JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Feasibility of Infrared and Raman Spectroscopies for Identification of Juvenile Black Seabream (*Sparus macrocephalus*) Intoxicated by Heavy Metals

Xiaojing Chen,[†] Di Wu,[§] Xiaochun Guan,*[,†] Bo Liu,[#] Gui Liu,*[,†] Maocang Yan,[#] and Huiling Chen[†]

[†]College of Physics and Electronic Engineering Information, Wenzhou University Chashan University Town, Wenzhou, Zhejiang Province, People's Republic of China

[§]Food Refrigeration and Computerised Food Technology (FRCFT), School of Biosystems Engineering, University College Dublin, National University of Ireland, Agriculture and Food Science Centre, Belfield, Dublin 4, Ireland

[#]Zhejiang Mariculture Research Institute, Zhejiang Key Laboratory of Exploitation and Preservation of Coastal Bio-resource, Wenzhou, Zhejiang Province 325005, People's Republic of China

Supporting Information

ABSTRACT: The potential application of infrared and Raman spectroscopies was explored as rapid and nondestructive tools for the identification of juvenile black seabream samples intoxicated by heavy metals (Zn, Cu, and Cd). Discrimination models were established on the basis of the infrared and Raman spectral data using three calibration methods, namely, partial least-squares discriminant analysis, least-squares support vector machines, and random forest. The combination of two spectroscopies was studied, in which three combination strategies were proposed and compared. Discrimination models achieved overall correct discriminations of 100% for identifying the fish intoxicated by one heavy metal or the heavy metal mixture. When samples intoxicated by different heavy metals were analyzed together, the discrimination accuracy remained >90%. Results confirmed the possibility of developing fast and reliable systems for the identification of juvenile black seabream intoxicated by heavy metals based on infrared and Raman spectroscopies.

KEYWORDS: *heavy metal, infrared spectroscopy, Raman spectroscopy, fish, multivariate analysis*

INTRODUCTION

Heavy metal pollution is a serious environmental threat all around the world, raising particular concerns about the potential degradation of the ecosystems and their associated biota. Fish is widely consumed around the world as an excellent source of nutrients. However, fish is also a source of heavy metals. Similar to others, the aquatic environment is continuously being contaminated with discharges of heavy metals in recent years. Fish accumulates heavy metals directly through the water, indirectly through ingested food, and nondietary routes such as uptake through absorbing epithelia (i.e., the gill).[1] Some metals in fish tissues may be readily excreted, whereas other metals may alter the biochemical composition of tissues.[2] Being at the top of the aquatic food chain, fish accumulate much more heavy metals than other aquatic biota. When the accumulation of heavy metals in fish organs reaches a certain level, fish organs become highly toxic. Ingestion of fish with large amounts of heavy metals would cause serious health hazards to humans. For these reasons, the identification of fish intoxicated by heavy metals is extremely important to human health.

Conventional methods for determination of heavy metals normally include flame atomic absorption spectrometry (AAS), graphite furnace atomic absorption spectrometry (GFAAS), inductively coupled plasma atomic emission spectroscopy (ICP-AES), and inductively coupled plasma mass spectrometry (ICP-MS).[3] However, these methods are expensive, labor-intensive, tedious, and complex, always take a long time to analyze, and need a lot of sample preparations. To overcome these short-comings, there is a need for advanced methodologies that can provide rapid, simple, and reliable identifications of fish intoxicated by heavy metals. Moreover, fish have also been widely regarded as bioindicators for monitoring toxic chemicals in aquatic environments. The possible risk of heavy metals in aquatic environments could be evaluated on the basis of the identification of fish intoxicated by heavy metals in a rapid and simple way.

Recently, vibrational spectroscopy techniques have been investigated as a potential tool for the automatic quality and safety evaluation of fish, including infrared spectroscopy[4] and Raman spectroscopy.[5] They have the advantages of rapidity, simplicity, accuracy, low maintenance cost, and minimal sample preparation. IR represents a direct absorption of light in the mid-infrared region, which contains useful information of the fundamental absorptions of hydrogen-containing bonds (O−H, C−H, N−H, S−H, and P−H).[6] Raman spectroscopy is based on inelastic scattering of light that occurs when laser light interacts with molecules and condensed matter.[7] Both spectroscopies are nonperturbing rapid techniques that could detect the vibration information of different functional groups

Received: July 25, 2013
Revised: November 20, 2013
Accepted: November 21, 2013
Published: November 21, 2013

of biomolecules such as proteins, lipids, and carbohydrates in biological tissues and cells.

The major objective of this study was to investigate the feasibility of infrared and Raman spectroscopies to distinguish healthy juvenile black seabream (*Sparus macrocephalus*) and those intoxicated with heavy metals of zinc (Zn), copper (Cu), and cadmium (Cd). Cd is a nonessential metal as it is toxic, even in traces. Cu and Zn are essential metals because they play an important role in biological systems. However, essential metals can also produce toxic effects when the metal intake is excessively elevated.[3] To the best of our knowledge, this is the first study identifying fish intoxicated by heavy metals using infrared/Raman spectroscopy individually and their combination in tandem with multivariate analysis. The specific objectives of the current work were to (1) acquire spectral data of tested fish samples in infrared (4000−400 cm^{-1}) and Raman (4000−100 cm^{-1}) regions; (2) establish discrimination models of samples intoxicated with one heavy metal alone (design I); (3) establish discrimination models of samples intoxicated with the mixture of all three heavy metals (design II); (4) establish discrimination models for samples with all treatments (intoxicated with one heavy metal and all three heavy metals, design III); and (5) investigate the discrimination results of one spectroscopy alone and the spectroscopic combination.

MATERIALS AND METHODS

Test Samples and Chemicals. Juvenile black seabream of length 5 ± 1 cm and weight 2 ± 1 g were procured from Zhejiang Mariculture Research Institute at Wenzhou, China. The fish samples were acclimatized to the laboratory condition for about 2 days in plastic pools with the size of 60.0 cm × 40.0 cm × 30.0 cm. AnalaR grade $Zn(C_2O_3H_5)_2$, $CuSO_4{\cdot}5H_2O$, and $CdC1_2{\cdot}2.5H_2O$ were obtained from Chemical Reagent Co. Ltd., Shanghai, China.

Experimental Study. The acclimated test samples were divided into five groups, each containing 15 fish. The water used for maintaining the fish in the fish tanks was seawater, which was prepared after sedimentation over 24 h and sand filtration and had a pH of 8.05 ± 0.1, water temperature of 29.8 ± 0.6 °C, dissolved oxygen of >6 mg/L, and salinity of 21‰. The water was changed every 24 h throughout the experiment. The containers were refilled and redosed with the metal toxicant. There were five fish in each tank. Each group had three tanks. The test fish in groups I, II, and III were exposed to higher sublethal concentrations of $Zn(C_2O_3H_5)_2$ (3.045 mg/L), $CuSO_4{\cdot}5H_2O$ (0.53 mg/L), and $CdC1_2{\cdot}2.5H_2O$ (11.505 mg/L), respectively, in water (subacute exposure). The concentrations referred to metal. The test fish from group IV were exposed to the mixture of three heavy metals at one-third of their sublethal concentrations. Group V was used as control and reared in seawater without heavy metal addition. Fish from all groups were reared for 10 days, which was enough for the accumulation of heavy metals. After the rearing period, the fish were sacrificed in a refrigerator at the temperature of −4 °C for 15 min and then used for spectral measurement. The daily diet was at the level that was about 5−10% of the fish's body weight.

Spectral Measurements. IR (4000−400 cm^{-1}) spectra were measured using a Tensor 27 spectrophotometer (Bruker, Germany) equipped with a Golden Gate Diamond ATR sampling accessory. The collection of all samples was completed in an airtight collection box. All samples were scanned 15 times and averaged using OMINIC software (version 5.2, Inc., Bruker). Raman spectra for 532 nm excitation were measured using a confocal Raman microscope (XploRA, Horiba Jobin Yvon, Paris, France). The objective 50× Olympus BX41microscope (Olympus Optical Co., Ltd., Tokyo, Japan) was used to focus the laser beam on the samples and to pick up the backscattered Raman signal. The exposure time was 10 s. The spectra of the samples were taken at excitation wavelengths of 532 nm using a power of 0.2 mW (0.1%).

Spectral Preprocessing. Spectral preprocessing is an integral part of spectral analysis to reduce effects from random noise, length variation of light path, and light scattering.[8] In this work, Savitzky−Golay first derivative (SG-1st Der) was used to decrease baseline effects in spectra and resolve nearby peaks. Besides, there were three preprocessing strategies used for the combination of infrared and Raman spectra. The first one simply combined the spectral data of infrared spectroscopy and Raman spectroscopy into one matrix (strategy I). The second strategy first executed mean normalization on infrared spectra and Raman spectra, respectively, and then combined the normalized infrared and Raman spectra together (strategy II). In the process of mean normalization, the areas below the spectra are made equal. The third strategy further applied the SG-1st Der on the data processed by the strategy II (strategy III). The performances of these three strategies were compared to determine the best spectral combination.

Spectral Calibration. Discrimination models were then established on the basis of the preprocessed data or the original spectra by employing three calibration methods, namely, partial least-squares discriminant analysis (PLS-DA), least-squares support vector machines (LS-SVM), and random forest (RF). PLS-DA is the discrimination version of partial least-squares regression (PLSR). As a bilinear modeling technique, PLSR extracts a set of orthogonal factors called latent variables (LVs) and explores the optimal function by minimizing the error of sum squares, which is typically performed by cross-validation.[9−11] Differing from PLSR that uses the reference values of the target attribute as the dependent variable, PLS-DA encodes the dependent variable with dummy variables describing the classes for the optimum separation of classes.[12] After encoding, PLS-DA is implemented in the usual way of PLSR.

LS-SVM is an optimized version of support vector machines (SVM) proposed by Suykens and Vandewalle.[13] It applies least-squares error in the training error function.[14] LS-SVM employs nonlinear map function and maps the input features to a high dimensional space, thus changing the optimal problem into equality constraint condition.[15] The minimum output coding was used to encode the integral numbers into a set of L binary classifiers.[12]

RF is an ensemble of classification trees with binary divisions to solve classification problems.[16] During the RF calculation, RF uses bagging or bootstrap aggregating to construct trees from the sample set, resulting in increasing diversity of the trees. The final classification is given by aggregating (majority vote or averaging) the vote cast by each tree for the class of the object. As an ensemble learning algorithm, RF can be more accurate and robust to noise than single classifiers.[17]

Model Evaluation. In the process of establishing discrimination models for designs I and II, there were 15 samples from group V and 15 samples from groups I, II, III, or IV, respectively. Therefore, there were 30 samples in four sample sets, respectively. The four sample sets were used for the intoxication analysis of Zn, Cu, Cd, and the mixture. For each sample set, there were 20 samples (10 samples from each group) used for calibration and the remaining 10 samples (5 samples from each group) for prediction. For design III, there were 75 samples (15 samples × 5 groups) from groups I, II, III, IV, and V, where 50 samples (10 samples from each group) were selected into the calibration sample matrix (X_C) and 25 samples (5 samples from each group) into the prediction sample matrix (X_P). For each group, fish samples in two tanks were used for the calibration (10 samples), and the fish samples in the last tank were for prediction (5 samples). Therefore, the samples used for the prediction were independent from the samples used for calibration. On the other hand, there were 3525 wavenumber variables in the range of 4000−600 cm^{-1} for the IR analysis, 2592 variables in the range of 4000−100 cm^{-1} for the Raman analysis, and 6117 variables for the combination of two spectroscopies. In addition, one column vector (Y_C) containing the integer numbers of intoxication status of samples from the calibration set was concatenated to matrix X_C, and the model calibration was executed based on X_C and Y_C. For the two-class problem (designs I and II), the integer numbers were set to 0 and 1, representing intoxicated samples and healthy samples, respectively. For the five-class problem (design III), the integer numbers were set to 0 and 1−4 representing samples

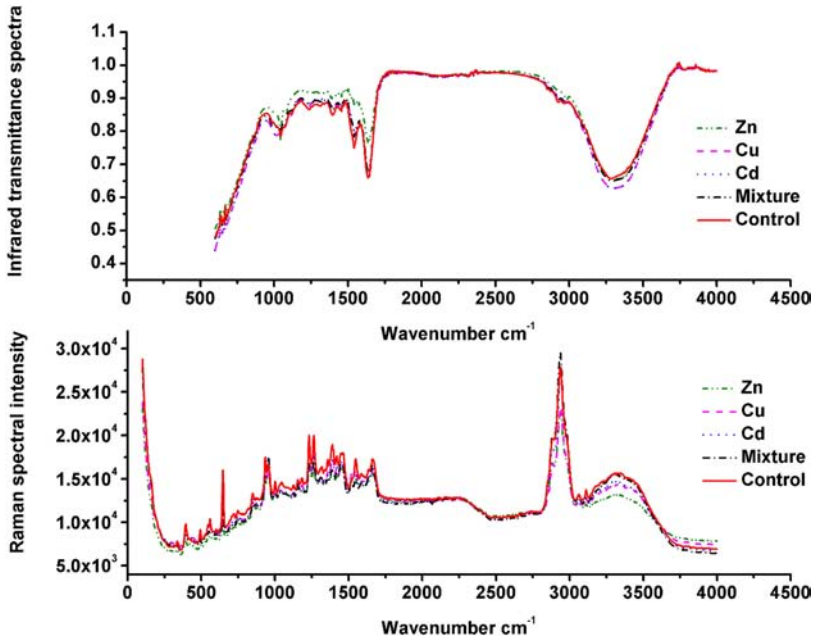

Figure 1. Representative infrared (a, top) and Raman spectra (b, bottom) of the control and heavy metal intoxicated samples of juvenile black seabream (*Sparus macrocephalus*).

intoxicated by Zn, Cu, Cd, and the mixture of three heavy metals and healthy samples, respectively.

The performance of the discrimination models was evaluated in terms of sensitivity, specificity, and accuracy for both calibration and prediction. Sensitivity is the number of positives (intoxicated samples) correctly classified by the model divided by the number of all positives. Specificity is defined as the number of negatives (healthy samples) correctly identified by the model divided by the number of all negatives. Accuracy is estimated as the number of correctly distinguished samples divided by the number of all samples. The threshold for the PLS-DA analysis was set to ±0.5. A sample with the predicted value within the actual integer number ±0.5 of this sample was determined as being correctly classified. All computations and chemometric analyses were operated using "The Unscrambler V9.7" (CAMOPROCESS AS, Oslo, Norway), and programs self-developed in Matlab 2011a software (The Mathworks, Inc., Natick, MA, USA).

RESULTS AND DISCUSSION

Analysis of Infrared and Raman spectra. Figure 1a shows the representative IR spectra of the control and intoxicated samples in the region of 4000−600 cm^{-1}. In general, the IR spectrum of fish was quite complex, containing several peaks contributed from different functional groups. The detailed wavenumber assignment was performed in three distinct ranges, namely, 3700−3000 cm^{-1}, 3000−2800 cm^{-1} (C−H stretching region of $-CH_2$ and $-CH_3$ groups of lipids and proteins), and 1800−950 cm^{-1} (fingerprint region). Specifically, the broad band centered at 3290 cm^{-1} in the 3700−3000 cm^{-1} region was assigned to the amide A, mainly N−H stretching mode of proteins.[18,19] In the region between 3000 and 2800 cm^{-1}, the absorptions observed at 2960 and 2874 cm^{-1} were caused by the CH_3 asymmetric and symmetric stretching groups. The former band is often used to determine the lipid structure, and the latter is mainly used to monitor proteins in the biological systems.[18−20] Another absorption band at 2930 cm^{-1} was assigned to the CH_2 asymmetric stretching of lipids.[18,19] The 1800−800 cm^{-1} region was generally dominated by the protein amide groups. Two sharp absorptions at 1640 and 1541 cm^{-1} were due to amide I and amide II vibration of structural proteins, respectively. The C═O stretching vibration of protein amide was mainly associated with the amide I absorption that is sensitive to protein conformation. The amide II absorption arises from amide N−H bending vibration (60%) coupled to C−N stretching vibration (40%) mode of the polypeptide and protein backbone.[20] The absorptions at 1456 and 1394 cm^{-1} were assigned to the CH_2 bending and COO− symmetric stretching modes, respectively.[18−20] The bands observed at 1237 cm^{-1} corresponds to PO_2^- asymmetric stretching of nucleic acids with little contribution from phospholipids.[19] The band observed at 1115 cm^{-1} was due to symmetric stretching of P−O−C,[21] and the band observed at 1044 cm^{-1} was assigned to C−O stretching vibrations in polysaccharides.[19]

Figure 1b shows the representative Raman spectra of the control and intoxicated samples in the region of 4000−100 cm^{-1}. Bands in the Raman spectra gave information based on both the relative amplitudes and frequencies of vibrational motions on amino acids, proteins, lipids, and nucleic acids. In general, some peaks from a variety of compounds exhibited strong Raman scattering, whereas others were weak. The band assignment of Raman spectra was carried out according to the literature.[5] A water broad Raman band at around 3300 cm^{-1} with the range of 3100−3500 cm^{-1} was attributable to O−H stretching motions. The strong and sharp peak at 2937 cm^{-1} was assigned to the CH stretching. In the amide I region, the Raman spectra at 1659 cm^{-1} with the range of 1650−1661 cm^{-1} was due to the C═O stretching mode of amide I of protein with high α-helical structure, whereas another peak centered at 1665 cm^{-1} with the range of 1661−1670 cm^{-1} was attributable to the amide I of protein with β-sheet structures and a high proportion of random coil or disorder structure.[5] The peak at 1659 cm^{-1} would also be attributable to the C═O stretching modes of fats and oils. Besides, the secondary structure of protein elucidated by amide I band could also use the C−C stretching vibrations between 900 and 1060 cm^{-1}. In detail, α-helices were assigned in 890−945 cm^{-1} and β-sheets were located in 1020−1060 cm^{-1}. In the amide III region,

vibrational spectroscopy of proteins produced a complex pattern of bands in the range of 1225−1350 cm^{-1}, where 1230−1245 and 1240−1255 cm^{-1} showed characteristics of β-sheet and random coil, respectively. An obvious peak at 1235 cm^{-1} was assigned to β-sheet of amide III. Raman bands observed in lipids near 1462, 1447, 1319, and 1263 cm^{-1} were assigned to the CH_2 scissoring modes, the C═C stretching modes, the CH_2 twisting modes, and the CH in-plane deformation modes. The peak at 1447 cm^{-1} was also attributed by the symmetric CH_2 and asymmetric CH_3 stretching modes of methylene. Other peaks of C−H deformations could also be found in the region of 1400−1500 cm^{-1}, such as 1462 and 1420 cm^{-1}. Other Raman characteristics included the COO− vibrational mode of carboxylate salt bridges (1390 cm^{-1}), the C−N stretching of proteins (1126 cm^{-1}), the C−H in-plane bending mode of phenylalanine (1045 cm^{-1}), tyrosine groups (850 cm^{-1}), and S−S stretching (492 cm^{-1}). In addition, several bands of tryptophan were found at 544, 559, 750, 875, 1001, 1343, 1360, 1548, and 1586 cm^{-1}.

Figure 1 indicates that both infrared and Raman spectra contained the spectral information of many chemical molecules and functional groups of constituents in fish, which was important to do the heavy metal discrimination. Some previous analyses have analyzed the structural changes of chemical/biological components linked with the metal intoxication caused by Zn, Cu, Cd, and their combinations in fish.[22−24] As both the healthy samples (negatives) and intoxicated samples (positives) were reared under the same conditions, the spectral differences were mainly caused by the intoxication of heavy metals. It should be noted that when more samples and/or more heavy metals were considered, their spectral profiles would be overlapped. It is difficult to identify an obvious visual difference of spectra caused by heavy metals. Thus, chemometrics including preprocessing and model calibration were employed for the discrimination purpose.

Analysis of Samples from Design I. Establishment of discrimination models for Zn analysis was executed using PLS-DA, LS-SVM, and RF algorithms based on the data of infrared spectra, Raman spectra, and their combination. Table S1 (Supporting Information) shows the discrimination results between healthy samples and those intoxicated by Zn. When infrared spectra were used for the model establishment, two models, namely, SG-1st Der-PLS-DA and SG-1st Der-RF, obtained the best discrimination results with 100% accuracy for both calibration and prediction. The discrimination results of Raman spectroscopy were poorer than those of infrared spectra, in which the prediction accuracy of all samples was <80%. The specificity of all samples was only 60% for prediction, showing that it was easy to classify healthy samples into intoxicated samples. The combination of infrared and Raman spectra did not improve the identification of Zn. Because of the poor discrimination results of Raman spectra, the involvement of Raman spectra and infrared spectra made the discrimination worse. Only the RF model with strategy I obtained 100% accuracy for both calibration and prediction. It was found that the calculation of SG-1st Der increased the discrimination accuracy in this case, compared with the analysis of the original spectra. Especially, the RF models established using the original spectra did not show good results for the analysis of Raman spectroscopy and spectral combination, but its SG-1st Der models had similar or better results than PLS-DA and LS-SVM models.

Analysis of design I (Cu) also considered three calibration algorithms and three spectral data sets like design I (Zn) did. Discrimination results between healthy samples and those intoxicated by Cu are shown in Table S2 (Supporting Information). Similar to the analysis of design I (Zn), infrared spectra outperformed Raman spectra in this case. There were five-sixths models for infrared spectra obtaining 100% accuracy for both calibration and prediction processes. On the other hand, discrimination using Raman spectra obtained the best results with 100% accuracy for calibration and 90% accuracy for prediction. Also, similar to the analysis of design I (Zn), the combination of two spectroscopies did not improve the discrimination performances. Only two models, namely, the LS-SVM model with strategy III and the RF model with strategy III, obtained 100% accuracy for both calibration and prediction.

In the discrimination between healthy samples and those intoxicated by Cd, the results were mostly acceptable with only a few models having slightly poorer results (Table S3, Supporting Information). There were a total of 21 models established, and 14 of them had 100% accuracy for calibration and 90% accuracy for prediction. The best discrimination result was obtained by the SG-1st Der-LS-SVM Raman model, yielding 100% accuracy for both calibration and prediction. The combination of two spectroscopies had no significant improvement compared with the use of single spectroscopy.

Analysis of Samples from Design II. Design I considered only the identification of one heavy metal. In design II, groups IV and V were considered where group IV included samples exposed to the mixture of Zn, Cu, and Cd. The results of 21 models established using PLS-DA, LS-SVM, and RF algorithms based on the data of infrared spectra, Raman spectra, and their combination are shown in Table S4 (Supporting Information). When infrared spectra were used for the model establishment, the PLS-DA models with or without preprocessing obtained the best discrimination results with 100% for both calibration accuracy and prediction accuracy. LS-SVM models had little worse discrimination than the PLS-DA models, in which the sensitivity of prediction decreased from 100 to 80%. RF models obtained the worst results with a prediction accuracy of only 70%. When Raman spectra were used for the model establishment, the best results with 100% accuracy for both calibration and prediction were obtained by the PLS-DA model with preprocessing and the LS-SVM model with preprocessing. Spectral preprocessing was proved to be efficient to improve the accuracy of the Raman analysis in this case. The result of the RF model was significantly improved when preprocessing was considered. The calibration accuracy was increased from 70 to 95%, and the prediction accuracy was increased from 60 to 100%. For the PLSR modeling, the consideration of preprocessing significantly decreased the number of latent variables from 16 to 2. In addition, although the infrared spectroscopy and Raman spectroscopy had good discriminations in some models when two spectroscopies worked separately, their combinations did not show much improvement. Because the amplitude of Raman spectra was much larger than that of infrared spectra, when strategy I was adopted, the infrared spectra had a little contribution. The results of models based on spectroscopic combination were similar to those based on Raman spectra. In strategy II, normalization was conducted to make the amplitudes of infrared and Raman spectra at the same level. However, strategy II was found to have poorer results compared with strategy I. To further

Table 1. Detailed Results of the RF Model with the Preprocessing of Strategy III for Design III Based on the Combined Data of Infrared and Raman Spectra[a]

		healthy	intoxicated by Zn	intoxicated by Cu	intoxicated by Cd	intoxicated by all three heavy metals
calibration	healthy	10	0	0	0	2
	intoxicated by Zn	0	10	0	0	0
	intoxicated by Cu	0	0	10	0	0
	intoxicated by Cd	0	0	0	9	0
	intoxicated by all three heavy metals	0	0	0	1	8
	accuracy (%)	100	100	100	90	80
prediction	healthy	4	0	0	0	1
	intoxicated by Zn	0	5	0	0	0
	intoxicated by Cu	0	0	5	0	0
	intoxicated by Cd	1	0	0	5	0
	intoxicated by all three heavy metals	0	0	0	0	4
	accuracy (%)	80	100	100	100	80

[a]Columns, actual classification of samples; rows, classification by model.

improve the discrimination results, strategy III was evaluated, in which the SG-1st Der was further carried out on the basis of the normalized spectra. It was found that strategy III efficiently combined two spectroscopies for the discrimination, especially for the LS-SVM and RF modeling.

Analysis of Samples from Design III. To obtain a more rigorous evaluation of infrared and Raman spectroscopies and their combination, their performance were determined on the basis of the combination of all the foregoing designs. The combined design (design III) included healthy samples, samples intoxicated with one heavy metal alone in design I, and samples intoxicated with three heavy metals in design II. When infrared spectra were used for the analysis, the results displayed in Table S5 (Supporting Information) show that the RF models had the best discrimination, followed by the LS-SVM models, whereas the PLS-DA models failed for the discrimination. The results of infrared spectra indicated that the nonlinear information in infrared spectra might be important for the discrimination, because the nonlinear calibration methods such as LS-SVM and RF performed better than the linear calibration method of PLS-DA. On the other hand, the preprocessing of SG-1st Der was found to be efficient for improving the discrimination in this case. The prediction accuracy was improved from 44 to 56%, from 64 to 76%, and from 80 to 88% for PLS-DA, LS-SVM, and RF methods, respectively. The best discrimination of infrared spectra was obtained by the RF model with preprocessing, where the calibration accuracy reached 90% and the prediction accuracy was 88%. When Raman spectra were considered, the models without preprocessing had prediction accuracy of <50%. The application of SG-1st Der improved the discrimination; especially the LS-SVM and RF models with preprocessing had prediction accuracy >70%. The best discrimination of Raman spectra was obtained by the RF model with preprocessing, where the accuracy was 76% for both calibration and prediction. In addition, RF models were found to be more robust than the other two methods, because there was not much difference of the RF models between the accuracy of calibration and prediction. However, other Raman models were overfitted as they had good discrimination for the calibration but poor results for the prediction.

When the infrared and Raman spectra were combined together on the basis of the three strategies, respectively, only two models, the LS-SVM model and the RF model, both with the preprocessing of strategy III, had prediction accuracies >80%. The PLS-DA models had poor discrimination results in this case, which was similar to the situations of the PLS-DA models based on either infrared or Raman spectra. Although the combination based on strategy II failed again as for design II, strategy III performed well for the LS-SVM model and the RF model. The best discrimination results for design III were obtained from the RF model with the preprocessing of strategy III, where both the calibration and prediction accuracies were >90%. The detailed discrimination results of this model are summarized in a confusion matrix presented in Table 1, in which the diagonal correspond to the correctly classified samples and the misclassified samples are shown on the off-diagonal. For the calibration and prediction in Table 1, the sum of the numbers in each column is the number of samples examined for each category. In the calibration, all 10 healthy samples were correctly classified to have the specificity of 100%. On the other hand, one sample intoxicated by Cd was misclassified as being intoxicated by all three heavy metals, and two samples intoxicated by all three heavy metals were misclassified as being healthy, accounting for 94% of correctly classified samples in the calibration. In the prediction, spectra of 25 samples were input to the established model. Of the five healthy samples, one was wrongly identified as being intoxicated by Cd. Another misclassified sample was the one intoxicated by all three heavy metals, which was wrongly identified as being healthy. Moreover, to evaluate the repeatability of the best discrimination model for design III (RF model with the preprocessing of strategy III), samples from a tank different from the previous work for each group were used as the prediction samples, whereas those from the remaining two tanks of the group were used as the calibration samples. The newly established model had the same accuracy for calibration (94%) and for prediction (92%) as the previously established model did. The difference between these two models was that the sensitivity and specificity for calibration were 95 and 90%, respectively, for the newly established model. The result shows that the best discrimination model for design III had a good repeatability and robustness.

Furthermore, of practical and useful interest is the ability of the method to distinguish between healthy and intoxicated samples, no matter by which heavy metal the latter is intoxicated. On this basis, samples from the calibration set for

design III were used for building a new RF calibration model with the preprocessing and strategy III. This model was then applied to distinguish the intoxication condition of the samples in the prediction set. The new RF model had a good calibration accuracy of 98% and a good prediction accuracy of 96%. In both the calibration and prediction processes, the intoxicated samples were correctly identified (e.g., with sensitivity of 100%), whereas only one healthy sample was wrongly identified as being intoxicated. As the misclassification of intoxicated samples as healthy samples is more serious than classifying healthy samples as intoxicated samples, the results of the new RF model were acceptable.

In general, good discriminations with 100% accuracy for both calibration and prediction were obtained for design I (Zn) and design I (Cu) by analyzing infrared spectra and for design I (Cd) and design II by analyzing Raman spectra. The spectroscopic combination did not show its advantages for designs I and II, but obtained the best discrimination for design III, which was much more complicated than other designs. Therefore, the spectroscopic combination was not required in this work for two-class problems (designs I and II), but was suggested for the identification of samples intoxicated by different heavy metals (design III). On the other hand, nonlinear calibration methods were found having better performances for complicated sample sets than linear methods for the identification of samples intoxicated by different heavy metals. Although RF did not show a distinct advantage over the other two calibration methods in designs I and II, it had better discriminations in design III, especially with the help of preprocessing. Actually, the spectral preprocessing of SG-first Der was found to be able to improve the discrimination in many cases and could keep the discrimination accuracy in other cases. In addition, it was found that strategy II was not efficient, as it made the discrimination worse. Strategy III did show some efficiency in the spectral combination of two spectroscopies, especially for design III. Strategy I gave results to similar those of the corresponding Raman model, as the amplitude of Raman spectra was between 5000 and 30000 and that of infrared spectra was only between 0 and 1.

The ability of infrared and Raman spectroscopies in identifying juvenile black seabream intoxicated by heavy metals lies in their provided spectral information related to the composition changes caused by the intoxication of heavy metals. Some works suggested that heavy metals could induce alteration on the major biochemical constituents such as lipids, proteins, and nucleic acids, which can be evidenced by infrared and Raman spectroscopies.[22−24] In this work, on the basis of previous works, the feasibility of infrared and Raman spectroscopies was evaluated on distinguishing healthy juvenile black seabream and those with accumulated heavy metals. The results presented in this work show that component changes induced by heavy metals in juvenile black seabream could be reflected by infrared and Raman spectroscopies. The intensity of infrared and Raman bands suggests the possible involvement of complex physical and chemical changes caused by the intoxication of heavy metals. Discrimination models established using PLS-DA, LS-SVM, and RF were reasonably efficient as rapid and convenient tools for identifying fish samples polluted by heavy metals. One hundred percent accuracy was obtained for identifying heavy metals in two-class problems for designs I and II, and >90% accuracy was obtained for design III, which was a combination of designs I and II.

Heavy metals are naturally occurring, but their discharges into natural ecosystems have increased significantly in the modern world due to anthropogenic activities, such as mining, automobiles, electroplating, paints and dye, coal burning, battery-making industries, and trash incineration. Remediation of heavy metals poses a different kind of challenge compared with organic pollutants because of heavy metals' indestructible nature through bioremediation and their stable and persistent existence. Another important factor that contributes to the deleterious effects of heavy metals as pollutants is their tendency to accumulate in the environment, especially in the bottom sediments of aquatic habitats in association with organic and inorganic matter.[25] Heavy metals would cause potential disease when they accumulate in the body.[26] Due to their toxicity, long persistence, and nonbiodegradable properties, it is important to identify fish intoxicated by heavy metals in a rapid and accurate way. The successful output of this work shows that healthy and intoxicated juvenile black seabream could be well distinguished, leading to improved safety assurance of the fish product. This study opens up an attractive prospect that fish intoxicated by heavy metals could be identified within seconds as opposed to hours or days. As the first research on rapid and noninvasive discrimination of healthy juvenile black seabream and those intoxicated by heavy metals, the results obtained here are very promising and will promote more efforts on investigating infrared and Raman spectroscopies for assessing heavy metal contamination of fish and other food products. In the next step, quantitative relationships will be established between the IR and Raman spectra of intoxicated fish samples and their contained concentration of heavy metals, the reference values of which will be measured using traditional methods, those being atomic absorption spectroscopy, etc.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional tables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*(X.G.) Phone: +86 577 86689027. Fax: +86 577 86689027. E-mail: guanxc@wzu.edu.cn.

*(G.L.) E-mail: gliu@wzu.edu.cn.

Funding

We acknowledge the financial support provided by the National Natural Science Foundation of China (31201355).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Horiba Jobin Yvon for the collection of Raman spectra.

■ REFERENCES

(1) Burger, J.; Gaines, K. F.; Boring, C. S.; Stephens, W. L.; Snodgrass, J.; Dixon, C.; McMahon, M.; Shukla, S.; Shukla, T.; Gochfeld, M. Metal levels in fish from the Savannah River: potential hazards to fish and other receptors. *Environ. Res.* **2002**, *89*, 85−97.

(2) Sloman, K.; Baker, D.; Ho, C.; McDonald, D.; Wood, C. The effects of trace metal exposure on agonistic encounters in juvenile rainbow trout *Oncorhynchus mykiss*. *Aquat. Toxicol.* **2003**, *63*, 187−196.

(3) Tüzen, M. Determination of heavy metals in fish samples of the middle Black Sea (Turkey) by graphite furnace atomic absorption spectrometry. *Food Chem.* **2003**, *80*, 119−123.
(4) Uddin, M.; Okazaki, E. Fish and related products. In *Infrared Spectroscopy for Food Quality Analysis and Control*, 1st ed.; Sun, D.-W., Ed.; Academic Press, Elsevier: San Diego, CA, USA, 2008; pp 215−240.
(5) Herrero, A. M. Raman spectroscopy a promising technique for quality assessment of meat and fish: a review. *Food Chem.* **2008**, *107*, 1642−1651.
(6) Zhu, X.; Huang, G.; Luo, S.; Guan, X.; Chen, X. Rapid determination of enantiomeric excess of *tert*-butoxycarbonyl (Boc-protected) amino acids based on infrared spectra technique with optimal wavelet packet transform decomposition frequency band. *Anal. Lett.* **2013**, *46*, 671−681.
(7) Schmidt, H.; Scheier, R.; Hopkins, D. L. Preliminary investigation on the relationship of Raman spectra of sheep meat with shear force and cooking loss. *Meat Sci.* **2012**, *93*, 138−143.
(8) Wu, D.; Sun, D.-W. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review − part I: fundamentals. *Innovative Food Sci. Emerging Technol.* **2013**, *19*, 1−14.
(9) Wu, D.; Chen, J.; Lu, B.; Xiong, L.; He, Y.; Zhang, Y. Application of near infrared spectroscopy for the rapid determination of antioxidant activity of bamboo leaf extract. *Food Chem.* **2012**, *135*, 2147−2156.
(10) Chen, X.; Wu, D.; He, Y.; Liu, S. Detecting the quality of glycerol monolaurate: a method for using Fourier transform infrared spectroscopy with wavelet transform and modified uninformative variable elimination. *Anal. Chim. Acta* **2009**, *638*, 16−22.
(11) Chen, X. J.; Lei, X. X. Application of a hybrid variable selection method for determination of carbohydrate content in soy milk powder using visible and near infrared spectroscopy. *J. Agric. Food Chem.* **2009**, *57*, 334−340.
(12) Wu, D.; Feng, L.; He, Y.; Bao, Y. Variety identification of Chinese cabbage seeds using visible and near-infrared spectroscopy. *Trans. ASABE* **2008**, *51*, 2193−2199.
(13) Suykens, J. A. K.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293−300.
(14) Suykens, J. A. K.; Van Gestel, T.; De Brabanter, J.; De Moor, B.; Vandewalle, J. *Least Squares Support Vector Machines*; World Scientific: Singapore, 2002.
(15) Wu, D.; Sun, D.-W. Potential of time series-hyperspectral imaging (TS-HSI) for non-invasive determination of microbial spoilage of salmon flesh. *Talanta* **2013**, *111*, 39−46.
(16) Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5−32.
(17) Rodriguez-Galiano, V.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.; Jeganathan, C. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93−107.
(18) Akkas, S.; Severcan, M.; Yilmaz, O.; Severcan, F. Effects of lipoic acid supplementation on rat brain tissue: an FTIR spectroscopic and neural network study. *Food Chem.* **2007**, *105*, 1281−1288.
(19) Cakmak, G.; Togan, I.; Severcan, F. 17β-Estradiol induced compositional, structural and functional changes in rainbow trout liver, revealed by FT-IR spectroscopy: a comparative study with nonylphenol. *Aquat. Toxicol.* **2006**, *77*, 53−63.
(20) Dogan, A.; Siyakus, G.; Severcan, F. FTIR spectroscopic characterization of irradiated hazelnut (*Corylus avellana* L.). *Food Chem.* **2007**, *100*, 1106−1114.
(21) Dovbeshko, G. I.; Gridina, N. Y.; Kruglova, E. B.; Pashchuk, O. P. FTIR spectroscopy studies of nucleic acid damage. *Talanta* **2000**, *53*, 233−246.
(22) Palaniappan, P.; Renju, V. FT-IR study of the effect of zinc exposure on the biochemical contents of the muscle of *Labeo rohita*. *Infrared Phys. Technol.* **2009**, *52*, 37−41.
(23) Palaniappan, P.; Vijayasundaram, V. Fourier transform infrared study of protein secondary structural changes in the muscle of *Labeo rohita* due to arsenic intoxication. *Food Chem. Toxicol.* **2008**, *46*, 3534−3539.
(24) Palaniappan, P.; Pramod, K. Raman spectroscopic investigation on the microenvironment of the liver tissues of zebrafish (*Danio rerio*) due to titanium dioxide exposure. *Vib. Spectrosc.* **2011**, *56*, 146−153.
(25) Sobha, K.; Poornima, A.; Harini, P.; Veeraiah, K. A study on biochemical changes in the fresh water fish, *Catla catla* (Hamilton) exposed to the heavy metal toxicant cadmium chloride. *Kathmandu Univ. J. Sci., Eng. Technol.* **2007**, *3*, 1−11.
(26) Palaniappan, P.; Vijayasundaram, V.; Prabu, S. M. A study of the subchronic effects of arsenic exposure on the liver tissues of *Labeo rohita* using Fourier transform infrared technique. *Environ. Toxicol.* **2011**, *26*, 338−344.