

Article



A Synergetic Strategy for Brand Characterization of Colla Corii Asini (Ejiao) by LIBS and NIR Combined with Partial Least Squares Discriminant Analysis

Ziyi Xia^{1,†}, Xiaoqing Che^{2,†}, Lei Ye¹, Na Zhao³, Dongxiao Guo⁴, Yanfang Peng⁵, Yongqiang Lin^{4,*} and Xiaona Liu^{1,*}

- ¹ College of Integrated Traditional Chinese and Western Medicine, Binzhou Medical University, Yantai 264003, China
- ² Shandong Runzhong Pharmaceutical Co., Ltd., Yantai 256603, China
- ³ Key Laboratory of Xinjiang Phytomedicine Resources and Utilization in Ministry of Education, School of Pharmacy, Shihezi University, Shihezi 832002, China
- ⁴ Shandong Institute of Food and Drug Inspection, Jinan 250101, China
- ⁵ Pharmacy Faculty, Hubei University of Chinese Medicine, Wuhan 430065, China
- * Correspondence: 13864067104@163.com (Y.L.); xiaonaliu@bzmc.edu.cn (X.L.)
- + Co-first author.

Abstract: A synergetic strategy was proposed to address the critical issue in the brand characterization of Colla corii asini (Ejiao, CCA), a precious traditional Chinese medicine (TCM). In all brands of CCA, Dong'e Ejiao (DEEJ) is an intangible cultural heritage resource. Seventy-eight CCA samples (including forty DEEJ samples and thirty-eight samples from other different manufacturers) were detected by laser-induced breakdown spectroscopy (LIBS) and near-infrared spectroscopy (NIR). Partial least squares discriminant analysis (PLS-DA) models were built first considering individual techniques separately, and then fusing LIBS and NIR data at low-level. The statistical parameters including classification accuracy, sensitivity, and specificity were calculated to evaluate the PLS-DA model performance. The results demonstrated that two individual techniques show good classification performance, especially the NIR. The PLS-DA model with single NIR spectra pretreated by the multiplicative scatter correction (MSC) method was preferred as excellent discrimination. Though individual spectroscopic data obtained good classification performance. A data fusion strategy was also attempted to merge atomic and molecular information of CCA. Compared to a single data block, data fusion models with SNV and MSC pretreatment exhibited good predictive power with no misclassification. This study may provide a novel perspective to employ a comprehensive analytical approach to brand discrimination of CCA. The synergetic strategy based on LIBS together with NIR offers atomic and molecular information of CCA, which could be exemplary for future research on the rapid discrimination of TCM.

Keywords: laser-induced breakdown spectroscopy; near-infrared spectroscopy; data fusion; *Colla Corii Asini*; brand characterization

1. Introduction

Ejiao (*Colla Corii Asini*, CCA) made from donkey hide has been widely used as a precious traditional Chinese medicine (TCM) in China for more than 2000 years [1,2]. As recorded in ancient and classic TCM monographs, CCA shows great efficacy in enriching blood and staunching bleeding [3]. Nowadays, multiple brands of CCA have been manufactured by different pharmaceutical factories in China. Quality difference of CCA exists among different brands, affecting nutritional value and market price. The nutritional properties of CCA are affected by such factors, as geographical origins, raw materials, water sources as well as pharmaceutical processes. Among these factors, geographical origin



Citation: Xia, Z.; Che, X.; Ye, L.; Zhao, N.; Guo, D.; Peng, Y.; Lin, Y.; Liu, X. A Synergetic Strategy for Brand Characterization of Colla Corii Asini (Ejiao) by LIBS and NIR Combined with Partial Least Squares Discriminant Analysis. *Molecules* **2023**, *28*, 1778. https://doi.org/ 10.3390/molecules28041778

Academic Editor: Már Másson

Received: 24 November 2022 Revised: 2 February 2023 Accepted: 7 February 2023 Published: 13 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). mainly confers nutritional qualities to CCA. Traditionally, CCA is evaluated according to its producing areas [3]. For example, Dong'e Ejiao (DEEJ) produced in Dong'e county of Shandong province has been acknowledged as authentic and top-quality CCA products. DEEJ is well known as one of the intangible cultural heritage resources. However, a similar manufacturing process of CCA caused similar appearance characteristics and physical features. It is difficult to distinguish DEEJ from CCA of other brands. A critical issue in the quality control of Ejiao is brand or manufacturer identification. Therefore, a way of discriminating against CCA is needed for the purpose of brand protection and competitive advantage.

To date, multiple analytical techniques have been successfully prioritized for quality assessment and differentiation of CCA, such as vibrational spectroscopic techniques (Fourier transform infrared (FTIR) spectroscopy and near-infrared (NIR) spectroscopy [4,5], liquid chromatography mass-spectrometry [6,7], gas chromatography-mass spectrometry (GC-MS) [4], and Polymerase Chain Reaction (PCR) [8], etc. The truth is that data obtained from different analytical techniques can provide complementary chemical information thus improving the qualitative or quantitative analysis performance [9–12]. However, these approaches applied only one analytical tool.

Spectroscopic techniques are appealing tools for fast quality control of traditional Chinese medicine (TCM) with no laborious sample preparation. For example, laser-induced breakdown spectroscopy (LIBS) and near-infrared spectroscopy (NIR) have many advantages such as minimal sample pretreatment, fast detection, less destruction, environmentally friendly and cost-effective properties of the analytical procedure, have been uesd in pharmaceuticals, food, and agriculture filed [13–15]. LIBS is a spectrochemical sensor technology for the acquisition of simultaneous multi-elemental analysis based on atomic emission [16,17]. In LIBS, a pulsed laser beam is focused on the sample to vaporize the sample and induce the formation of a plasma. The emission of the plasma containing excited atoms and ions present in the sample is collected and analyzed through an optical system and a spectrometer in order to extract the spectroscopic information [18]. Qualitative and quantitative information can be obtained by identification of the spectral lines and analysis of their peak intensities, respectively. LIBS has the unique advantages of fast analysis, in-situ or remote measurements, and ability to analyze various samples such as solids, liquids, and gases with little or no sample preparation [17]. It is well accepted that some inorganic elements of TCM play significant roles in biological activity, which are responsible for the clinical effect [19]. Hence, elemental fingerprints are useful to identify the geographical origin and species of TCM.

Another promising technique is NIR, which has been successful in the quality control of food and drug. As a non-destructive technique, NIR employs measurements of photon energy to obtain qualitative and quantitative information based on the interaction of the sample with NIR radiation [20,21]. It covers the wavelength range of 800–2500 nm and mainly reflects vibration information of hydrogen bonds in a molecular structure, such as carbon–hydrogen (C-H), oxygen–hydrogen (O-H), nitrogen–hydrogen (N-H) [22,23]. Particularly, hydrogen bonds own high spectral stability in the NIR region, which aid in qualitative and quantitative analyses.

The combination of LIBS and NIR could provide complementary information on unknown species, involving not only the elemental compositions but also the molecular components. Nowadays, the two techniques have been prevailing analytical tools for quality control of TCM with the rapid development of powerful computers and data analysis tools [24–26]. In addition, the potential applications of LIBS coupled with NIR on analyses of food and vegetable achieve considerable success [18,27]. However, brand characterization of CCA by combining LIBS with NIR has not been reported yet.

Usually, the application of LIBS and NIR spectroscopy in classification analysis needs the assistance of chemometrics approaches to construct the discriminant model. Among these approaches, partial least squares discriminant analysis (PLS-DA) is an effective algorithm used to optimize separation between different groups of samples [28]. PLS-DA is based on the classical PLS regression algorithm, which models the response matrix Y (class membership) through the predictive matrix X (original data) [28,29]. The Y matrix in PLS-DA is a set of binary variables (0 and 1), representing the membership of each sample. The main advantage of PLS-DA is that it can offer a graphical visualization and understanding of the different data patterns and relations by latent variables (LVs) and loadings [30].

The main objective of the present study was to classify CCA samples according to their species. CCA samples mainly contain organic and inorganic components including collagen, proteins, amino acids, polysaccharides, trace elements, water, etc. [3]. We focused on the application of a synergetic strategy to build robust classification models for brand characterization of CCA collected from different pharmaceutical factories in China. LIBS and NIR were fully investigated to offer atomic and molecular information on CCA samples. In this work, PLS-DA was chosen for model development based on LIBS and NIR spectroscopy processed separately. Considering the potential comprehensive information of CCA species provided by LIBS and NIR, a low-level data fusion strategy was also employed. After spectral pretreatment, two types of PLS-DA models were established, namely, single models and data fusion models. Then, the classification performance of PLS-DA models was evaluated by three statistical indicators such as total accuracy, sensitivity (Se), and specificity (Sp) [31,32].

The methodology is found to be helpful for the quality control of CCA, and exhibits the promise of addressing similar needs in other precious TCM products in the future.

2. Results and Discussion

2.1. Spectral Features of LIBS and NIR

The normalized LIBS spectra of the CCA sample is shown in Figure 1. The elemental compositions of CCA were identified and marked according to the NIST (US, National Institute of Standards and Technology) database [33]. LIBS spectra are quite complex due to the emission of multi-elements from samples. The characteristic spectral lines of Ca, Na, Mg, K, and Fe were observed in the CCA LIBS spectra. In addition, spectral lines of organic compositions, such as C, N and O can also be observed, which are actually both from CCA samples and the air. Additionally, some molecular spectra, such as C-N emissions appear in the LIBS spectrum. While the C-N bands arise from the interaction of carbon with atmospheric nitrogen. Selecting the appropriate characteristic spectra is significant to the classification results. The 47 spectral lines including 13 elements and four C-N molecular bands are summarized in Table 1 and chosen for further analysis.



Figure 1. The representative normalized LIBS spectrum of the CCA sample.

Elements	Wavelength (nm)	Elements	Wavelength (nm)	
С	247.722	Ва	455.360, 493.385	
	279.422, 280.124,			
Ma	285.085, 383.825	Fo	526 987	
1415	516.730, 517.270,	i c	020.007	
	518.359			
Si	288.031	Na	588.958, 589.551	
Ca	315.866, 317.921,			
	370.621, 393.378			
	396.814, 422.639,	Ν	742.400, 744.294,	
Cu	428.287, 430.226		746.927	
	442.643, 443.500,			
	445.444, 558.850			
	612.923, 616.233,			
	643.966, 646.212	Li	670.746	
	649.394, 714.856,			
	720.259			
C-N	385.461, 386.105,	Н	656.309	
	387.087, 388.302	T/		
Al	394.422, 396.096	K	766.515, 769.947	
Sr	407.786, 421.500	0	777.216, 777.505	

Table 1. Selected spectral lines of LIBS Spectra.

Figure 2 shows plots of the original NIR spectra of CCA samples. The major features of the raw NIR spectra exhibit no distinct differences among different samples except the transmittance (indirect reflecting the absorption intensity). The NIR spectra reflects the organic molecules' information containing molecular bonds, such as C-H, O-H, N-H, etc [34]. Three characteristic absorption bands can be viewed at 5784 cm⁻¹, 6650 cm⁻¹, 8417 cm⁻¹, which are abundant in the molecular structures of amino acids, peptides, proteins, etc. In addition, the relatively slow absorption peak appeared at approximately 5151 cm⁻¹, which may correspond to the frequency of the O–H bond in water molecules. The NIR spectroscopy-based fingerprint contains complex attributes of the CCA samples sufficiently, which are related to brand characterization.



Figure 2. NIR spectra of Ejiao samples.

2.2. PLS-DA Model from LIBS Data

The PLS-DA model was built on a single LIBS data block using the selected variables (listed in Table 2). When the number of LVs set as 11, the classification accuracy of the PLS-DA model reached the max value. Hence, the optimal number of LVs was selected as 11. Subsequently, the PLS-DA model was constructed to validate the training set and predict the testing set. Figure 3 displays the classification result of the PLS-DA model using the LIBS spectra. As can be seen from Figure 3, all samples are properly classified in the training set while the DEEJ samples are misclassified as non-DEEJ samples and one non-DEEJ sample is misclassified as a DEEJ sample in the testing set. The total accuracy values of the PLS-DA model are 100% and 85.18% for different class samples in the training set and the testing set, respectively. The sensitivity values of the PLS-DA model are 100% and 91.67% for different class samples in the training and testing sets, while the specificity values are 100% and 92.30% for different class samples, respectively. The misclassified sample in the top right corner of Figure 3 was sample No. 74 (belonging to Shandong Fupai Ejiao Co., Jinan, China, Ltd., listed in Table 2). We compared the LIBS spectra. It was found that the LIBS spectra of non-DEEJ sample No. 74 was very similar to DEEJ sample No. 38 (listed in Table 2), which indicated that they have the same elemental compositions and similar peaks intensity.

Table 2. Detailed information of all the CCA samples.

Sample No.	Manufacturer
1–40	Shandong Dong'e Ejiao Co., Ltd. Liaocheng city, Shandong, China
41–48	Shandong Huaxin Pharmaceutical Group Co., Ltd. Heze city, Shandong, China
49–61	Shandong Yanggu Guajing Ejiao factory Liaocheng city, Shandong, China
62	Shandong Dong'a Xiuyuan Ejiao biological group Neihuang Ejiao Pharmaceutical Co., Ltd. Anyang city, Henan, China
63	Shandong Jishui Ejiao Co., Ltd. Heze city, Shandong, China
64–65	Shandong Hongjitang Pharmaceutical Group Co., Ltd. Jinan city, Shandong, China
66–72	Shandong Dong'e Guojiaotang Ejiao Pharmaceutical Co., Ltd. Liaocheng city, Shandong, China
73–75	Shandong Fupai Ejiao Co., Ltd. Jinan city, Shandong, China
76–78	Shandong Yixiaotang Ejiao group Bainian Pharmaceutical Co., Ltd. Zaozhuang city, Shandong, China



Figure 3. The PLS-DA model using the LIBS spectra.

The performance of PLS-DA models for NIR spectra with different pretreatment methods is shown in Table 3. The total accuracy values of the PLS-DA model based on raw NIR spectroscopy are 100% and 88.89% for the training set and the testing set, respectively. The sensitivity values of the PLS-DA model are 100% and 82.35% for different class samples, while the specificity values are 100% both in the training set and testing set, respectively. As shown in Figure 4, all samples are properly assigned in the training set but four Non-DEEJ samples are incorrectly assigned as DEEJ samples in the testing set, which means that, except for these four incorrectly classified samples, the model correctly predicted the class of the other samples.

Table 3. The discrimination results of the PLS-DA model based on NIR spectra (in the region of $5500-9000 \text{ cm}^{-1}$) by different preprocessing methods.

Pretreatment	LVs	Training Set			Testing Set		
		Se (%)	Sp (%)	Ta (%)	Se (%)	Sp (%)	Ta (%)
Raw	7	100	100	100	93.33	100	96.3
MSC	10	100	100	100	100	100	100
SNV	7	100	100	100	100	92.86	96.3
SG9	9	100	100	100	82.35	100	88.89
1st d	6	100	100	100	92.86	92.31	96.3

Some preprocessing methods for the NIR data were compared, such as multiplicative scatter correction (MSC), standard normal variate transformation (SNV), Savitzky–Golay smoothing with 9 points (SG9), and first derivative (1st d).



Figure 4. The PLS-DA model using the raw NIR spectra.

Screening of specific NIR wavebands was employed to increase the classification performance between DEEJ and Non-DEEJ samples. The screened wavebands include 5500–9000 cm⁻¹, which might represent the special spectroscopic characteristic of the discriminatory markers. The discrimination ability of PLS-DA models building with original and different pretreatments of NIR spectra were computed and compared. As shown in Table 3, all PLS-DA models constructed with NIR spectra show good prediction performance. Meanwhile, the discrimination potential in the testing set is better than that in the testing set. The PLS-DA model established by the raw NIR spectra exhibits satisfactory performance. The total accuracy values are 100% and 96.3% for different class samples in

the training set and in the testing set, respectively. The sensitivity values are 100% and 93.33% for different class samples in the training set and testing set, respectively, while specificity values reach 100% for different class samples in the training set and in the testing set, respectively. Theoretically, as might be expected, PLS-DA models built on the NIR data by different data preprocessing methods can improve the discrimination. But the models obtained by Savitzky-Golay smoothing (SG) and first derivative (1st d) pretreatment do not show better performance than that of screened original NIR spectra.

Among all the discriminant models, the PLS-DA model with NIR spectra pretreated by the multiplicative scatter correction (MSC) method offers perfect performance than the others. When the optimal number of LVs set as 10, this PLS-DA model provides correct classification for all samples both in the training and testing sets. The PLS-DA discriminatory information of NIR spectra pretreated by the MSC method is shown in Figure 5. Therefore, the MSC method is identified as the optimal pretreatment method for NIR spectra. The capability of NIR spectroscopy for CCA authentication can attribute to the differences in the amount of hydrogen bonds (C-H, O-H, and N-H) absorption for the organic compounds presenting in CCA samples.



Figure 5. The PLS-DA model of NIR spectra pretreated by multiplicative scatter correction method.

2.4. PLS-DA Model from Data Fuion of LIBS and NIR

Data fusion is the integration of data blocks from different analytical technologies into a single model [35,36]. Though, single LIBS or single NIR can provide good discrimination ability. The low-level fusion strategy based on concatenating LIBS and NIR spectra was also investigated in this work. The discrimination ability of PLS-DA models based on the low-level fusion with different pretreatments is shown in Table 4. The total accuracy values of PLS-DA models based on LIBS combined with the screened original NIR spectra are 98.04% and 96.30% for different class samples in the training set and testing set, respectively. The sensitivity values are 96.15% and 93.33%, whereas specificity values are 96.15% and 100% for different class samples in the training sets and in the testing sets, respectively.

Pretreatment	LVs	Training Set			Testing Set		
		Se (%)	Sp (%)	Ta (%)	Se (%)	Sp (%)	Ta (%)
Raw	7	96.15	96.15	98.04	93.33	100	96.3
MSC	11	100	100	100	100	100	100
SNV	10	100	100	100	100	100	100
SG9	13	100	100	100	82.35	100	88.89
1st d	7	100	100	100	93.33	100	96.3

Table 4. The discrimination results of the partial least squares-discriminant analysis models based on low-level data fusion.

Some preprocessing methods for the NIR data were compared, such as multiplicative scatter correction (MSC), standard normal variate transformation (SNV), Savitzky-Golay smoothing with 9 points (SG9), and first derivatives (1st d).

Taking the classification efficiency into account, SNV and MSC were the best data preprocessing methods for discriminating DEEJ from Non-DEEJ, which can eliminate redundant information and manifest differences among samples. When the optimal number of LVs set as 10, the total accuracy values of the PLS-DA model based on low-level data fusion with SNV data pretreatment reach 100% for different class samples both in the training sets and testing sets. The sensitivity values and specificity values of the PLS-DA model are 100% for different class samples both in the training sets and testing sets. The data fusion model also exhibits excellent discrimination performance with MSC data pretreatment. When the optimal number of LVs set as 11, the total accuracy values, sensitivity values, and specificity values all reach 100%, indicating no sample is misclassified.

3. Materials and Methods

3.1. Sample Preparation

Seventy-eight CCA samples were collected from different manufacturers in China, including forty DEEJ samples and thirty-eight CCA samples from other factories called Non-DEEJ. Sample storage may influence the CCA quality. As a drug, CCA was sealed with a box. Thus, all samples were stored in a cool, dark area before analysis. The detailed information on all CCA samples is listed in Table 1. Due to the solid and glossy surface of CCA samples, no special sample preparation was performed in LIBS and NIR experiments.

3.2. Sample Measurement

The LIBS system used for CCA analysis comprises a Q-switched Nd: YAG laser operating at 1064 nm with a pulses duration of 3–5 ns at a frequency of 2 Hz. Echelle spectrometer (LTB, ARYELLE 150, Berlin, Germany) records the spectrum from 193 to 840 nm with an approximate resolution of 0.1 nm. The laser beam was focused on the sample surface by a microscope objective lens system with an 80 mm focal length (Avantes UV-74). The focused spot size on the target surface was about 100 μ m in diameter. An optical fiber was connected to the spectrometer with an intensified gated CCD camera (LTB ARYELLE 200 and iStar from Andor Technology) to disperse and record the spectra. A laser pulse energy of 40 mJ, a gate width of 100 μ s, and a delay time of 0.6 μ s resulted in the maximum signal-to-noise ratio, then being selected for LIBS experiments. All LIBS experiments were performed in ambient air. "LTB-Sophi" software was used to record the emission spectrum of the species and extract analytical information. Four laser shots per location of twelve sampling locations were collected and spectra averaged aiming to limit the relative standard deviation of the peak-intensity ratios.

NIR spectra were collected by Antaris Nicolet FT-NIR system (Thermo Fisher Scientific Inc., Waltham, MA, USA). All spectra were measured in integrating sphere diffuse mode. Each spectrum was an average of 64 scans across the wavenumbers from 10,000 to 4000 cm^{-1} with a resolution of 4 cm⁻¹ at ambient temperature. The spectra was recorded as Log(1/R) using air as a reference.

3.3. Spectral Pretreatment

The original LIBS spectra are often affected by various sources of interference from light scattering, instrument noise, and background [27]. Thus, conventional minimummaximum normalization was employed to transform the original LIBS data to the range of 0–1 [27]. The raw spectra acquired from the NIR spectrometer are often characterized by baseline shift and unwanted spectral variation produced from light scattering effects, variation in optical path length, random noise, sample matrix, and others. In order to build a robust and stable model, data pretreatment must be taken prior to data analysis to minimize interference of NIR spectra. Here, four basic NIR spectral pretreatment methods were applied to reduce interference and optimize model performance, including standard normal variate (SNV), multiplicative scatter correction (MSC), Savitzky-Golay smoothing (SG), first derivative (1st d). Moreover, the optimal pretreatment method is selected by the highest classification accuracy rule in both training and testing sets.

3.4. Multivariate Analysis and Latent Variable Analysis

The spectral preprocessing and model calculation steps were performed by the Unscrambler 9.7 software (CAMO software AS, Oslo, Norway) and homemade routines in Matlab version R2009a with Statistical Toolbox. Partial least squares discriminant analysis (PLS-DA) algorithm was investigated for a single data block or fusion block respectively. To obtain robust multivariate calibration models, 78 samples were split into two sets of training (calibration) and testing (validation) sets in a ratio of 2:1 by KennardStone (KS) algorithm. Training set is used for building the model while the testing set is used for assessing the robustness of the model [29]. Before PLS-DA, several pretreatments were performed on LIBS and NIR spectra. It is generally known that the number of latent variables (LVs) is a critical parameter that is obtained from the evaluation of the class border by Bayesian decision [30]. The optimal number of LVs for each model was determined by 10-fold cross-validation based on the highest classification accuracy. Then the optimized model was used for the testing set. Besides total accuracy (Tc), sensitivity (Se) and specificity (Sp) were mainly computed to evaluate the performance of the established PLS-DA model based on preprocessed data. Sensitivity, also called the true positive rate (TPR), represents the positive correctly classified samples to the total number of positive samples as in Equation (2). Whereas specificity, also called the true negative rate(TNR), is expressed as the proportion of the correctly classified negative samples to the total number of negative samples as in Equation (3).

The three evaluation parameters were calculated as follows:

$$Tc = (TP + TN)/(TP + FP + FN + TN)$$
(1)

$$Se = TP/(TP + FN)$$
(2)

$$Sp = TN/(FP + TN)$$
(3)

where, *TP* and *TN* indicate the number of true positive samples and true negative samples, respectively. While *FP* and *FN* indicate the number of false positive samples and false negative samples. For each class, a positive sample means itself, and a negative sample is the other class. The sensitivity corresponds to the model's ability to correctly predict the class of the samples, while specificity reflects its ability to incorrectly prediction of samples from other classes [29]. Perfect discrimination of the PLS-DA model yields the highest values of total accuracy, sensitivity, and specificity, indicating all samples are successfully classified.

3.5. Data Fusion Strategy

Data fusion can produce more accurate information by integrating multiple complementary data sources. Basically, it can be categorized into three levels: low level, mid level, and high level [37–39]. Low-level data fusion strategy was adopted in this work. The scheme of data analysis in this experiment is shown in Figure 6. To prepare for low-level fusion, raw data from all CCA are straightforwardly concatenated into a new data matrix. Before data concatenation, all variables were normalized to the same scale aimed at rectifying the dimensional imbalance between raw data matrices.



Figure 6. The scheme of data analysis.

4. Conclusions

In this study, the feasibility of individual and combining information from LIBS and NIR techniques was researched aiming to address brand discriminant of Ejiao. LIBS and NIR spectroscopies bring atomic and molecular information of species. The discrimination results of PLS-DA models showed that individual techniques had good performance at authenticating brands of CCA. Especially, the PLS-DA model based on the single NIR data pretreated with the MSC method achieved a perfect classification effect, and it obtained 100% correct classification of each class for all samples. A data fusion strategy was also investigated for complementary information from two spectroscopic techniques. Data fusion models based on LIBS and NIR with SNV and MSC pretreatment also show good predictive ability. In general, the synergetic strategy of LIBS and NIR spectra demonstrated the possibility of brand discrimination of Ejiao, and it can be used to provide a reliable method for brand protection or geographical authenticity evaluation of TCM.

Author Contributions: Conceptualization, X.L. and Y.L.; methodology, X.L.; X.C. and Z.X.; software, X.L.; X.C.; N.Z. and Z.X.; formal analysis, X.C. and N.Z.; investigation, X.L.; L.Y. and Y.P.; resources, X.L.; D.G. and Y.L.; writing—original draft preparation, X.L. and Z.X.; writing—review and editing, X.L. and X.C.; project administration, X.L. and Z.X.; funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Natural Science Foundation of China (No. 81903796), and the National Key Research and Development Program of China (No. 2019YFC1711200).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Authors are thankful to the Shandong Institute of Food and Drug Inspection for supplying the samples, and to Tao Xu from Fujian Normal University for helping with the LIBS analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. National Pharmacopiea Committee. *Pharmacopoeia of the People'S Republic of China*, 1, 10th ed.; China Medical Science Press: Beijing, China, 2020; pp. 189–190.
- Li, X.; Shi, F.; Gong, L.; Hang, B.; Li, D.; Chi, L. Species-specific identification of collagen components in Colla corii asini using a nano-liquid chromatography tandem mass spectrometry proteomics approach. *Int. J. Nanomed.* 2017, 12, 4443–4454. [CrossRef]
- Zhang, S.; Xu, L.; Liu, Y.-X.; Fu, H.-Y.; Xiao, Z.-B.; She, Y.-B. Characterization of Aroma-Active Components and Antioxidant Activity Analysis of E-jiao (Colla Corii Asini) from Different Geographical Origins. *Nat. Prod. Bioprosp.* 2018, *8*, 71–82. [CrossRef] [PubMed]
- 4. Dong, W.-W.; Zeng, Z.-D.; Au, D.; Chan, C.-O.; Yang, D.-J.; Li, X.-B. Quality control of Colla corii asini using near-infrared spectroscopy and chemometrics clustering techniques. *J. Food Drug Anal.* **2012**, *20*, 152–158. [CrossRef]
- Li, W.-L.; Han, H.-F.; Zhang, L.; Zhang, Y.; Qu, H.-B. Manufacturer identification and storage time determination of "Dong'e Ejiao" using near infrared spectroscopy and chemometrics. J. Zhejiang Univ. Sci. B 2016, 17, 382–390. [CrossRef]
- Shen, L.; Chen, H.; Zhu, Q.; Wang, Y.; Wang, S.; Qian, J.; Wang, Y.; Qu, H. Identification of bioactive ingredients with immunoenhancementand anti-oxidative effects from Fufang-Ejiao-Syrup by LC–MSn combined with bioassays. *J. Pharm. Biomed. Anal.* 2016, 117, 363–371. [CrossRef]
- 7. Huang, H.; Liu, S.; Du, J.; Lin, J.; Liang, Q.; Liu, S.; Wei, Z. Structural analysis of glycosaminoglycans from Colla corii asini by liquid chromatography-electrospray ion trap mass spectrometry. *Glycoconj. J.* **2020**, *37*, 201–207. [CrossRef]
- Zhang, W.; Cui, S.; Cheng, X.-l.; Wei, F.; Ma, S. An optimized TaqMan real-time PCR method for authentication of ASINI CORII COLLA (donkey-hide gelatin). J. Pharm. Biomed. 2019, 170, 196–203. [CrossRef]
- 9. De Queiroz Baddini, A.L.; de Paula Santos, J.L.V.; Tavares, R.R.; de Paula, L.S.; da Costa Araújo Filho, H.; Freitas, R. PLS-DA and data fusion of visible Reflectance, XRF and FTIR spectroscopy in the classification of mixed historical pigments. Spectrochim. *Acta A Mol. Biomol.* **2022**, *265*, 120384. [CrossRef]
- Liang, J.; Li, M.; Du, Y.; Yan, C.; Zhang, Y.; Zhang, T.; Zheng, X.; Li, H. Data fusion of laser induced breakdown spectroscopy (LIBS) and infrared spectroscopy (IR) coupled with random forest (RF) for the classification and discrimination of compound salvia miltiorrhiza. *Chemometr. Intell. Lab. Syst.* 2020, 207, 104179. [CrossRef]
- 11. Song, X.-C.; Canellas, E.; Asensio, E.; Nerín, C. Predicting the antioxidant capacity and total phenolic content of bearberry leaves by data fusion of UV–Vis spectroscopy and UHPLC/Q-TOF-MS. *Talanta* **2020**, *213*, 120831. [CrossRef]
- Dai, S.; Lin, Z.; Xu, B.; Wang, Y.; Shi, X.; Qiao, Y.; Zhang, J. Metabolomics data fusion between near infrared spectroscopy and high resolution mass spectrometry: A synergetic approach to boost performance or induce confusion. *Talanta* 2018, 189, 641–648. [CrossRef]
- 13. Hahn, D.W.; Omenetto, N. Laser-Induced Breakdown Spectroscopy (LIBS), Part II: Review of instrumental and methodogical approaches to material analysis and applications to different fields. *Appl. Spectrosc.* **2012**, *66*, 347–419. [CrossRef] [PubMed]
- 14. Nagy, M.; Wang, S.P.; Farag, M. Quality analysis and authentication of nutraceuticals using near IR (NIR) spectroscopy: A comprehensive review of novel trends and applications. *Trends Food Sci. Technol.* **2022**, *123*, 290–309. [CrossRef]
- 15. Guo, L.-B.; Zhang, D.; Sun, L.-X.; Yao, S.-C.; Zhang, L.; Wang, Z.-Z.; Wang, Q.-Q.; Ding, H.-B.; Lu, Y.; Hou, Z.-Y.; et al. Development in the application of laser-induced breakdown spectroscopy in recent years: A review. *Front. Phys.* **2021**, *16*, 22500. [CrossRef]
- Wang, Z.; Afgan, M.S.; Gu, W.; Song, Y.; Wang, Y.; Hou, Z.; Song, W.; Li, Z. Recent advances in laser-induced breakdown spectroscopy quantification: From fundamental understanding to data processing. *Trends Analyt. Chem.* 2021, 143, 116385. [CrossRef]
- 17. Zhao, S.Y.; Song, W.R.; Hou, Z.Y.; Wang, Z. Classification of ginseng according to plant species, geographical origin, and age using laser-induced breakdown spectroscopy and hyperspectral imaging. J. Anal. At. Spectrom. 2021, 36, 1704–1711. [CrossRef]
- De Oliveira, D.M.; Fontes, L.M.; Pasquini, C. Comparing laser induced breakdown spectroscopy, near infrared spectroscopy, and their integration for simultaneous multi-elemental determination of micro- and macronutrients in vegetable samples. *Anal. Chim. Acta* 2019, 1062, 28–36. [CrossRef]
- 19. Liu, X.; Che, X.; Li, K.; Wang, X.; Lin, Z.; Wu, Z.; Zheng, Q. Geographical authenticity evaluation of Mentha haplocalyx by LIBS coupled with multivariate analyzes. *Plasma Sci. Technol.* **2020**, *22*, 074006. [CrossRef]
- 20. Chen, Q.; Lin, H.; Zhao, J. Near-infrared spectroscopy technology in food. In *Advanced Nondestructive Detection Technologies in Food*; Springer: Singapore, 2021.
- 21. Tsuchikawa, S.; Ma, T.; Inagaki, T. Application of near-infrared spectroscopy to agriculture and forestry. *Anal. Sci.* 2022, *38*, 635–642. [CrossRef]
- 22. Wang, Y.; Yang, Y.; Sun, H.; Dai, J.; Zhao, M.; Teng, C.; Ke, Z.; Yang, M.; Zhong, L.; Zhu, W. Application of a data fusion strategy combined with multivariate statistical analysis for quantification of puerarin in Radix puerariae. *Vib. Spectrosc.* **2020**, *108*, 103057. [CrossRef]
- 23. Tao, L.; Via, B.; Wu, Y.; Xiao, W.; Liu, X. NIR and MIR spectral data fusion for rapid detection of Lonicera japonica and Artemisia annua by liquid extraction process. *Vib. Spectrosc.* **2019**, *102*, 31–38. [CrossRef]
- Shen, T.; Li, W.; Zhang, X.; Kong, W.; Liu, F.; Wang, W.; Peng, J. High-Sensitivity Determination of Nutrient Elements in Panax notoginseng by Laser-induced Breakdown Spectroscopy and Chemometric Methods. *Molecules* 2019, 24, 1525. [CrossRef] [PubMed]

- Ma, L.; Liu, D.; Du, C.; Lin, L.; Zhu, J.; Huang, X.; Liao, Y.; Wu, Z. Novel NIR modeling design and assignment in process quality control of Honeysuckle flower by QbD. Spectrochim. *Acta A Mol. Biomol.* 2020, 242, 118740. [CrossRef] [PubMed]
- Fu, H.; Shi, Q.; Wei, L.; Xu, L.; Guo, X.; Hu, O.; Lan, W.; Xie, S.; Yang, T. 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. J. Raman Spectrosc. 2019, 2019, 2467185. [CrossRef]
- 27. Aints, M.; Paris, P.; Laan, M.; Piip, K.; Riisalu, H.; Tufail, I. Determination of Heating Value of Estonian Oil Shale by Laser-Induced Breakdown Spectroscopy. J. Spectrosc. 2018, 2018, 4605925. [CrossRef]
- Li, Z.; Song, J.; Ma, Y.; Yu, Y.; He, X.; Guo, Y.; Dou, J.; Dong, H. Identification of aged-rice adulteration based on near-infrared spectroscopy combined with partial least squares regression and characteristic wavelength variables. *Food Chem. X* 2023, 17, 100539. [CrossRef]
- 29. Barker, M.; Rayens, W. Partial least squares for discrimination. J. Chemom. 2003, 17, 166–173. [CrossRef]
- De Oliveira, V.M.A.T.; Baqueta, M.; Março, P.; Valderrama, P. Authentication of organic sugars by NIR spectroscopy and partial least squares with discriminant analysis. *Anal. Methods* 2020, *12*, 701–705. [CrossRef]
- Liu, X.; Zhang, Q.; Wu, Z.; Shi, X.; Zhao, N.; Qiao, Y. Rapid elemental analysis and provenance study of Blumea balsamifera DC using laser-induced breakdown spectroscopy. *Sensors* 2015, 15, 642–655. [CrossRef]
- Sun, F.; Chen, Y.; Wang, K.-Y.; Wang, S.-M.; Liang, S.-W. Identification of Genuine and Adulterated Pinellia ternata by Mid-Infrared (MIR) and Near-Infrared (NIR) Spectroscopy with Partial Least Squares—Discriminant Analysis (PLS-DA). *Anal. Lett.* 2020, 53, 937–959. [CrossRef]
- Kramida, A.; Ralchenko, Y.; Reader, J.; NIST ASD Team. NIST Atomic Spectra Database (Version 5.10); National Institute of Standards and Technology: Gaithersburg, MD, USA, 2022. Available online: https://physics.nist.gov/asd (accessed on 10 February 2023). [CrossRef]
- 34. Ozaki, Y. Near-Infrared Spectroscopy-Its Versatility in Analytical Chemistry. Anal. Sci. 2012, 28, 545–563. [CrossRef]
- 35. Zhang, H.; Liu, Z.; Zhang, J.; Zhang, L.; Wang, S.; Wang, L.; Chen, J.; Zou, C.; Hu, J. Identification of Edible Gelatin Origins by Data Fusion of NIRS, Fluorescence Spectroscopy, and LIBS. *Food Anal. Methods* **2021**, *14*, 525–536. [CrossRef]
- Biancolillo, A.; Bucci, R.; Magrì, A.L.; Magrì, A.D.; Marini, F. Data fusion for multiplatform characterization of an Italian craft beer aimed at its authentication. *Anal. Chim. Acta* 2014, 820, 23–31. [CrossRef] [PubMed]
- Assis, C.; Gama, E.M.; Nascentes, C.C.; de Oliveira, L.S.; Anzanello, M.J.; Sena, M.M. A data fusion model merging information from near infrared spectroscopy and X-ray fluorescence. Searching for atomic-molecular correlations to predict and characterize the composition of coffee blends. *Food Chem.* 2020, 325, 126953. [CrossRef] [PubMed]
- Spiteri, M.; Dubin, E.; Cotton, J.; Poirel, M.; Corman, B.; Jamin, E.; Lees, M.; Rutledge, D. Data fusion between high resolution 1H-NMR and mass spectrometry: A synergetic approach to honey botanical origin characterization. *Anal. Bioanal. Chem.* 2016, 408, 4389–4401. [CrossRef]
- Sampaio, P.S.; Calado, C.R.C. Comparison of partial least squares-discriminant analysis and soft independent modeling of class analogy methods for classification of Saccharomyces cerevisiae cells based on mid-infrared spectroscopy. J. Chemom. 2021, 35, e3340. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.