard samples and n data points (wave-numbers) for each spectrum,

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}.$$

If the rank of X is r , the matrix can be orthogonally decomposed into score matrix T and loading matrix P through the Gram-Schmidt algorithm,

$$X = TP = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1r} \\ t_{21} & t_{22} & \cdots & t_{2r} \\ \vdots & \vdots & & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mr} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rn} \end{bmatrix}, \quad (1)$$

where $p_j = [p_{j1}, p_{j2}, \dots, p_{jn}]$ ($j=1, 2, \dots, r$) is a set of vectors orthogonal to each other. Therefore, the following formula is established,

$$T = XP^T. \quad (2)$$

If the spectral data of the sample to be evaluated is x_{new} , its score vector t_{new} can be calculated according to Eq. (2),

$$t_{\text{new}} = x_{\text{new}} P^T = [s_1, s_2, \dots, s_r]. \quad (3)$$

When $j=1$, define $e_{i1} = |s_1 - t_{i1}|^2$ ($i=1, 2, \dots, m$) and find the minimum from $e_{11}, e_{21}, \dots, e_{m1}$, written as e_{k1} , and the associated t_{k1} is written as t_1' . With the same method, calculate t_2', t_3', \dots, t_r' , when $j=2, 3, \dots, r$. The residual spectrum e can be obtained according to the following formula,

$$e = [s_1 - t_1', s_2 - t_2', \dots, s_r - t_r'] \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_r \end{bmatrix}. \quad (4)$$

The residual spectrum e represents the difference between the sample to be evaluated and the standard sample. The SMV can be obtained according to the following formula,

$$\text{SMV} = \left(1 - \frac{\|e\|}{\|x_{\text{new}}\|}\right) \times 100. \quad (5)$$

SMV can range from 0 to 100; the higher the SMV, the greater the similarity between the sample to be evaluated and the standard sample (McCarthy, 1992; Li et al., 2013).

It is important to determine the distribution of SMV because if it follows a normal distribution, its threshold can be determined according to the 3σ principle in the statistical process control (Zhang et al., 2009). If the SMV is normally distributed, the threshold can be set as the average value of SMV minus three standard deviations (SD). The SMV of the sample to be evaluated can be compared with the threshold: if it is greater than or equal to the threshold, the sample can be considered to be in conformity with the standard sample, and vice versa.

2.3.2 Principal component analysis

There is a great deal of correlated or redundant information in data obtained from NIR spectra. PCA is a favorite tool for data compression and information extraction, which can significantly aid interpretation and further analysis. Soft independent modeling of class analogy (SIMCA) is a PCA-based discriminant analysis algorithm. Hotelling T^2 and DModX are two important statistics associated with SIMCA models. Thus, SIMCA was adopted here for the discriminant analysis of DEEJ, and Hotelling T^2 and DModX were taken as the criteria for the consistency test.

2.3.3 Partial least squares-discriminant analysis

To determine the storage time of DEEJ, the PLS-DA method was used. PLS-DA is a straightforward

extension of partial least squares regression (PLSR), which combines PLSR with the discrimination power of Bayesian classification (González-Arjona *et al.*, 1999), and has been widely used in TCM research (Ortiz *et al.*, 2006; Liu *et al.*, 2008; Niu *et al.*, 2008).

All spectral pre-treatments and chemometric analyses were implemented using the TQ analyst (V8.0, Thermo Fisher, USA), the Unscrambler (V9.6, CAMO PROCESS AS, Oslo, Norway), and the SIMCA-P+ 12.0 (MKS Umetrics AB, Umeå, Sweden).

3 Results and discussion

3.1 Spectral collection

3.1.1 Spectral collecting mode

Two spectral collecting modes, an integrating-sphere mode and a fibre-optical mode, were compared. Spectra of five randomly chosen samples were collected using each of the two modes, shown in Fig. 1. It can be seen that the spectra obtained using the fibre-optical mode are marked by indentations in the waveband of 7000–4000 cm^{-1} , which may result from fibre-optical absorption. We therefore decided to collect the spectra using the integrating-sphere mode.

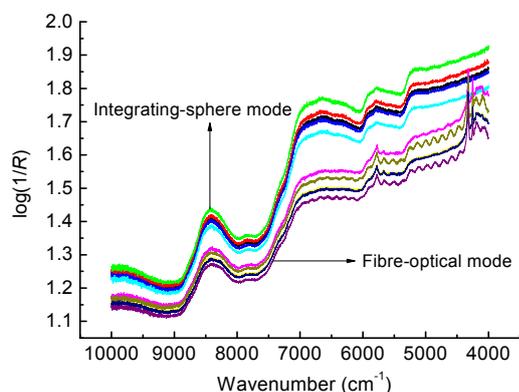


Fig. 1 NIR spectra collected using integrating-sphere mode and fibre-optical mode on five Ejiao samples

3.1.2 Optimization of resolution and number of scans

The collection of spectra with better signal-to-noise ratio and repeatability is a key step in establishing NIR spectroscopy analytical methods. We thus optimized the conditions for spectral collection, including resolution and number of scans. In order to characterize the stability and repeatability of the NIR

spectra, Ni *et al.* (2009) reported a method that used the standard variance spectrum of repeat spectra (SVSRS) under the same collecting condition to reflect the repeatability of the NIR spectra. However, there is still no quantitative index that can be used as a criterion. In this research, we adopted Mahalanobis distance for this purpose. Spectra were collected repeatedly for different resolutions and numbers of scans, and then the average Mahalanobis distance (AMD), which represents the average distance of every spectrum to the average spectrum, was calculated and compared. The spectral collection condition, which resulted in the minimum AMD, was subsequently adopted. The AMDs of the spectra gained with different resolutions and for five scans are listed in Table 2.

Table 2 Average Mahalanobis distance (AMD) of the spectra with different resolutions and for five scans

Spectral collecting condition	AMD
Spectrum resolution (cm^{-1})	
2	1.154
4	1.015
8	0.485
16	0.823
Number of repetition	
16	1.055
32	1.006
64	0.473
128	0.400

3.1.3 Character of the NIR spectra

Spectra of all the samples were collected using the conditions described above. The stacked spectra of the Ejiao from different manufacturers are shown in Fig. 2a, and the average spectra of the DEEJ samples with different storage time are shown in Fig. 2b. As can be seen, Fig. 2a indicated that the Ejiao samples from different manufacturers have similar spectral shapes, and from Fig. 2b, it was observed that the spectral absorption decreased with increased storage time. In the DEEJ spectra, four characteristic absorption peaks can be viewed and interpreted as: OH stretch+OH deformation of water (5151 cm^{-1}), N–H (2v) in CONH_2 (6735 cm^{-1}), CONH_2 specifically due to C=O hydrogen bonded to the N–H of the peptide link termed the α -helix structure (5738 cm^{-1}), N–H (3v) methyl C–H (8377 cm^{-1}) (Workman and Weyer, 2008). All these bonds are abundant in the molecular structures of water, amino acids, peptides, proteins, and fatty acids (which exist in soybean oil, a

raw material for Ejiao), so the NIR spectra can be considered to reflect the chemical information of the Ejiao sufficiently.

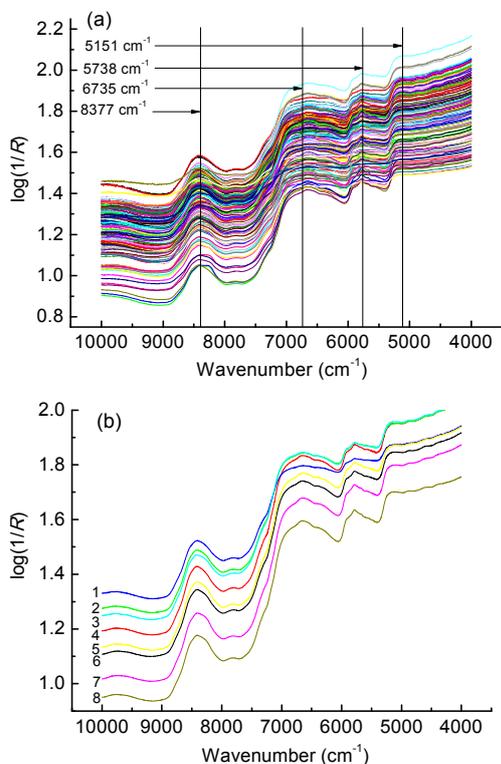


Fig. 2 NIR spectra of the Ejiao samples from different manufacturers (a) and storage time (1–8 years) (b)

3.2 Fingerprint analysis

Currently, the chromatographic fingerprint is widely used in the quality control of natural products. The analysis is often based on chromatographic separation techniques such as gas chromatography (GC), high-performance liquid chromatography (HPLC), and capillary electrophoresis (CE), combined with ultraviolet (UV), diode array detector (DAD), evaporative light scattering detector (ELSD), or mass spectrometry (MS) (Xie *et al.*, 2006; Liu *et al.*, 2007; Alaerts *et al.*, 2010; Ding *et al.*, 2011). Such fingerprints can reflect the information of a few compounds with specific chromatographic properties and characteristic detection signals, but are unsuitable for holistic chemical composition. Spectroscopic analysis, such as mid-infrared (MIR) (Sun *et al.*, 2003) and NIR (Woodcock *et al.*, 2009; Lucio-Gutiérrez *et al.*, 2011), can supplement and provide a more comprehensive analysis than the chromatographic fingerprint

alone. The NIR spectrum is sensitive to organic molecules containing molecular bonds, such as C–H, N–H, and O–H. Thus the NIR spectroscopy-based fingerprint of the sample reflects the complex attributes of Ejiao, which are related to raw material content and technical processing.

The calculation of the evaluation index is the key stage in fingerprint analysis. Cosine ratio, correlation coefficient, and Euclidean distance are three commonly used indicators. However, these parameters are often calculated by comparing the chromatogram of the new sample with the average spectrum of standard samples, which does not reflect the inherent flexibility of the standard samples. We therefore adopted two new evaluation methodologies for the NIR spectroscopy-based fingerprint analysis.

3.2.1 Similarity match model

Qu *et al.* (2006) reported the satisfactory use of the SMVs of the NIR spectroscopy-based method for the identification of donkey hide glue (Ejiao). However, the method has only been applied to the identification of authentic Ejiao samples and counterfeits, but not to the identification of Ejiao from individual manufacturers with different brands.

In the present research, 30 DEEJ samples were used as standard samples, and the similarity match model was established with the Gram-Schmidt algorithm after the spectra were pre-treated with a combination of the “MSC+2nd D+Norris derivative filter (15, 5),” that is a multiplicative scattering correction followed by a 2nd derivative and Norris derivative filter (window width=15, polynomial order=5). The SMVs of other samples were calculated and are shown in Fig. 3.

It can be seen that the remaining 32 DEEJ samples had higher SMVs, with 96.52 as the minimum value. A normality test ($\alpha=0.05$) indicated that the SMVs were normally distributed with $P=0.691 (>0.05)$. We therefore set the threshold SMV at $97.91-3 \times 0.89=95.24$. Samples from other manufacturers possessed lower SMVs, which were distributed in a range between 77.38 and 97.97. The maximum SMVs reached for the three other animal glues (*Colla Cornus Cervi*, *Colla Carapacis et Plastris Testudinis*, *Oxhide Gelatin*) were 91.79, 91.35, and 96.73, respectively. Although the SMVs vary with the different samples, the DEEJ samples cannot be completely

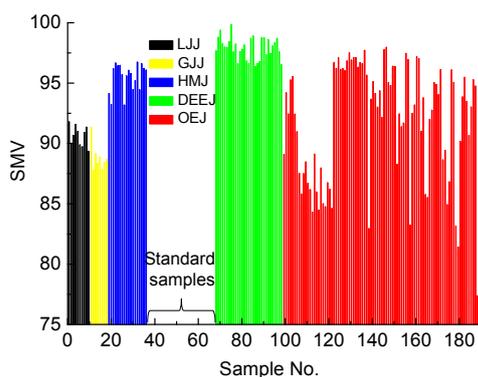


Fig. 3 Similarity match values (SMVs) of the validation samples

LJJ, black; GJJ, yellow; HMJ, blue; DEEJ, green; OEJ, red (Note: for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article). The same for the following Figs. 4 and 5

differentiated from other Ejiao samples or other kinds of animal glue. From the SMV calculations, it can also be seen that samples from the same manufacturer can vary greatly between batches, indicating a relatively low production level. A different data processing method is required to distinguish the DEEJ.

3.2.2 SIMCA model

SIMCA is a supervised classification technique, which describes each class separately in a principal component (PC) space. New measurements are considered to belong to the class if their Euclidean distance towards the constructed PC space is not significantly larger than the Euclidean distance of the class objects towards their PC space (de Maesschalck *et al.*, 1999). In our research, 30 DEEJ samples and 45 OEJ (Ejiao (*Colla Corii Asini*)) samples were used as the calibration set to establish the PC space and the corresponding Euclidean distance. Other samples were used as the validation set to challenge the established model. The optimum PC number ($n=4$) was determined using leave-one-out (LOO) cross validation. Fig. 4 gives the t_1-t_2 scattergram of the established SIMCA model. Among the 113 validation samples, three other kinds of animal glue were accurately distinguished; only 2 DEEJ samples and 5 other OEJ samples were incorrectly categorized, the misjudgment rate being 6.2%.

3.2.3 Evaluation by Hotelling T^2 and DModX

After the PC spaces were established with 30 DEEJ samples randomly selected for the fingerprint

analysis, Hotelling T^2 and DModX were used to evaluate quality consistency and fast recognition of abnormal samples (Setarehdan *et al.*, 2002; Kourti, 2005; Chiang and Colegrove, 2007). Fig. 5 gives the Hotelling T^2 and DModX control charts of the samples. It can be concluded that all the DEEJ samples are kept within the control limit and also the warning limit. Several other samples also have Hotelling T^2 or DModX values lower than the control limit.

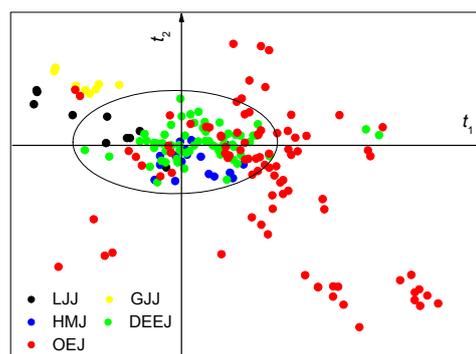


Fig. 4 t_1-t_2 scattergram of the established SIMCA model

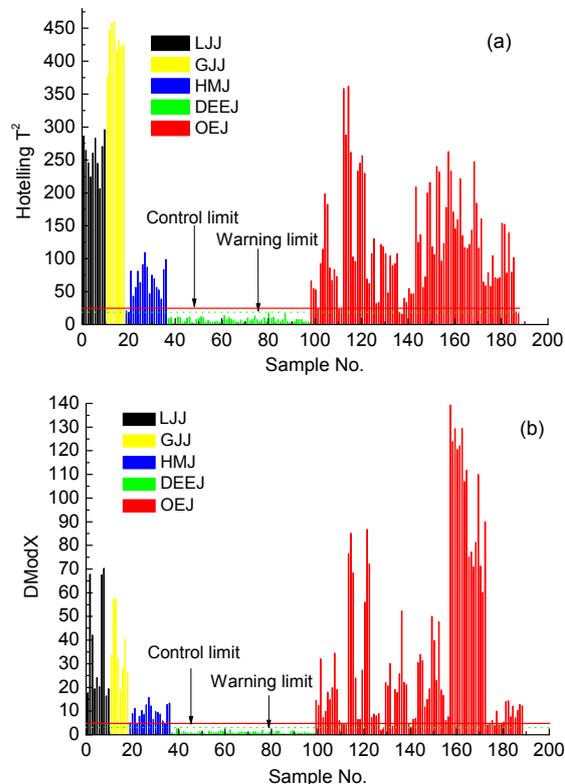


Fig. 5 Hotelling T^2 (a) and DModX (b) control charts of the samples

All the above mentioned methods indicate that some OEJ samples are close to the authentic DEEJ with similar statistics, but it is interesting to note that if we use SMV, Hotelling T^2 , and DModX as discriminate criteria simultaneously, no other samples will be recognized as authentic DEEJ, which indicates that the NIR spectroscopy-based fingerprint can be used as a brand-protection technology for DEEJ.

3.3 Storage time determination

3.3.1 Division of the sample sets and spectral data preprocessing

About 20 samples a year from each of the last 8 years were randomly selected to construct the calibration set ($n=163$), the remaining samples being used as the validation set ($n=77$). The outliers were detected using Dixon's test: spectral data were processed with PCA, and the Mahalanobis distance (MD) between the average spectrum and each spectrum was calculated with the first few PCs. Samples were then sorted in descending order according to MD value, the samples with extreme MD values being considered as outliers. Using Dixon's test, 6 samples were removed from the calibration set as outliers.

Different wavenumber ranges and spectral pretreatment methods were compared in accordance with the predictive ability of corresponding calibration models. The waveband of $8664.36\text{--}5066.06\text{ cm}^{-1}$ and the pretreatment combination of "MSC+2nd D+ Norris derivative filter (15, 5)" were chosen.

3.3.2 Calibration and validation of the PLS-DA model

After the division of sample sets, waveband selection, and spectral data preprocessing, the spectral data matrix of the calibration set was used for the PLS-DA modeling, and based upon the LOO cross validation, an optimal model was built on the first 4 PCs, which have cumulative variance contribution rates of 99.6%. Among 77 samples of the validation set used to evaluate the established model, 65 samples were predicted correctly, giving an accuracy rate of 84.4%, and the deviations of the other 12 samples were all within ± 1 year. The correlation coefficients of the calibration set and the validation set reached 0.9736 and 0.9707, respectively, which indicated that the established model fits satisfactorily. Correlation plots of actual values versus NIR predictions are shown in Fig. 6.

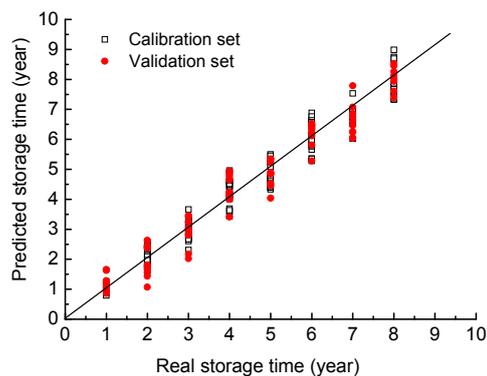


Fig. 6 Correlation plots of the real storage time and the NIR predictions of DEEJ

In conclusion, the quality of DEEJ with different storage time can be determined by NIR spectroscopy. The storage time difference has confirmed the rationality of the proverb in the TCM theory: "new ginseng is better than old, while old Ejiao is better than new." We speculate that such a difference originates from degradation of the collagen protein as well as polymerization of amino acids in DEEJ, which results in DEEJ's characteristic peptide with strong pharmacological activity. However, the reaction process often takes a long time to complete, which also provides evidence for why there was no significant difference between the samples produced in neighboring years. The change in DEEJ with different storage time is reflected in the NIR spectra. To verify this speculation, we are now conducting research using size exclusion chromatography, which can separate collagen protein, peptides, and amino acids effectively.

4 Conclusions

NIR spectra can comprehensively provide quality information about DEEJ, addressing two important issues: discriminant analysis among different manufacturers and storage time determination. The use of NIR spectroscopy and chemometric algorithms indicates that NIR spectroscopy can be used as a brand protection tool for DEEJ, and the proposed methods may be equally applicable to the quality control of other precious TCM products.

Acknowledgements

We sincerely thank the collaborators from Shandong Dong'e Ejiao Co., Ltd., Liaocheng, China, for supporting this work and for permission to publish the above results.

Compliance with ethics guidelines

Wen-long LI, Hai-fan HAN, Lu ZHANG, Yan ZHANG, and Hai-bin QU declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Alaerts, G., Merino-Arévalo, M., Dumarey, M., et al., 2010. Exploratory analysis of chromatographic fingerprints to distinguish rhizoma chuanxiong and rhizoma ligustici. *J. Chromatogr. A*, **1217**(49):7706-7716. <http://dx.doi.org/10.1016/j.chroma.2010.10.010>
- Chiang, L.H., Colegrove, L.F., 2007. Industrial implementation of on-line multivariate quality control. *Chemometr. Intell. Lab.*, **88**(2):143-153. <http://dx.doi.org/10.1016/j.chemolab.2007.02.005>
- de Maesschalck, R., Candolfi, A., Massart, D.L., et al., 1999. Decision criteria for soft independent modelling of class analogy applied to near infrared data. *Chemometr. Intell. Lab.*, **47**(1):65-77. [http://dx.doi.org/10.1016/S0169-7439\(98\)00159-2](http://dx.doi.org/10.1016/S0169-7439(98)00159-2)
- Ding, Y., Wu, E.Q., Liang, C., et al., 2011. Discrimination of cinnamon bark and cinnamon twig samples sourced from various countries using HPLC-based fingerprint analysis. *Food Chem.*, **127**(2):755-760. <http://dx.doi.org/10.1016/j.foodchem.2011.01.011>
- Fernández-Novales, J., López, M.I., Sánchez, M.T., et al., 2009. Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, winemaking, and aging of white and red wines. *Food Res. Int.*, **42**(2):285-291. <http://dx.doi.org/10.1016/j.foodres.2008.11.008>
- Gao, X., Zhong, G., 2012. Science of Chinese Materia Medical, 2nd Ed. People's Medical Publishing House, Beijing, China, p.1756-1761 (in Chinese).
- Ghasemi-Varnamkhasi, M., Forina, M., 2014. NIR spectroscopy coupled with multivariate computational tools for qualitative characterization of the aging of beer. *Comput. Electron. Agric.*, **100**:34-40. <http://dx.doi.org/10.1016/j.compag.2013.10.001>
- González-Arjona, D., López-Pérez, G., González, A.G., 1999. Performing procrustes discriminant analysis with HOLMES. *Talanta*, **49**(1):189-197. [http://dx.doi.org/10.1016/S0039-9140\(98\)00354-3](http://dx.doi.org/10.1016/S0039-9140(98)00354-3)
- Kourtí, T., 2005. Application of latent variable methods to process control and multivariate statistical process control in industry. *Int. J. Adapt. Control Signal. Pr.*, **19**(4): 213-246. <http://dx.doi.org/10.1002/acs.859>
- Latorre, C.H., Crecente, R.M.P., Martín, S.G., et al., 2013. A fast chemometric procedure based on NIR data for authentication of honey with protected geographical indication. *Food Chem.*, **141**(4):3559-3565. <http://dx.doi.org/10.1016/j.foodchem.2013.06.022>
- Li, W., Cheng, Z., Wang, Y., et al., 2013. Quality control of *Lonicerae Japonicae Flos* using near infrared spectroscopy and chemometrics. *J. Pharmaceut. Biomed. Anal.*, **72**(2):33-39. <http://dx.doi.org/10.1016/j.jpba.2012.09.012>
- Liu, A.H., Lin, Y.H., Yang, M., et al., 2007. Development of the fingerprints for the quality of the roots of *Salvia miltiorrhiza* and its related preparations by HPLC-DAD and LC-MSⁿ. *J. Chromatogr. B*, **846**(1-2):32-41. <http://dx.doi.org/10.1016/j.jchromb.2006.08.002>
- Liu, L., Cozzolino, D., Cynkar, W.U., et al., 2008. Preliminary study on the application of visible-near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. *Food Chem.*, **106**(2):781-786. <http://dx.doi.org/10.1016/j.foodchem.2007.06.015>
- Lucio-Gutiérrez, J.R., Coello, J., Maspoch, S., 2011. Application of near infrared spectral fingerprinting and pattern recognition techniques for fast identification of *Eleutherococcus senticosus*. *Food Res. Int.*, **44**(2):557-565. <http://dx.doi.org/10.1016/j.foodres.2010.11.037>
- McCarthy, W.J., 1992. TQ Analyst User's Guide. Thermo Nicolet Corp., Madison, WI, USA, p.46-47.
- Ministry of Public Health of the People's Republic of China, 2015. Pharmacopoeia of the People's Republic of China. China Pharmaceutical Technology Press, Beijing, China, p.189-190 (in Chinese).
- Ni, L.J., Zhang, L.G., Xie, J., et al., 2009. Pattern recognition of Chinese flue-cured tobaccos by an improved and simplified K-nearest neighbors classification algorithm on near infrared spectra. *Anal. Chim. Acta*, **633**(1):43-50. <http://dx.doi.org/10.1016/j.aca.2008.11.044>
- Niu, X.Y., Yu, H.Y., Ying, Y.B., 2008. The application of near-infrared spectroscopy and chemometrics to classify Shaoxing wines from different breweries. *Trans. ASABE*, **51**(4):1371-1376. <http://dx.doi.org/10.13031/2013.25220>
- Ortiz, M.C., Sarabia, L., Garcia-Rey, R., et al., 2006. Sensitivity and specificity of PLS-class modelling for five sensory characteristics of dry-cured ham using visible and near infrared spectroscopy. *Anal. Chim. Acta*, **558**(s1-2): 125-131. <http://dx.doi.org/10.1016/j.aca.2005.11.038>
- Qu, H.B., Yang, H.L., Cheng, Y.Y., 2006. Fast and nondestructive discrimination of donkeyhide glue by near-infrared spectroscopy. *Spectrosc. Spectr. Anal.*, **26**(1):60-62 (in Chinese).
- Setarehdan, S.K., Soraghan, J.J., Littlejohn, D., et al., 2002. Maintenance of a calibration model for near infrared spectrometry by a combined principal component analysis-partial least squares approach. *Anal. Chim. Acta*, **452**(1): 35-45. [http://dx.doi.org/10.1016/S0003-2670\(01\)01446-5](http://dx.doi.org/10.1016/S0003-2670(01)01446-5)
- Sun, S.Q., Tang, J.M., Yuan, Z.M., et al., 2003. FTIR and classification study on trueborn tuber dioscoreae samples. *Spectrosc. Spectr. Anal.*, **23**(2):258-261 (in Chinese).
- Teye, E., Huang, X., Lei, W., et al., 2014. Feasibility study on the use of Fourier transform near-infrared spectroscopy

- together with chemometrics to discriminate and quantify adulteration in cocoa beans. *Food Res. Int.*, **55**:288-293. <http://dx.doi.org/10.1016/j.foodres.2013.11.021>
- Wang, L., Liu, F., He, Y., 2008. Fast detection of white vinegar varieties and pH by Vis/NIR spectroscopy. *Spectrosc. Spectr. Anal.*, **28**(4):813-816 (in Chinese).
- Woodcock, T., Downey, G., Donnell, C.P.O., 2009. Near infrared spectral fingerprinting for confirmation of claimed PDO provenance of honey. *Food Chem.*, **114**(2):742-746. <http://dx.doi.org/10.1016/j.foodchem.2008.10.034>
- Workman, J., Weyer, L., 2008. *Practical Guide to Interpretive Near-Infrared Spectroscopy*. CRC Press, Boca Raton.
- Xie, P., Chen, S., Liang, Y., et al., 2006. Chromatographic fingerprint analysis—a rational approach for quality assessment of traditional Chinese herbal medicine. *J. Chromatogr. A*, **1112**(1-2):171-180. <http://dx.doi.org/10.1016/j.chroma.2005.12.091>
- Yu, H., Lin, H., Xu, H., et al., 2008. Prediction of enological parameters and discrimination of rice wine age using least-squares support vector machines and near infrared spectroscopy. *J. Agric. Food Chem.*, **56**(2):307-313. <http://dx.doi.org/10.1021/jf0725575>
- Zhang, F., Liu, J.S., Liu, Z.C., et al., 2009. Quality control in tobacco flavor by near-infrared transmittance spectroscopy. *Acta Tabacaria Sin.*, **15**(3):12-16.

中文概要

题目: 基于近红外光谱技术的东阿阿胶生产厂家和存放时间鉴定方法研究

目的: 对东阿阿胶的生产厂家和存放时间进行鉴定, 为东阿阿胶的品牌保护和质量控制提供技术手段。

创新点: 首次报道采用近红外光谱指纹图谱技术和相关的化学计量学方法对东阿阿胶进行质量控制。

方法: 收集有代表性的阿胶样品, 其中包括来自不同厂家的阿胶、黄明胶、龟甲胶、鹿角胶样品 188 份和来自东阿阿胶股份有限公司 2005~2012 年间生产的阿胶, 每年 30 批, 共计 240 份东阿阿胶样品。采集样品光谱, 采用 Hotelling T^2 、DModX 和相似度匹配值等统计量作为判断标准, 进行东阿阿胶品牌的鉴定; 采用偏最小二乘判别分析 (PLS-DA) 方法进行东阿阿胶存放年限的判别分析。

结论: 所建立的方法可有效区分东阿阿胶与黄明胶、龟甲胶、鹿角胶、以及其他品牌的阿胶, 可用于东阿阿胶的品牌保护; 利用 PLS-DA 可对不同存放年限的东阿阿胶进行判别分析, 结果令人满意。

关键词: 东阿阿胶; 近红外光谱; 生产厂家鉴定; 存放时间鉴定; 质量控制