Research Article

The Feasibility of Using Near Infrared Spectroscopy for Rapid Discrimination of Aged Shiitake Mushroom (*Lentinula edodes*) after Long-Term Storage

Lu Xu,[1,2] Xian-Shu Fu,[3] Chen-Bo Cai,[4] and Yuan-Bin She[2]

[1]*College of Material and Chemical Engineering, Tongren University, Tongren, Guizhou 554300, China*
[2]*College of Chemical Engineering, Zhejiang University of Technology, Hangzhou 310014, China*
[3]*College of Life Sciences, China Jiliang University, Hangzhou 310018, China*
[4]*College of Chemistry and Life Science, Chuxiong Normal University, Chuxiong 675000, China*

Correspondence should be addressed to Lu Xu; lxchemo@163.com and Yuan-Bin She; sheyb@zjut.edu.cn

Long-term storage can largely degrade the taste and quality of dried shiitake mushroom (*Lentinula edodes*). This paper aimed at developing a rapid method for discrimination of the regular and aged shiitake by near infrared (NIR) spectroscopic analysis and chemometrics. Regular (n = 197) and aged (n = 133) samples of shiitake were collected from six main producing areas in two successive years (2013 and 2014). NIR reflectance spectra (4000–12000 cm^{-1}) were measured with finely ground powders. Different data preprocessing method including smoothing, taking second-order derivatives (D2), and standard normal variate (SNV) were investigated to reduce the unwanted spectral variations. Partial least squares discriminant analysis (PLSDA) and least squares support vector machine (LS-SVM) were used to develop classification models. The results indicate that SNV and D2 can largely enhance the classification accuracy. The best sensitivity, specificity, and accuracy of classification were 0.967, 0.953, and 0.961 obtained by SNV-LS-SVM and 0.933, 0.930, and 0.932 obtained by SNV-PLSDA, respectively. Moreover, the low model complexity and the high accuracy in predicting objects produced in different years demonstrate that the classification models had a good generalization performance.

1. Introduction

Shiitake (*Lentinula edodes*) is an edible medicinal mushroom native to China and originally cultivated in East Asia [1, 2]. Presently, shiitake mushroom is the second most cultivated edible mushroom around the world and accounts for about a quarter of the worldwide production [3, 4]. Shiitake mushroom is also the fastest-growing species among all cultivated mushrooms [1]. In the United States, the shiitake mushroom production has exceeded 9 million pounds per year during 2009 [5]. It has recently attracted much attention as physiologically functional food as well as a source for the development of novel drugs [6]. Its importance can be attributed to its nutritional value and healthy effects, as well as its special taste [7]. As a popular nonstaple food, shiitake is rich in dietary fiber, minerals, and vitamins and low in fat [8]. Several important components have been separated from its basidiocarp, mycelium, and culture medium [9, 10], including biologically active polysaccharides, ergothioneine, phenolics, amino acids, dietary fiber, ergosterol, vitamins B1, B2, and C, and minerals. By experimental researches, shiitake mushroom has demonstrated many important therapeutic properties, including antitumoral, hipocolesterolemic, antifungic, and antimicrobial activities, along with a high antioxidant potential [11–13].

Fresh shiitake mushroom is highly perishable and tends to lose quality quickly after harvest. It has a short shelf life because of its high respiration rate, tendency to turn brown and lose water, and lack of physical protection against microbial attack [14]. Therefore, in China, only a small fraction of the total shiitake yield is freshly consumed and the bulk is made into dried mushroom to obtain a much longer

shelf life. The quality of dried shiitake can be influenced by many factors, including species, origins, cultivations, and processing. Moreover, long-term storage can cause aging and the quality degradation of shiitake [15]; for example, during storage, dried shiitake tends to absorb moisture in the air, which may cause the loss and reactions of soluble components and even mustiness. Although producers and sellers usually will recover the appearance by some physical or/and chemical processing, the taste and quality of aged shiitake can no longer be recovered. The aged shiitake should be sorted out and sold as lower-grade products. However, for economic reasons, most producers will just get rid of the seriously aged and mildewy shiitake and prefer to mix and sell the regular and aged shiitake together to get a good price. Therefore, it is necessary to develop a rapid and effective method to distinguish the aged and regular shiitake.

As a promising alternative approach to the traditional analytical methods, near infrared (NIR) spectrometry, when combined with chemometrics, has demonstrated great potential for rapid analysis of food products. NIR has some advantages over traditional chemical analysis and sensory methods, including less sample preparation, reduced analysis time and cost, simultaneous multicomponent analysis, and the potential use for online analysis. This had made NIR widely used in quality control of various foods and agricultural products [16, 17].

The objective of this work was to develop a rapid and accurate method to discriminate regular and aged shiitake by NIR spectrometry combined with chemometrics. Considering the diversity of practical objects, shiitake objects were collected from six main producing areas in successive two years. To ensure the generalization performance of classification models, different data preprocessing methods and classification models were studied and compared to obtain the models with least complexity and preprocessing.

2. Materials and Methods

2.1. Collection of Shiitake Samples. The shiitake objects were collected from six main producing areas in China in successive two years (2013 and 2014). The regular objects were made from fresh shiitake mushrooms harvested in the current year. All the aged objects had been stored for over one year by the time of NIR analysis. All the regular and aged shiitake objects were provided and labeled by the producers. The seriously dampened and mildewy shiitake objects were manually sorted out, because such seriously dampened shiitake can be readily identified and it is unnecessary to include them in the classification models. As a result, regular (197 objects) and dampened (133 objects) shiitake objects were obtained and the detailed information is listed in Table 1.

2.2. Sample Preparation and NIR Spectroscopy Analysis. Before grinding, all the shiitake samples were fully dried in the sun. Each intact object (full basidiocarp with pileus and stipe) was ground by a disintegrator and then the particles were filtered through a 200-mesh screen. Impacted sample powders were analyzed in a quartz sample cup using

TABLE 1: Origins and types of shiitake mushroom.

Types[a]	Producing areas	Years[b]	Objects
R	Nanyang	2013	12
R	Suizhou	2013	15
R	Qingyuan	2013	13
R	Hanzhong	2013	17
R	Ningde	2013	11
R	Sanmenxia	2013	13
R	Nanyang	2014	22
R	Suizhou	2014	13
R	Qingyuan	2014	21
R	Hanzhong	2014	18
R	Ningde	2014	22
R	Sanmenxia	2014	20
A	Nanyang	2013	12
A	Suizhou	2013	11
A	Hanzhong	2013	12
A	Nanyang	2014	22
A	Suizhou	2014	20
A	Qingyuan	2014	15
A	Hanzhong	2014	10
A	Ningde	2014	17
A	Sanmenxia	2014	14

[a]R = regular shiitake; A = aged shiitake.
[b]The time of sample collection and NIR analysis.

a Bruker-TENSOR37 FTIR spectrometer (Bruker Optics, Ettlingen, Germany) in the reflectance mode. The spectra were measured using a PbS detector with an internal gold background as the reference. The working range of spectrometer was 4000–12000 cm^{-1}. Each sample was measured triply while being stirred and impacted before each scanning. The spectrum was taken as the average of the three scans. The instrumental resolution was 4 cm^{-1} with a scanning interval of 1.929 cm^{-1}, so each raw spectrum had 4148 wavelengths. The temperature was kept at around 25°C and the humidity was kept at a stable level during analysis.

2.3. Preprocessing and Data Splitting. All the data preprocessing and statistical analysis were performed on Matlab 7.0.1 (Mathworks, Sherborn, MA). Considering the uncertainty in the NIR analysis and the measured spectra, different preprocessing methods were studied to select the most proper method. Smoothing can reduce white noise in the spectra and improve the signal-to-noise ratio (SNR). The algorithm of polynomial fitting [18] was used for smoothing for its simplicity and effectiveness. Taking derivatives can improve spectral resolution and remove linear baseline shifts, so first-order (D1) and second-order (D2) derivatives were also applied. To avoid decreasing of the SNR by direct differencing, D1 and D2 spectra were also computed by polynomial fitting algorithms. Moreover, standard normal variate (SNV) [19] was used to reduce the effects of the difference in powder sizes. The Kennard and Stone (K&S) algorithm [20] was then
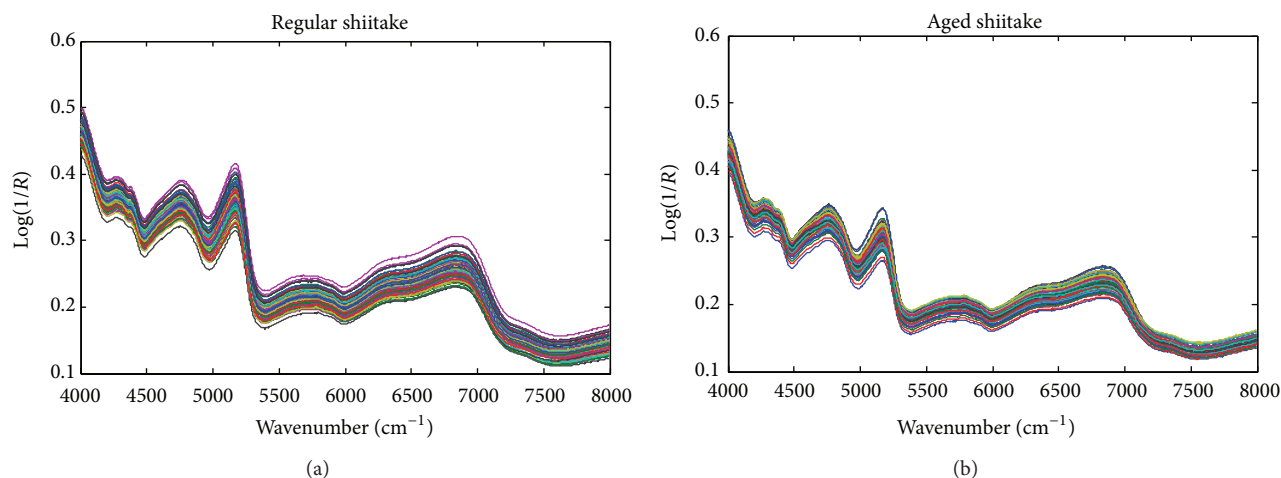
Figure 1: Raw NIR spectra of regular and aged shiitake objects.

applied to split the measured objects into a representative training set and test set. K&S algorithm selects the farthest points based on the Euclidean distance to be included in the training set to cover the widest range of experimental pace.

2.4. Classification Models. Partial least squares discriminant analysis (PLSDA) is a classification technique using partial least squares (PLS) regression. The usefulness and effectiveness of PLSDA as a classification technique have been demonstrated by the relationship between PLS, canonical correlation analysis (CCA), and linear discriminant analysis (LDA) [21]. For the two-class problem in this work, for the spectral matrix with objects arranged in rows, a dummy response vector was generated containing +1 and −1 to denote regular and aged samples, respectively. A PLS regression was built to relate the spectral matrix to the dummy response vector. For classification, an object with a predicted response value above/under 0 would be assigned to the regular/dampened group.

Support vector machine (SVM) [22] is a modeling technique based on the regularization of regression coefficients. By balancing the capacity and generalization performance of a learned model, SVM has been proved to be an effective tool to solve regression and classification problem. Least squares SVM (LS-SVM) [23] is a simplified version of SVM. Unlike the traditional SVM algorithms, which are based on quadratic programming, LS-SVM can obtain the solution by solving a set of linear equations. Moreover, nonlinear relationship can be learned using a nonlinear radial basis function (RBF).

In this work, linear PLSDA and nonlinear LS-SVM were used to classify the regular and dampened shiitake. For LS-SVM, the commonly used Gaussian RBF was adopted for nonlinear transformation. The number of PLSDA components and the parameters of LS-SVM were estimated using Monte Carlo cross validation (MCCV) [24].

2.5. Evaluation of Classification Performance. Sensitivity and specificity [25] were used to compare the classification models. In this work, the regular shiitake objects were denoted

as "positives" and the dampened ones as "negatives"; model sensitivity (Sen) and specificity (Spe) were defined as

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$
$$\text{Spe} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{1}$$

where true positives (TP), false negatives (FN), true negatives (TN), and false positives (FP) can be obtained from the numbers of misclassified objects. As an overall criterion to evaluate classification models, total accuracy (Acc) of classification was also used:

$$\text{Acc} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}}. \tag{2}$$

3. Results and Discussions

For the raw data, the spectral range of $8000–12000 \, \text{cm}^{-1}$ was significantly influenced by baseline shifts and scattering, so this spectral range was not used for further data analysis. Some of the raw NIR spectra ($4000–8000 \, \text{cm}^{-1}$) of regular and aged shiitake are shown in Figure 1. Seen from Figure 1, the regular and dampened samples have very similar peak patterns. Generally, due to the low resolution and serious peak overlapping, accurate assignments of most peaks were very difficult. Therefore, chemometrics methods were required to extract the useful information from spectral data for classification.

Figure 2 demonstrates the PCA plot of raw NIR spectra of regular and aged shiitake. The first two principal components (PCs) accounted for 92.7% of the total data variances. It can be seen that the projections of the objects in both classes onto the 2-PC subspace are very disperse, which can be attributed to the fact that both regular and aged shiitake included samples collected from different producing areas and years. The PCA plot also indicates that the difference in NIR data of the two classes is significant and can be used for classification. However, due to overlapping of classes,

TABLE 2: Classification of regular and aged shiitake by PLSDA.

Preprocessing	LVs[a]	MRMCCV	Sensitivity[b]	Specificity[c]	Accuracy
Raw data	4	0.230	0.800 (48/60)	0.837 (36/43)	0.816
Smoothing	4	0.228	0.817 (49/60)	0.837 (36/43)	0.825
D2	3	0.133	0.917 (55/60)	0.930 (40/43)	0.922
SNV	3	0.096	0.933 (56/60)	0.930 (40/43)	0.932

[a]The number of PLS latent variables.
[b]The numbers in the brackets indicate TP/(TP + FN).
[c]The numbers in the brackets indicate TN/(TN + FP).

TABLE 3: Classification of regular and aged shiitake by LS-SVM.

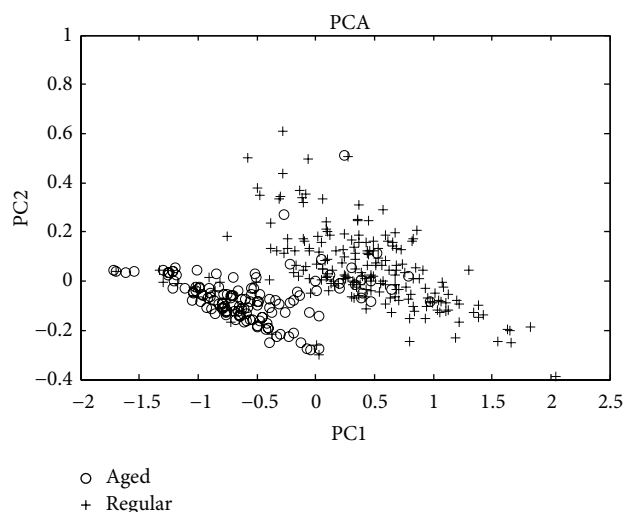| Preprocessing	σ^2, γ	MRMCCV	Sensitivity[a]	Specificity[b]	Accuracy
Raw data	984.4, 596.4	0.252	0.833 (50/60)	0.791 (34/43)	0.816
Smoothing	705.2, 529.7	0.259	0.833 (50/60)	0.791 (34/43)	0.816
D2	758.9, 140.3	0.184	0.883 (53/60)	0.930 (40/43)	0.903
SNV	215.3, 1.26	0.105	0.967 (58/60)	0.953 (41/43)	0.961

[a]The numbers in the brackets indicate TP/(TP + FN).
[b]The numbers in the brackets indicate TN/(TN + FP).

FIGURE 2: PCA plot of the 197 regular and 133 aged shiitake objects.

data preprocessing and classification models are required to accurately discriminate the regular and aged shiitake objects.

The preprocessed spectra of regular and aged shiitake samples are shown in Figure 3. By comparison of the raw spectra with smoothed spectra, although smoothed spectra can slightly improve the SNR, it might lose some useful high-frequency information in the raw data. Compared with first-order derivative spectra, the D2 spectra can remove most of the baselines and enhance some detailed information and peak resolution. SNV spectra can remove some spectral variations while enhancing others. The effects of data preprocessing should be evaluated by classification results.

The K&S algorithm was performed on the regular and aged samples separately. Each group was split into training and prediction objects, which were then combined to generate the final training and test sets. The final training set had 227 objects (137 regular and 90 aged objects) and 103 test samples (60 regular and 43 aged objects).

LS-SVM and PLSDA models were developed based on raw and preprocessed NIR data. For LS-SVM, the values of two parameters, γ and σ, need to be tuned. The magnitude of kernel width parameter, σ, is directly related to the nonlinear nature of the regression. The regularization parameter, γ, controls the tradeoff between the risk of model and the learning error. Therefore, γ and σ were optimized simultaneously by MCCV. Misclassification rate of MCCV (MRMCCV) was used to screen different combinations of γ and σ to obtain the lowest MRMCCV value. For MCCV, the number of random pieces of data splitting was 100 and each time 20% of the training objects were left out for prediction. The number of PLSDA latent variables was also estimated by MCCV.

The prediction results and optimized parameters of PLSDA and LS-SVM with different data preprocessing were presented in Tables 2 and 3. Obviously, data preprocessing by taking SNV and D2 spectra could significantly improve the classification accuracy. Data smoothing obtained almost the same classification results as using the raw data. Comparison of the results of different preprocessing methods indicates reducing baseline shifts and the influence of scattering effects was more important than improving the SNR of spectra. For the best classification accuracy, the sensitivity, specificity, and accuracy were 0.967, 0.953, and 0.961 and 0.933, 0.930, and 0.932 by SNV-LS-SVM and SNV-PLSDA, respectively. The SNV-LS-SVM model obtained slightly better results than the SNV-PLSDA model but the difference was insignificant. The prediction results by SNV-LS-SVM and SNV-PLSDA are shown in Figure 4. Compared with LS-SVM with RBF nonlinear transformation, PLSDA is a linear model and the risk of overfitting is lower. Moreover, the model complexity of PLSDA models was low (3 to 4 components), so PLSDA can be expected to have a better generalization performance when the model is applied to different types of shiitake objects and more objects should be included in the training set.
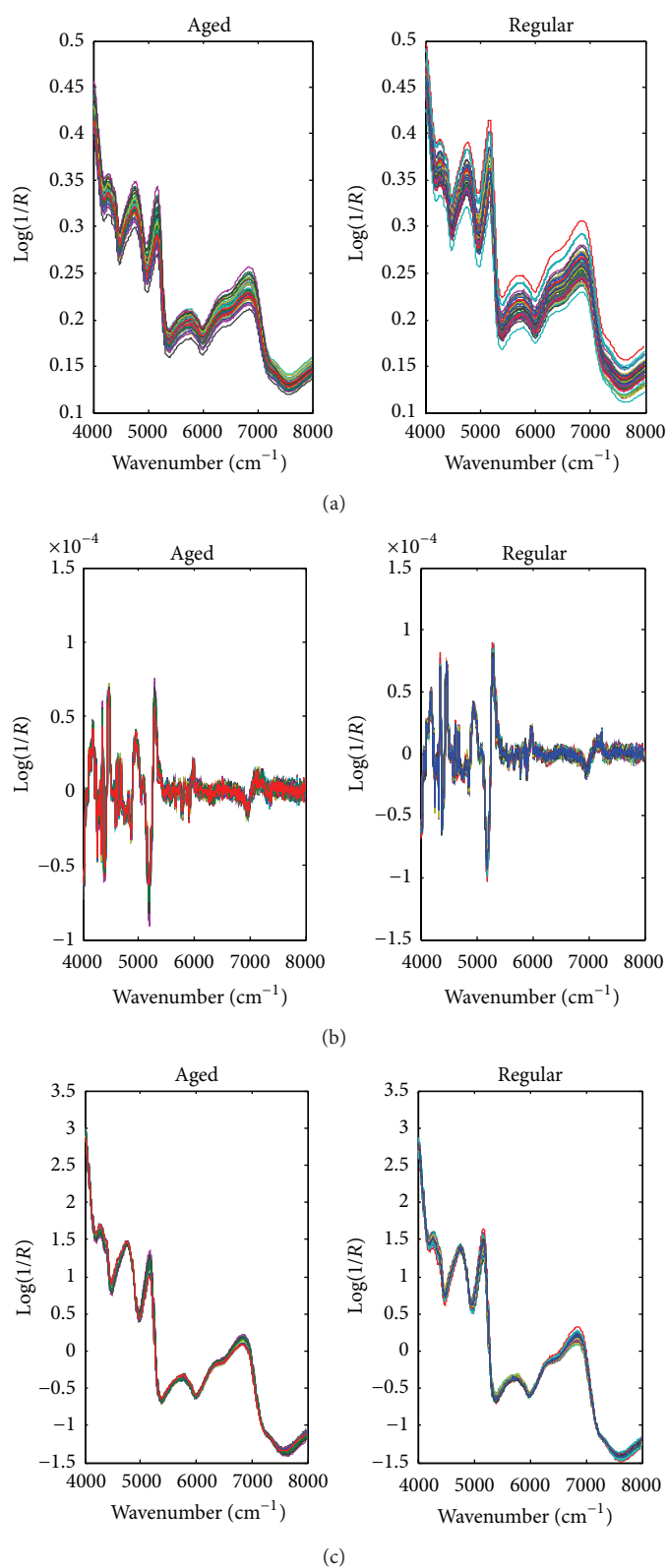
FIGURE 3: Preprocessed NIR spectra of regular and aged shiitake preprocessed by smoothing (a), SNV taking D2 (b), and SNV (c).
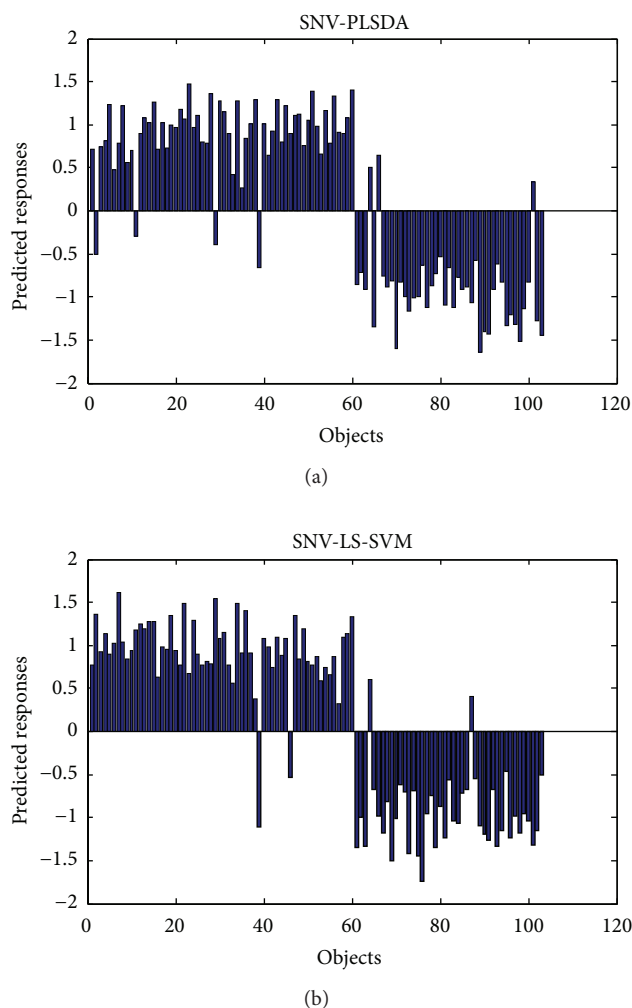
(a)

(b)

FIGURE 4: The predicted response values by SNV-PLSDA and SNV-LS-SVM models. Objects 1–60 are regular (positive) objects and 61–103 are aged (negative) objects.

4. Conclusion

The feasibility of using NIR spectroscopy for rapid discrimination of regular/aged shiitake mushroom was investigated. Both D2 and SNV could improve the classification performance significantly compared with the raw data, indicating that it is important to control the unwanted variations caused by baselines and scatter effects. With SNV transformation, both the linear PLSDA and nonlinear LS-SVM obtained accuracy discrimination of aged shiitake. The PLSDA had a low model complexity and the model can be generalized to practical applications by including different types of shiitake for training.

Conflict of Interests

Lu Xu declares no conflict of interests; Xian-Shu Fu declares no conflict of interests; Chen-Bo Cai declares no conflict of interests; Yuan-Bin She declares no conflict of interests.

Authors' Contribution

Lu Xu and Xian-Shu Fu equally contributed to this work.

References

[1] D. J. Royse and J. E. Sanchez, "Ground wheat straw as a substitute for portions of oak wood chips used in shiitake (*Lentinula edodes*) substrate formulae," *Bioresource Technology*, vol. 98, no. 11, pp. 2137–2141, 2007.

[2] Q. Shen, P. Liu, X. Wang, and D. J. Royse, "Effects of substrate moisture content, log weight and filter porosity on shiitake (*Lentinula edodes*) yield," *Bioresource Technology*, vol. 99, no. 17, pp. 8212–8216, 2008.

[3] M. Dermiki, N. Phanphensophon, D. S. Mottram, and L. Methven, "Contributions of non-volatile and volatile compounds to the umami taste and overall flavour of shiitake mushroom extracts and their application as flavour enhancers in cooked minced meat," *Food Chemistry*, vol. 141, no. 1, pp. 77–83, 2013.

[4] T. Llorente-Mirandes, M. Barbero, R. Rubio, and J. F. López-Sánchez, "Occurrence of inorganic arsenic in edible Shiitake (*Lentinula edodes*) products," *Food Chemistry*, vol. 158, pp. 207–215, 2014.

[5] L. C. Chandra, B. J. Smith, S. L. Clarke, D. Marlow, J. M. D'Offay, and S. R. Kuvibidila, "Differential effects of shiitake- and white button mushroom-supplemented diets on hepatic steatosis in C57BL/6 mice," *Food and Chemical Toxicology*, vol. 49, no. 12, pp. 3074–3080, 2011.

[6] N. F. M. Yahaya, M. A. Rahman, and N. Abdullah, "Therapeutic potential of mushrooms in preventing and ameliorating hypertension," *Trends in Food Science & Technology*, vol. 39, no. 2, pp. 104–115, 2014.

[7] J.-H. Yang, H.-C. Lin, and J.-L. Mau, "Non-volatile taste components of several commercial mushrooms," *Food Chemistry*, vol. 72, no. 4, pp. 465–471, 2001.

[8] N. J. Dubost, B. Ou, and R. B. Beelman, "Quantification of polyphenols and ergothioneine in cultivated mushrooms and correlation to total antioxidant capacity," *Food Chemistry*, vol. 105, no. 2, pp. 727–735, 2007.

[9] M. Wojtas, T. Bieńkowski, S. Tateyama et al., "Polyisoprenoid alcohols from the mushroom *Lentinus edodes*," *Chemistry and Physics of Lipids*, vol. 130, no. 2, pp. 109–115, 2004.

[10] T. Bito, F. Teng, N. Ohishi et al., "Characterization of vitamin B_{12} compounds in the fruiting bodies of shiitake mushroom (*Lentinula edodes*) and bed logs after fruiting of the mushroom," *Mycoscience*, vol. 55, no. 6, pp. 462–468, 2014.

[11] X. Li, H. Zhang, and H. Xu, "Analysis of chemical components of shiitake polysaccharides and its anti-fatigue effect under

vibration," *International Journal of Biological Macromolecules*, vol. 45, no. 4, pp. 377–380, 2009.

[12] T. Jiang, M. M. Jahangir, Z. Jiang, X. Lu, and T. Ying, "Influence of UV-C treatment on antioxidant capacity, antioxidant enzyme activity and texture of postharvest shiitake (*Lentinus edodes*) mushrooms during storage," *Postharvest Biology and Technology*, vol. 56, no. 3, pp. 209–215, 2010.

[13] J. R. Rao, B. C. Millar, and J. E. Moore, "Antimicrobial properties of shiitake mushrooms (*Lentinula edodes*)," *International Journal of Antimicrobial Agents*, vol. 33, no. 6, pp. 591–592, 2009.

[14] G. Antmann, G. Ares, P. Lema, and C. Lareo, "Influence of modified atmosphere packaging on sensory quality of shiitake mushrooms," *Postharvest Biology and Technology*, vol. 49, no. 1, pp. 164–170, 2008.

[15] T. Jiang, L. Feng, and J. Li, "Changes in microbial and postharvest quality of shiitake mushroom (*Lentinus edodes*) treated with chitosan-glucose complex coating under cold storage," *Food Chemistry*, vol. 131, no. 3, pp. 780–786, 2012.

[16] M. Bevilacqua, R. Bucci, S. Materazzi, and F. Marini, "Application of near infrared (NIR) spectroscopy coupled to chemometrics for dried egg-pasta characterization and egg content quantification," *Food Chemistry*, vol. 140, no. 4, pp. 726–734, 2013.

[17] S. Laroussi-Mezghani, P. Vanloot, J. Molinet et al., "Authentication of Tunisian virgin olive oils by chemometric analysis of fatty acid compositions and NIR spectra. Comparison with Maghrebian and French virgin olive oils," *Food Chemistry*, vol. 173, pp. 122–132, 2015.

[18] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.

[19] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Applied Spectroscopy*, vol. 43, no. 5, pp. 772–777, 1989.

[20] R. W. Kennard and L. A. Stone, "Computer aided design of experiments," *Technometrics*, vol. 11, no. 1, pp. 137–148, 1969.

[21] M. Barker and W. Rayens, "Partial least squares for discrimination," *Journal of Chemometrics*, vol. 17, no. 3, pp. 166–173, 2003.

[22] O. Devos, G. Downey, and L. Duponchel, "Simultaneous data pre-processing and SVM classification model selection based on a parallel genetic algorithm applied to spectroscopic data of olive oils," *Food Chemistry*, vol. 148, pp. 124–130, 2014.

[23] H. Zheng and H. Lu, "A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.)," *Computers and Electronics in Agriculture*, vol. 83, pp. 47–51, 2012.

[24] Q.-S. Xu and Y.-Z. Liang, "Monte Carlo cross validation," *Chemometrics and Intelligent Laboratory Systems*, vol. 56, no. 1, pp. 1–11, 2001.

[25] M. Forina, C. Armanino, R. Leardi, and G. Drava, "A class-modelling technique based on potential functions," *Journal of Chemometrics*, vol. 5, no. 5, pp. 435–453, 1991.