

RESEARCH ARTICLE

Differentiation and authentication of fishes at the species level through analysis of fish skin by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

Hongyan Bi¹  | Chunyi Zhong¹ | Mingke Shao¹ | Chengyu Wang¹ | Jia Yi² | Liang Qiao² | Junbo Zhang¹

¹ College of Food Science and Engineering, and College of Marine Sciences, Shanghai Ocean University, Hucheng Ring Road 999, Pudong New District, 201306 Shanghai, China

² Department of Chemistry, Fudan University, Songhu Road 2005, Yangpu District, 200438 Shanghai, China

Correspondence

H. Y. Bi, College of Food Science and Engineering, and College of Marine Sciences, Shanghai Ocean University, Hucheng Ring Road 999, Pudong New District, 201306 Shanghai, China.
Email: hybi@shou.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Number: 21804087; Science and Technology Commission of Shanghai Municipality, Grant/Award Number: 18050502200

Rationale: Authentication of fish is of importance in the view of toxins, allergen warnings and economic fraud control. Traditional methods in the authentication of fish, e.g. morphological, genetic and proteomic analysis, are either at low throughput or at high-cost.

Methods: A high-throughput matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF MS)-based approach was developed to analyze biomaterials from fish skin, and mass spectra from different fish species were compared by chemometric methods to differentiate fish species.

Results: A total of 51 fish samples were used to generate more than 150 fingerprinting mass spectra. The fish belonging to the same genus can be identified at species level. A mass spectral database of different fishes can be built as reference for authentication. The analysis can be performed based on micrograms of fish-skin sample and accomplished in 1–3 hours.

Conclusions: The developed strategy holds potential to be applied to fish authentication in the fishing industry and as a scientific method to avoid mislabeling. It has promise to be practically used for fast and effective identification of closely related fish species to guarantee the quality of fishery products to consumers.

1 | INTRODUCTION

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF MS) is a technique developed in late 1980s to characterize large biomolecules,¹ e.g. proteins, peptides, lipids, polynucleotides, sugars, etc. Compared with liquid chromatography (LC)/MS, MALDI TOF MS analyzes dried samples on a target plate, and generates mass spectra over a large mass range. The technique holds advantages of relatively simple sample pretreatment and high throughput in analysis. MALDI TOF MS is also widely used to generate fingerprint mass spectra for the identification of microorganisms.^{2–4} Biological molecules, typically small proteins originated from cell surfaces, intracellular membranes

and ribosomes, constitute the fingerprint mass spectrum of a bacterium.⁵ Fingerprint mass spectra acquired from bacterial samples are usually compared with reference spectra of purified known strains for bacterial identification. With a similar concept, MALDI TOF MS is also a promising method for the authentication of meat and fish products.

Fish and fishery products are one of the most highly traded commodities in the world. Due to the globalization of the market, economically motivated adulteration and/or fraudulent labeling are becoming important concerns.⁶ According to the estimated value by the Food and Agriculture Organization of the United Nations (FAO), the annual global seafood industry is worth almost 120 billion dollars. The FAO reports that by virtue of seafood mislabeling this

equates to fraud of the 120 billion dollars to be worth 400 billion dollars.⁷⁸ *Time* magazine reported that 58% of fraudulent seafood products in US were species that could cause health complications because of toxins and allergens. Besides, mislabeling illegally caught fish or vulnerable fish may evade supervision from authorities.⁸ To prevent economic fraud, possible health hazards, and potential ecological destruction, it is highly important to develop rapid, simple, economical and reliable techniques for fish species authentication. Identification of fish traditionally relies on morphological analysis, including body shape, pattern of colors, scale size and count, number and relative position of fins, number and type of fin rays, or various relative measurements of body parts.⁹ Genetic methods have been used to identify species. DNA is largely independent of tissue type, age and status, and presents in all cell types. Genetic methods, such as polymerase chain reaction (PCR), show high specificity and sensitivity.⁹ However, the genetic methods usually need complicated pretreatment to extract DNA from original fish samples, which is labor-intensive and time-consuming. Selection of appropriate primers for DNA amplification is another issue to be addressed for successful identification of fish.¹⁰ Furthermore, the risk of cross-contamination is linked to the intrinsic sensitivity of the methods, which can be exacerbated by the generation of billions of copies of targeted sequences during gene amplification.¹¹

Two-dimensional electrophoresis (2-DE) profiles of water-soluble muscular proteins have been used for fish discrimination,¹² but with difficulty in the interpretation of the 2-DE image.¹³ Bossier et al isolated water-soluble proteins from fish muscle tissue and used the isoelectric focusing (IEF) technique to identify fish.¹⁴ IEF patterns of various specimens were used to form a databank which can be considered as reference, and used for identifying flatfish products at the species level.¹⁴ Ortea et al used reversed-phase capillary liquid chromatography coupled with tandem mass spectrometry to analyze peptides digested from shrimps for shrimp identification.¹⁵

Proteomic methods were also developed for seafood identification. Proteomic methodologies usually authenticate fish by targeting species-specific peptide with liquid chromatography coupled with tandem mass spectrometry (LC/MS/MS).¹⁶ Large-scale comparison of tandem mass spectra for authentication of fish products was proposed by Palmblad et al in 2013.¹⁷ In 2016, an untargeted data acquisition mode was utilized by Palmblad et al to build a spectral library to identify unknown samples by comparing the LC/MS-acquired data with the spectral library, which allows fish authentication at the genus or family level.¹⁸ The developed proteomic methods for fish authentication mostly need protein extraction, digestion, and/or separation procedures prior to MS analysis.^{15,16} For LC-based methods, the pretreatment of samples, the optimization of chromatographic conditions, and the choice of proper columns must be carefully considered. To date, the proteomics-based fish identification methods are still time-consuming, expensive, and complicated in methodology. Closely related seafood species are still difficult to be identified by morphological characteristics. Studies have demonstrated that identification of closely related fish species can be accomplished by

using MS/MS,^{17,18} and that the detection confidence is supposed to increase when traveling up the taxonomy since the difference of fishes in between is getting bigger. However, studies on fish authentication at the species level based on mass spectrometric techniques still needs further development.

To date, there are few studies that have used MALDI TOF MS to identify fish. Mazzeo et al developed a strategy for fish authentication by using MALDI TOF, where proteins extracted from fish muscle with molecular weights around 11 kDa were chosen as species-specific biomarkers.¹⁹ However, the differences in mass-to-charge ratio (m/z) between the selected biomarkers for fishes belonging to the same genus were very small, e.g. 0.5 m/z , which are too small to be well distinguished by a routine linear TOF mass analyzer.

Herein, we developed a rapid and facile strategy based on skin samples of fish to identify fishes at the species level for the first time by using MALDI TOF MS to collect fingerprint mass spectra of different fishes and multivariate statistical analysis methods to classify the collected mass spectra. Only simple pretreatments, including soaking fish skin in alkaline solution for bio-extraction and desalting, were performed before MALDI TOF measurements. Closely related species from the same genera, e.g. *Larimichthys crocea* (large yellow croaker) and *Larimichthys polyactis* (small yellow croaker), *Cynoglossus abbreviatus*, *Cynoglossus gracilis*, *Cynoglossus joyneri* and *Cynoglossus semilaevis*, can be well distinguished by the method. Fishes from a local seafood market, *Pampus argenteus*, *Trachinotus ovatus*, and *Zeus faber*, were further analyzed, demonstrating that the strategy can be used to classify different fish species. The analysis can be performed based on fish-skin samples and accomplished in 1–3 h, meaning it has the potential to be applied to fast authentication in the fishing industry. With the mass spectral datasets of blind samples and known samples, we demonstrated that the strategy could be further used to identify unknown fish samples at the species level. Therefore, the method holds promise for practical application for fast and effective fish authentication to guarantee the quality of fishery products.

2 | EXPERIMENTAL

2.1 | Chemicals and materials

Sodium hydroxide ($\geq 96\%$, NaOH) was from Shanghai Dahe Chemical Co., Ltd (Shanghai, China). Hydrochloric acid (36.0–38.0%, HCl) was from Sinapharm Chemical Reagent Co. Ltd (Shanghai, China). Acetonitrile (ACN, 99.9%) and trifluoroacetic acid (TFA, HPLC grade) were from Beijing Bailingwei Technology Co., Ltd (Beijing, China). Methanol (gradient grade for liquid chromatography) was from Merck (Darmstadt, Germany). Sinapinic acid (SA, $\geq 99.0\%$) was from Bruker (Bremen, Germany). Cytochrome C from equine heart ($\geq 95\%$) and myoglobin from equine skeletal muscle (95–100%), insulin from bovine pancreas, and ubiquitin human ($\geq 95\%$) were from Sigma-Aldrich Inc. (St Louis, MO, USA). Deionized (DI) water

(18.2 M Ω -cm) was purified using a Smart-Q deionized water system (Hitech Pure Water Technology, Shanghai, China) and used in all aqueous solutions.

2.2 | Preparation of fish skin samples

Authenticated tonguefishes, *Cynoglossus abbreviates*, *Cynoglossus gracilis*, *Cynoglossus joyneri*, *Cynoglossus semilaevis*, and *Collichthys lucidus*, used in the present study were caught off the Huanghai Sea of China at 34°53'44"N and 120°03'50"E, and two species of yellow croaker (*Larimichthys crocea*, *Larimichthys polyactis*) were collected in the Donghai Sea of China at 30°29'48"N, 122°47'46"E. Dead fish samples were obtained from a local fisherman. Authenticated *Pampus argenteus*, *Trachinotus ovatus*, and *Zeus faber* were obtained from a local seafood market (Shanghai Oriental International Fisheries Market, Shanghai, China). All the fish samples were stored in ice, and transported to the Shanghai Ocean University within 36 h. The fish samples were stored at -20°C. The skin of fish was taken within 2 months after fish storage. The skin of fish was collected, and the scales and meat were removed from the skin with a scalpel if necessary, and then thoroughly flushed with DI water. The cleaned fish skins were sealed in a plastic bag and stored at -20°C until use.

2.3 | Extraction of non-collagen proteins in fish skin

Fresh or thawed fish skin samples were flushed by DI water. The extra water was removed by paper towels. Samples of fish skins (0.2 g) were weighed, placed in a 15 mL centrifuge tube, and mixed with 2 mL of 1.3 M NaOH solution. The mixture was incubated under 25°C with the assistance of a perpendicular mixer (HS-3, Shanghai Chuanxiang Bio-technology Co., China) at its maximum rotational speed (60 rpm) for different periods of time. The NaOH treatment can extract non-collagen proteins from fish skin according to a previous study.²⁰

2.4 | Sample desalting

From the above resultant liquid 150 μ L were taken, and neutralized with HCl: ca. 20 μ L of 6 M HCl was needed to adjust the pH to 7 or to a weak acidic state. The sample was centrifuged in case there was precipitate after pH adjustment on a centrifuge (Himac CT 15RE, Hitachi, Tokyo, Japan), and the supernatant was taken for further treatment. Pierce C18 spin columns (Thermo Scientific) were utilized to clean salts from the extracted samples. The desalted samples were stored at -20°C or directly analyzed by MALDI TOF MS.

2.5 | MALDI TOF MS analysis

Desalted fish skin protein samples (2 μ L) were deposited on MALDI plates and dried at ambient condition; then 1 μ L of 15 mg/mL SA in 50% ACN, 49.9% H₂O and 0.1% TFA was deposited on the same spot and dried under ambient conditions. When the matrix had

dried, MALDI TOF MS analysis was performed on a MicroFlex LRF mass spectrometer (Bruker) in linear positive mode. Instrumental parameters were: 75% laser intensity, laser attenuator with 35% offset and 40% range, accumulation from 400 laser shots, 10.3 $\times$ detector gain, and 350 ns delayed extraction time. An aqueous solution containing cytochrome c, myoglobin, insulin and ubiquitin was used for external mass calibration.

2.6 | Statistical analysis

MS raw data were processed using the R package MALDIquant.²¹ Peak lists were extracted from the processed MS data with the information of *m/z* value and relative intensity for the peaks with signal-to-noise ratio (S/N) > 3 as aligned matrices. Multivariate statistical analysis to the matrices was performed by using SIMCA-P 13.0 (MKS Umetrics, Umeå, Sweden) and MetaboAnalyst 3.0 (McGill University, Montreal, Canada).

2.6.1 | Safety considerations

Note that sodium hydroxide is irritating and corrosive. Exposure to sodium hydroxide (solid or solution) can cause skin and eye irritation.

3 | RESULTS AND DISCUSSION

3.1 | Mass spectrometric analysis of fish skin sample

To the best of the authors' knowledge, there is to date no research on fish authentication by the analysis of biomaterials from fish skin by using MALDI TOF MS. Herein, we endeavored to develop a straightforward method for fish authentication by virtue of fish skin and the MALDI MS technique. Figure 1 shows the workflow of the method. *Larimichthys crocea* (large yellow croaker, short as LC) and *Larimichthys polyactis* (small yellow croaker, short as LP), belonging to the order *Perciformes*, the family *Sciaenidae* and the genus *Larimichthys*, were taken as model samples to illustrate the workflow. The morphological features of the two species are very similar, as shown in Figure S1 (supporting information). Due to the similarity in morphology of the two species, disguising between LC and LP often happens for economic purposes, which is one of the seafood mislabeling phenomena.

Six fish of each species were taken to prepare the samples for mass spectrometric measurement. Table S1 (supporting information) lists the details of the LC and LP utilized in the present work, including size and weight. The extraction of fish skin proteins by 1.3 M NaOH solution was performed overnight. It is known that boiling proteins in strong acid for 18–96 h, or in 2–4 M NaOH for 4–8 h, can lead to a complete hydrolysis of proteins.²² In the present study, the way of extracting biomaterials from fish skin is relatively mild without any heating. Therefore, proteins and polypeptides are supposed to be obtained rather than amino acids. Based on a previous study, non-collagenous proteins from fish skin

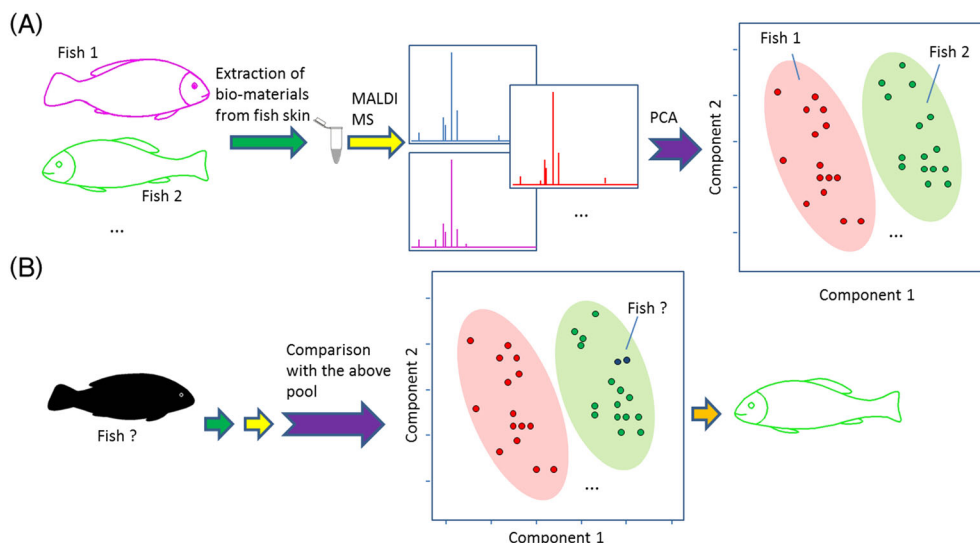


FIGURE 1 A, Workflow for the classification of fish by analyzing bio-samples extracted from fish skin using MALDI TOF MS and chemometric methods. B, Scheme of fish identification based on MALDI TOF mass spectra of unknown samples and the reference built by MALDI TOF mass spectra of known species. The mass spectra and results of chemometrics analysis in the figure are not real data [Color figure can be viewed at wileyonlinelibrary.com]

are supposed to be extracted by the present method.²⁰ The obtained biomaterials were desalted and then analyzed by MALDI TOF MS. Three common matrices for proteins and peptides, i.e. 2,5-dihydroxybenzoic acid (DHB), α -cyano-4-hydroxycinnamic acid (CHCA), and 3,5-dimethoxy-4-hydroxycinnamic acid (SA), were tested, and SA was chosen for subsequent experiments, as it helped to generate mass spectra from the biomaterials with best quality when considering the number and intensity of peaks.

Figure 2 shows the representative mass spectra of samples from LC and LP obtained in positive ion mode. Peaks with m/z from 1000 to 5000 can be well observed. The peaks are originated from proteins of fish skin or fragments of the proteins. As labeled in Figure 2, distinctive peaks at $m/z = 1162, 1223, 2190, 2132, 2426, 2654, 3056$ for LC and $m/z = 1156, 2183, 2419, 2651, 3048$ for LP can be observed, showing that this is a promising way to distinguish LC and LP based on the mass spectra.

3.2 | Fish classification based on their MALDI TOF mass spectral datasets

During the analysis of large yellow croaker (*Larimichthys crocea*, LC) and small yellow croaker (*Larimichthys polyactis*, LP), a dataset consisting of 36 mass spectra corresponding to 12 fish samples listed in Table S1 (supporting information) were obtained. The mass spectra have a similar pattern, as shown in Figure 2. It is not easy to distinguish different fish samples by virtue of directly observing the mass spectra. Multivariate analysis, e.g. principal component analysis (PCA) and partial least-squares regression discriminant analysis (PLS-DA), was employed to distinguish the mass spectra from different samples. PCA is a chemometric method, and can reduce the dimensionality of large datasets by creating new uncorrelated

variables to increase the interpretability of the data and in the same time minimize information loss.²³ The reduced set of variables may favor the identification and description of data in a new pattern or in another sense. Nowadays, PCA is a data analysis tool that has been explored for many applications in numerous disciplines, e.g. color technology,²⁴ water studies,²⁵ and wine quality control.⁶ PLS-DA modeling, which is a supervised method, is more powerful to discriminate groups of datasets.²⁶ Figure 3A shows the PCA score

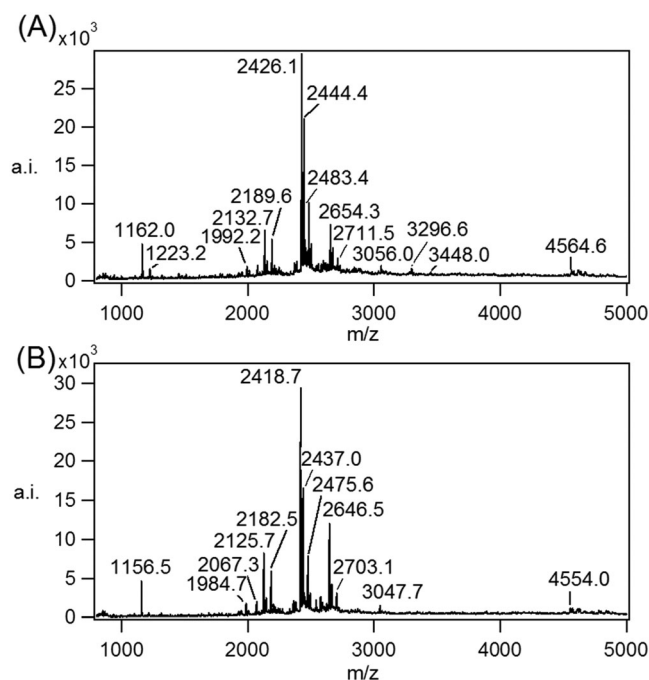


FIGURE 2 Representative MALDI TOF mass spectra of bio-materials extracted from the skin of A, *Larimichthys crocea* and B, *Larimichthys polyactis*. a.i.: Arbitrary intensity. m/z : mass to charge

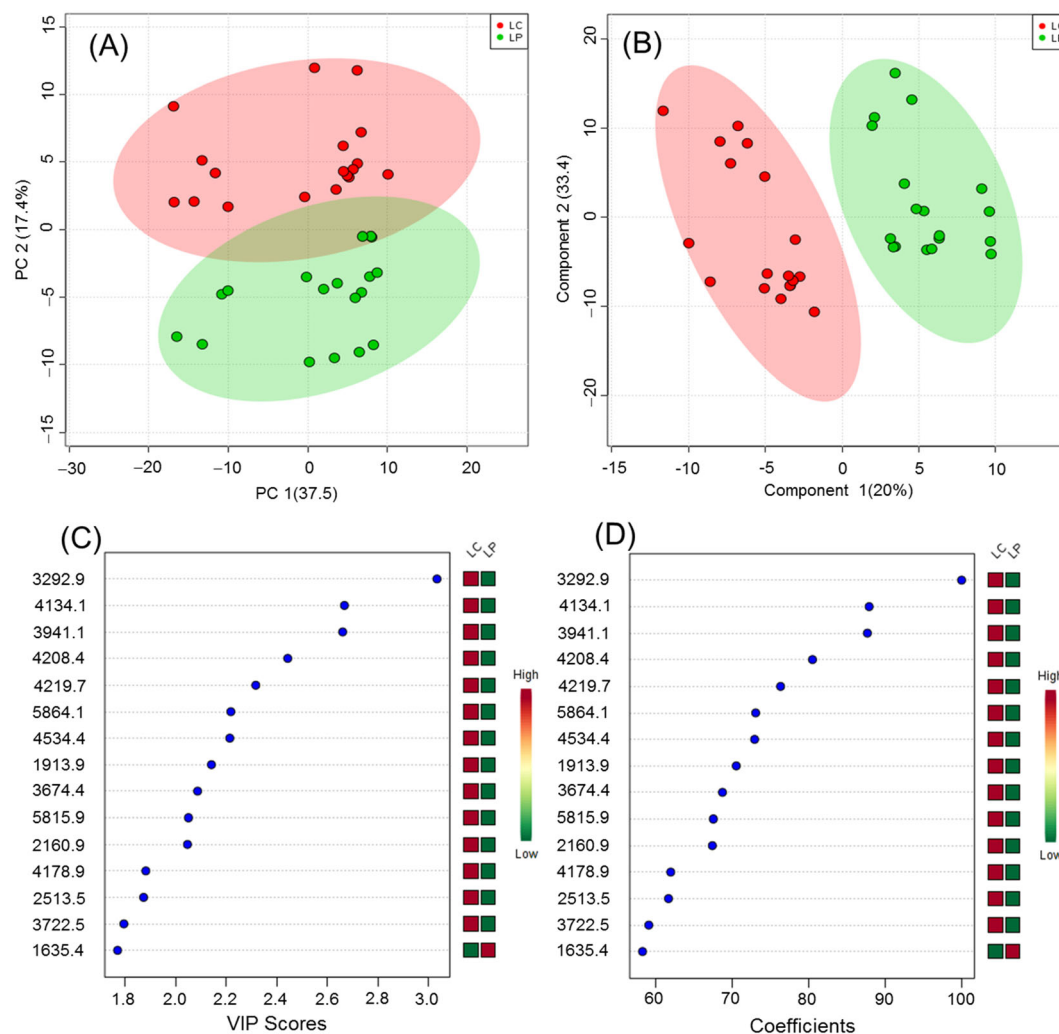


FIGURE 3 Score plots of A, PCA and B, PLS-DA for MALDI TOF mass spectra of *Larimichthys crocea* and *Larimichthys polyactis*, respectively. Plots of C, VIP and D, weighted sum of absolute regression coefficients from PLS-DA. The colored boxes on the right indicate the relative abundance of the peaks for each fish under mass spectrometric study. Three mass spectra were collected for each fish from three different skin samples [Color figure can be viewed at wileyonlinelibrary.com]

plot for the mass spectra from LC and LP. The ellipses in the plot display areas with 95% confidence level, which in statistics is a measure of the reliability of a result.

It can be seen that the two species can be well distinguished by PCA with their MALDI TOF mass spectra with small overlap within the 95% confidence level. Using PLS-DA to analyze the same mass spectrometric datasets, the two species can be clearly discriminated within the 95% confidence level (Figure 3B). The results show that the biomaterials extracted from fish skins exhibit highly obvious differences during MALDI TOF analysis.

Cross validation (CV) is often used to assess the prediction performance of a model and to see how well the model predicts on test data. Figure S2 (supporting information) shows the CV of the MALDI TOF mass spectra datasets collected from LC and LP obtained from PLS-DA. Q2, calculated via CV, is an estimate of the predictive ability of the model. As shown in Figure S2 (supporting information), Q2 has an average value >0.90, indicating that it can predict (to within >90% accuracy) an unknown

sample such as LC or LP by using the mass spectrometric data of the unknown sample and the mass spectrometric data from known samples.

The plot of variable importance in projection (VIP) scores can be obtained from PLS-DA. High VIP scores usually indicate a high significance of the corresponding mass spectrometric signals (labeled with m/z values on the figure) on the classification of the samples. Fifteen mass spectrometric signals with highest VIP scores and PLS regression coefficients are listed in Figures 3C and 3D. The colored boxes on the right of Figures 3C and 3D indicate the relative abundance of the corresponding mass spectrometric signals in each fish group under study.

To further validate the reliability of the present strategy, four tonguefishes, *Cynoglossus abbreviatus* (CA), *Cynoglossus gracilis* (CG), *Cynoglossus joyneri* (CJ), and *Cynoglossus semilaevis* (CS), from the genus *Cynoglossus*, were analyzed. Figure 4 shows the PLS-DA score plot and VIP scores of the four species, showing that the fishes can be discriminated at the species level within the 95% confidence level

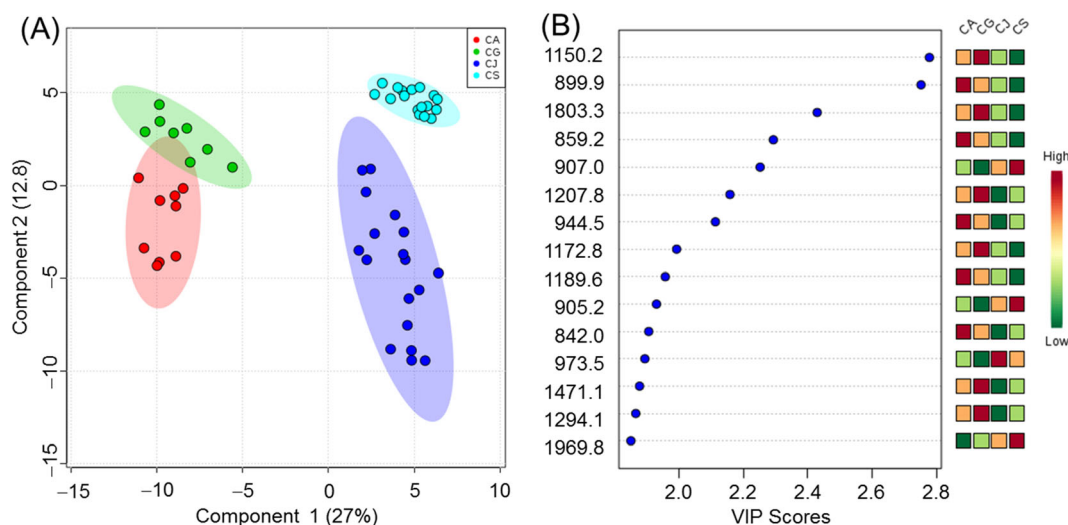


FIGURE 4 A, Score plot of PLS-DA for MALDI TOF mass spectra of skin protein samples of four tonguefishes and B, 15 mass spectrometric signals with highest VIP values among the tonguefish samples of *Cynoglossus abbreviatus* (CA), *Cynoglossus gracilis* (CG), *Cynoglossus joyneri* (CJ), and *Cynoglossus semilaevis* (CS). The colored boxes on the right of the VIP scores plots indicate the relative abundances of the corresponding mass spectrometric signals in each fish under study. 3 CA fish, 3 CG fish, 6 CJ fish and 6 CS fish were taken for MALDI TOF analysis. Three mass spectra were collected for each fish from three different skin samples [Color figure can be viewed at wileyonlinelibrary.com]

by MALDI TOF and PLS-DA. As shown in the VIP scores in Figure 4, the relative abundances of the corresponding mass spectrometric signals in each species are completely different, illustrating the difference between the four species. The results further demonstrate that the developed strategy can authenticate fish belonging to the same genus, i.e., at the species level.

When a method can classify fishes at the species level, it is supposed to be capable of discriminating fishes at higher level (e.g. family or order) as well, since the difference between fishes from different genera is normally much larger than those belonging to the same genus. As listed in Tables S2 and S3 (supporting information), *Collichthys lucidus* (CL), which belongs to the same family as LC and LP, i.e. *Sciaenidae*, was selected to demonstrate the classification power of the present strategy for fishes at the genus level. Figure 5A shows the PCA scores plot of the three

species, illustrating that LC and LP are much more similar when compared with CL. PLS-DA scores plot can classify the three species within the 95% confidence level. The 15 MS signals with the highest VIP scores are listed in Figure 5C, illustrating that the relative abundance of these ions for the three species is different, which can be ascribed to the difference between the three species. The results further demonstrate that the present method can classify different fish species. Moreover, the fish species *Pampus argenteus*, *Trachinotus ovatus*, and *Zeus faber*, obtained from a local seafood market, were analyzed to further demonstrate the present strategy for fish classification. Figure S3 (supporting information) shows the mass spectra of these fishes, and Figure S4 shows PCA score plots for the fishes. It can be seen that the fishes can also be discriminated, further demonstrating that the present strategy can be used to classify different fishes.

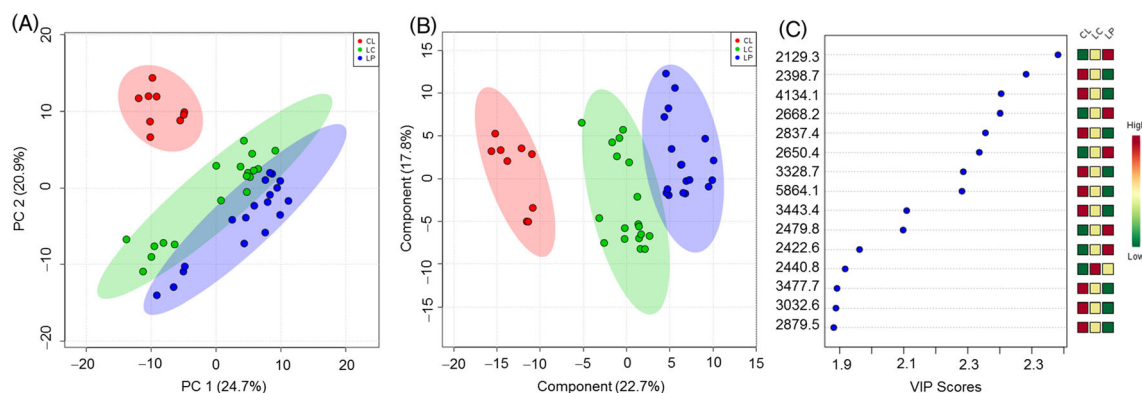


FIGURE 5 Score plots of A, PCA and B, PLS-DA for MALDI TOF mass spectra of fishes from the family *Sciaenidae*; C, Measure of scores of VIP in PLS-DA. The colored boxes on the right indicate the relative abundances of the corresponding mass spectrometric signals in each fish under study. CL, LC and LP in the inset indicate the fish of *Collichthys lucidus* (3 fish), *Larimichthys crocea* (6 fish) and *Larimichthys polyactis* (6 fish), respectively. Three mass spectra were collected for each fish from three different skin samples [Color figure can be viewed at wileyonlinelibrary.com]

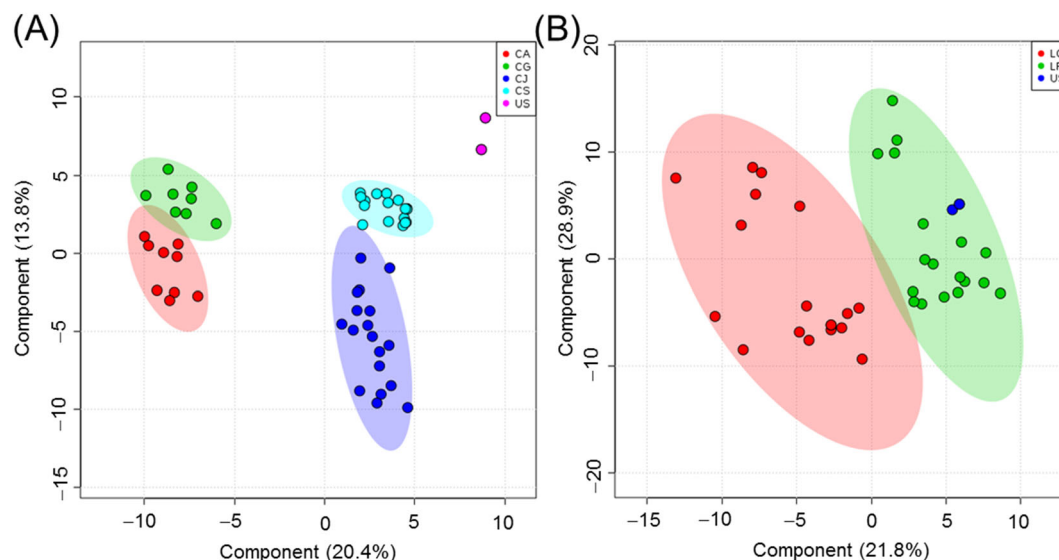


FIGURE 6 A, Score plots of PCA for MALDI TOF mass spectra of CA, CS, CG, CJ and US, respectively; B, score plots of PLS-DA for MALDI TOF mass spectra of LC, LP and US, respectively. CA: *Cynoglossus abbreviatus*, CG: *Cynoglossus gracilis*, CJ: *Cynoglossus joyneri*, CS: *Cynoglossus semilaevis*, LC: *Larimichthys crocea*, LP: *Larimichthys polyactis*, US: unknown sample [Color figure can be viewed at wileyonlinelibrary.com]

3.3 | Identification of unknown fish species

Authentication is highly useful to avoid adulteration. Gerbig et al used a miniature mass spectrometer for real-time milk authentication.²⁷ Sidwick et al used LC/MS to differentiate normally slaughtered and dead-on-arrival poultry meat.²⁸ To apply the presently developed method to fish authentication, we endeavored to build a PCA score plot pool as a reference with the mass spectral datasets of many individual fishes from one genus, and then determine the species of an unknown fish sample by comparing the mass spectra of the unknown sample with the datasets of known samples, as shown in Figure 1B. Figure 6 shows the PCA score plots for the investigation of an unknown sample (US). From Figure 6A, it can be seen that the US is rather different from all the four fishes from the genus *Cynoglossus*. From Figure 6B, it can be observed that the US is very close to LC. We duplicated the authentication procedure to the same test sample, and reached an identical conclusion. The results were in accordance with the fact that a new individual of LC fish was used to prepare the US, which was kept unknown to the examiner. Thus it is promising to identify fish using the developed strategy.

3.4 | Concerns about sample preparation and sensitivity

The optimum time for sample preparation and detection sensitivity was investigated. The findings shown in Figures S5 and S6 (supporting information) illustrate that it is very possible to shorten the extraction time of biosampling from the fish skin to 1 h, and thus the whole measurement, including soaking, desalting, MS analysis and multivariable statistical analysis, can be finished within 1–3 h for identification of ambiguous fishes. In addition, more experiments

were performed to survey how small amounts of sample can achieve the authentication of fishes. The results in Figure S7 (supporting information) indicate that fish skin samples in amounts of several μg can generate sufficient (*ca.* dozens of μL) biosample for MALDI TOF based fish identification. In practical routine analysis, to ensure the obtained sample is a representative sample of the population, hundreds of milligrams of sample (*ca.* 3–7 cm^2 of fish skin) from several individuals (≥ 6) is suggested.

4 | CONCLUSIONS

In the present study, an innovative method based on MALDI TOF MS has been developed and successfully applied to fish authentication at the species level. Ten different fish species from six genera were studied. The method was testified based on large-scale MALDI MS datasets and chemometric techniques. The results showed that all the studied fish could be clearly classified with the assistance of PCA and PLS-DA, including highly closely related species. Only trace amounts of fish skin sample and 1–3 h were required for analysis. Such a protocol is promisingly useful for rapid assessment of fraudulent practice of substitution of low-quality fish for commercially important species. It is technically possible to explore more strategies based on a similar concept to achieve applications in authentication of other food species.

4.1 | Live subject statement

The fish samples were taken and processed in accordance with the law and regulations of China and World Health Organization (WHO)'s guidelines for fish and seafood.

ACKNOWLEDGEMENTS

This work was supported by the Science and Technology Commission of Shanghai Municipality for Ability Construction of Regional Universities and Colleges Project (No. 18050502200), and the National Natural Science Foundation of China (NSFC 21804087). The authors are grateful to Miss Dandan Li and Mr. Shike Gao for their assistance during experiments.

ORCID

Hongyan Bi  <https://orcid.org/0000-0003-2162-1829>

REFERENCES

1. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T. Protein and polymer analyses up to m/z 100000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*. 1988;2(8):151-153.
2. Cain TC, Lubman DM, Weber WJ. Differentiation of bacteria using protein profiles from matrix-assisted laser-desorption ionization time-of-flight mass-spectrometry. *Rapid Commun Mass Spectrom*. 1994;8(12):1026-1030.
3. Krishnamurthy T, Ross PL, Rajamani U. Detection of pathogenic and non-pathogenic bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*. 1996;10(8):883-888.
4. Liang XL, Zheng KF, Qian MG, Lubman DM. Determination of bacterial protein profiles by matrix-assisted laser desorption/ionization mass spectrometry with high-performance liquid chromatography. *Rapid Commun Mass Spectrom*. 1996;10(10):1219-1226.
5. Giebel R, Worden C, Rust SM, Kleinheinz GT, Robbins M, Sandrin TR. Microbial fingerprinting using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS): Applications and challenges. *Adv Appl Microbiol*. 2010;71:149-184.
6. Bi H, Xi M, Zhang R, Wang C, Qiao L, Xie J. Electrostatic spray ionization-mass spectrometry for direct and fast wine characterization. *ACS Omega*. 2018;3(12):17881-17887.
7. M&A International Inc. The seafood industry: A sea of buyers fishing for M&A opportunities; 2013. Available: http://web.tmcapi.com/tmc/TMC_IMG/MAI/Reports/MAI_F&B_2013.pdf Retrieved March 11, 2019.
8. Black C, Chevallier OP, Haughey SA, et al. A real time metabolomic profiling approach to detecting fish fraud using rapid evaporative ionisation mass spectrometry. *Metabolomics*. 2017;13(12):153.
9. Teletchea F. Molecular identification methods of fish species: Reassessment and possible applications. *Rev Fish Biol Fish*. 2009;19(3):265-293.
10. Rasmussen RS, Morrissey MT. DNA-based methods for the identification of commercial fish and seafood species. *Compr Rev Food Sci Food Saf*. 2008;7(3):280-295.
11. Varadi L, Luo JL, Hibbs DE, et al. Methods for the detection and identification of pathogenic bacteria: Past, present, and future. *Chem Soc Rev*. 2017;46(16):4818-4832.
12. Ortea I, O'Connor G, Maquet A. Review on proteomics for food authentication. *J Proteomics*. 2016;147:212-225.
13. Nhek S, Mosleth EF, Hoy M, et al. Nonlinear visualisation and pixel-based alignment of 2D electrophoresis images. *Chemom Intel Lab Syst*. 2012;118:97-108.
14. Bossier P, Cooreman K. A databank able to be used for identifying and authenticating commercial flatfish (Pleuronectiformes) products at the species level using isoelectric focusing of native muscle proteins. *Int J Food Sci Technol*. 2000;35(6):563-568.
15. Ortea I, Canas B, Gallardo JM. Selected tandem mass spectrometry ion monitoring for the fast identification of seafood species. *J Chromatogr A*. 2011;1218(28):4445-4451.
16. Barik SK, Banerjee S, Bhattacharjee S, Das Gupta SK, Mohanty S, Mohanty BP. Proteomic analysis of sarcoplasmic peptides of two related fish species for food authentication. *Appl Biochem Biotechnol*. 2013;171(4):1011-1021.
17. Wulff T, Nielsen ME, Deelder AM, Jessen F, Palmblad M. Authentication of fish products by large-scale comparison of tandem mass spectra. *J Proteome Res*. 2013;12(11):5253-5259.
18. Nessen MA, van der Zwaan DJ, Grevers S, et al. Authentication of closely related fish and derived fish products using tandem mass spectrometry and spectral library matching. *J Agric Food Chem*. 2016;64(18):3669-3677.
19. Mazzeo MF, Siciliano RA. Proteomics for the authentication of fish species. *J Proteomics*. 2016;147:119-124.
20. Woo J-W, Yu S-J, Cho S-M, Lee Y-B, Kim S-B. Extraction optimization and properties of collagen from yellowfin tuna (*Thunnus albacares*) dorsal skin. *Food Hydrocoll*. 2008;22(5):879-887.
21. Gibb S, Strimmer K. MALDIquant: A versatile R package for the analysis of mass spectrometry data. *Bioinformatics*. 2012;28(17):2270-2271.
22. Walker JM. *The Protein Protocols Handbook*. 2nd ed. University of Hertfordshire Hatfield, UK: Humana Press; 2009.
23. Jolliffe IT, Cadima J. Principal component analysis: A review and recent developments. *Philos Trans R Soc A Math Phys Eng Sci*. 2016;374(2065):16.
24. Tzeng DY, Berns RS. A review of principal component analysis and its applications to color technology. *Color Res Appl*. 2005;30(2):84-98.
25. Olsen RL, Chappell RW, Loftis JC. Water quality sample collection, data treatment and results presentation for principal components analysis – literature review and Illinois River watershed case study. *Water Res*. 2012;46(9):3110-3122.
26. Szymanska E, Saccenti E, Smilde AK, Westerhuis JA. Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics*. 2012;8(1):S3-S16.
27. Gerbig S, Neese S, Penner A, Spengler B, Schulz S. Real-time food authentication using a miniature mass spectrometer. *Anal Chem*. 2017;89(20):10717-10725.
28. Sidwick KL, Johnson AE, Adam CD, Pereira L, Thompson DF. Use of liquid chromatography quadrupole time-of-flight mass spectrometry and metabolomic profiling to differentiate between normally slaughtered and dead on arrival poultry meat. *Anal Chem*. 2017;89(22):12131-12136.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Bi H, Zhong C, Shao M, et al. Differentiation and authentication of fishes at the species level through analysis of fish skin by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*. 2019;33:1336-1343. <https://doi.org/10.1002/rcm.8474>

Supplementary Materials

Differentiation and Authentication of Fishes at Species Level

Through Analysis of Fish Skin by MALDI TOF MS

Hongyan Bi^{1,*}, Chunyi Zhong¹, Mingke Shao¹, Chengyu Wang¹, Jia Yi¹, Liang Qiao², Junbo Zhang¹

¹, College of Food Science and Engineering, and College of Marine Sciences, Shanghai Ocean University, Hucheng Ring Road 999, Pudong New District, 201306 Shanghai, China

², Department of Chemistry, Fudan University, Songhu Road 2005, Yangpu District, 200438 Shanghai, China

^{*}, To whom correspondence should be addressed. Tel.: +86-21-61900364; Fax: +86-21-61900365; E-mail address: hybi@shou.edu.cn;

Supplementary data associated with the primary article, including figures and tables, are listed as follows, to illustrate more details and results during the experiment. In detail, the data include, 1) photos of large yellow croaker and small yellow croaker, as models to show the similarity of morphological features of fishes within the same genus; 2) table listed the size of large yellow croaker, small yellow croaker and other fishes used in the present study; 3) cross validation (CV) of the datasets collected from large yellow croaker and small yellow croaker; 4) table listed the scientific classification of fishes used in the present study; 5) mass spectra of bio-materials when soaking fish skin sample in alkali for 1 and 3 hours respectively, and 6) mass spectra of bio-materials from fish skin when dilution was made.



Figure S1. Photos of (A) large yellow croaker and (B) small yellow croaker, respectively.

Table S1. Sizes of large yellow croaker and small yellow croaker used for the experiments in the present study.

Species	Large Yellow Croaker				Small Yellow Croaker			
Scientific name	<i>Larimichthys Crocea</i>				<i>Larimichthys Polyactis</i>			
	Sample code	Length (L) /mm	Weight (W) /g	Ratio of W to L / (g/mm)	Sample code	Length/mm	Weight/g	Ratio of W to L / (g/mm)
	LC1	229	225.18	0.98	LP1	161	54.34	0.34
	LC2	210	212.09	1.01	LP2	148	44.93	0.30
	LC3	223	230.68	1.03	LP3	142	42.34	0.30
	LC4	216	200.86	0.93	LP4	148	34.22	0.23
	LC5	223	236.17	1.06	LP5	149	43.46	0.29
	LC6	224	237.92	1.06	LP6	166	61.42	0.37
Average	-	220.8±6.7	223.8±14.6	1.01±0.05	-	152.3±9.1	46.8±9.6	0.30±0.05

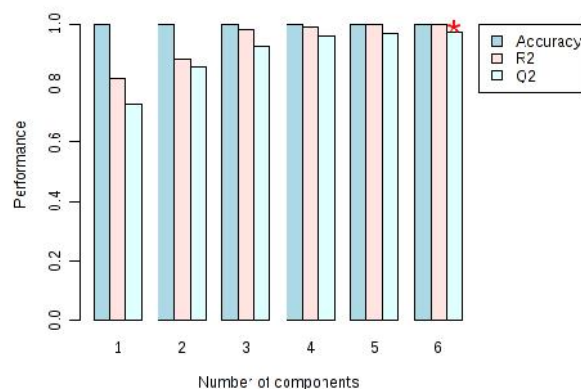


Figure S2. Cross validation (CV) of the datasets collected from large yellow croaker and small yellow croaker. During modeling, 10-fold cross validation was used to choose the number of components where 6 components were set. Q2, calculated via CV, is an estimate of the predictive ability of the model. In each CV, the predicted data are compared with the original data, and the sum of squared errors is calculated. The prediction error is then summed over all samples, which is Predicted Residual Sum of Squares or short as PRESS. For convenience, the PRESS is divided by the initial sum of squares and subtracted from 1 to resemble the scale of the R2.

Table S2. Scientific classification of the fishes used in the present study. All these fishes belong to kingdom of *Animalia*, Phylum of *Chordata*, and Class of *Actinopterygii*. The classification of fish refers to <http://www.marinespecies.org> and fishbase.de, and Wikipedia.

Abbreviation in the present study	Scientific name	Common name	Order	Family	Genus	Species
CA	<i>Cynoglossus Abbreviatus</i>	Three-lined tongue sole	<i>Pleuronectiformes</i>	<i>Cynoglossidae</i>	<i>Cynoglossus</i>	<i>C. abbreviates</i>
CG	<i>Cynoglossus Gracilis</i>	~	<i>Pleuronectiformes</i>	<i>Cynoglossidae</i>	<i>Cynoglossus</i>	<i>C. gracilis</i>
CJ	<i>Cynoglossus Joyneri</i>	~	<i>Pleuronectiformes</i>	<i>Cynoglossidae</i>	<i>Cynoglossus</i>	<i>C. joyneri</i>
CS	<i>Cynoglossus Semilaevis</i>	~	<i>Pleuronectiformes</i>	<i>Cynoglossidae</i>	<i>Cynoglossus</i>	<i>C. semilaevis</i>
CL	<i>Collichthys Lucidus</i>	~	<i>Perciformes</i>	<i>Sciaenidae</i>	<i>Collichthys</i>	<i>C. lucidus</i>
LC	<i>Larimichthys Crocea</i>	Large Yellow Croaker	<i>Perciformes</i>	<i>Sciaenidae</i>	<i>Larimichthys</i>	<i>L. crocea</i>
LP	<i>Larimichthys Polyactis</i>	Small Yellow Croaker	<i>Perciformes</i>	<i>Sciaenidae</i>	<i>Larimichthys</i>	<i>L. polyactis</i>
PA	<i>Pampus Argenteus</i>	Silver pomfret; White pomfret	<i>Scombriformes</i>	<i>Stromateidae</i>	<i>Pampus</i>	<i>P. argenteus</i>
TO	<i>Trachinotus Ovatus</i>	Pompano	<i>Perciformes</i>	<i>Carangidae</i>	<i>Trachinotus</i>	<i>T. ovatus</i>
ZF	<i>Zeus Faber</i>	Zeus faber Linnaeus; John Dory; Target perch	<i>Zeiformes</i>	<i>Zeidae</i>	<i>Zeus</i>	<i>Z. faber</i>

Table S3. Sizes of fishes used in the present study.

Scientific name	Length (cm)
<i>Cynoglossus Joyneri</i>	23.1 ± 2.1 (n = 6)
<i>Cynoglossus Semilaevis</i>	18.4 ± 1.2 (n = 6)
<i>Cynoglossus Abbriviates</i>	20.8 ± 1.1 (n = 3)
<i>Cynoglossus Glacilis</i>	19.0 ± 0.4 (n = 3)
<i>Collichthys Lucidus</i>	10.8 ± 0.3 (n = 3)
<i>Pampus Argenteus</i>	29.8 ± 1.2 (n = 6)
<i>Trachinotus Ovatus</i>	23.8 ± 0.9 (n = 6)
<i>Zeus Faber</i>	22.2 ± 1.8 (n = 6)

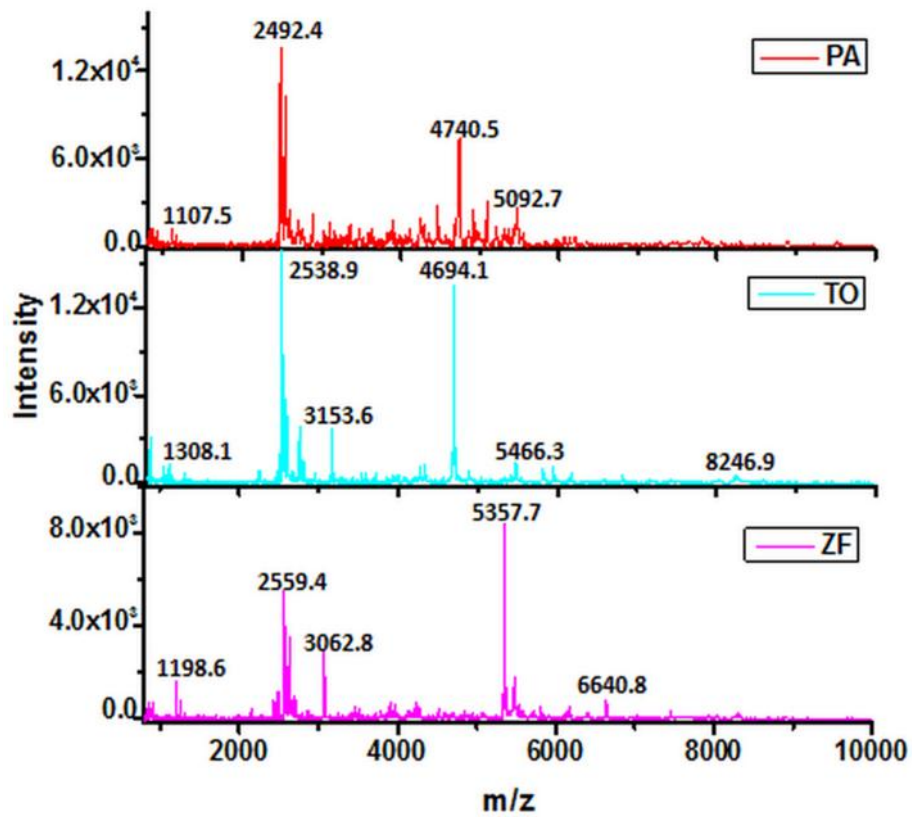


Figure S3. MALDI-TOF mass spectra of bio-materials extracted from the skin of *Pampus argenteus* (PA), *Trachinotus ovatus* (TO), and *Zeus faber* (ZF). The skin samples of these fishes were pretreated in NaOH solution for 3 hours.

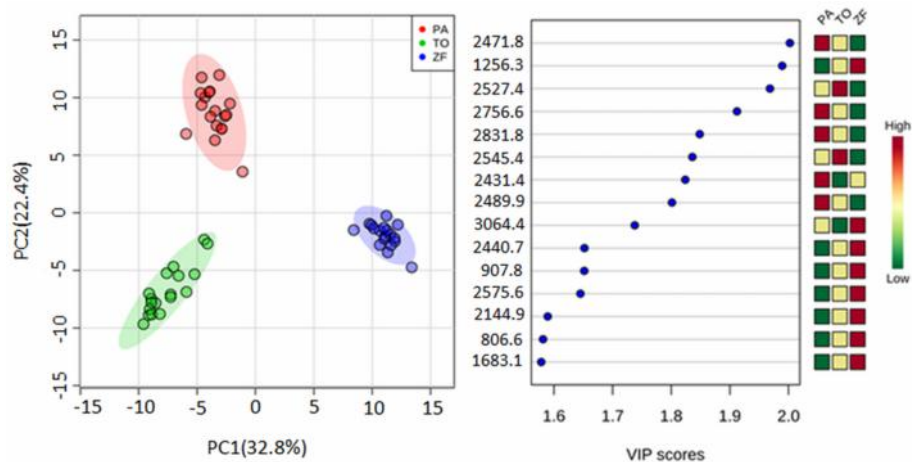


Figure S4. (A) PCA for the MALDI-TOF mass spectra of *Pampus argenteus* (PA), *Trachinotus ovatus* (TO), and *Zeus faber* (ZF). (B) VIP score from PLS-DA.

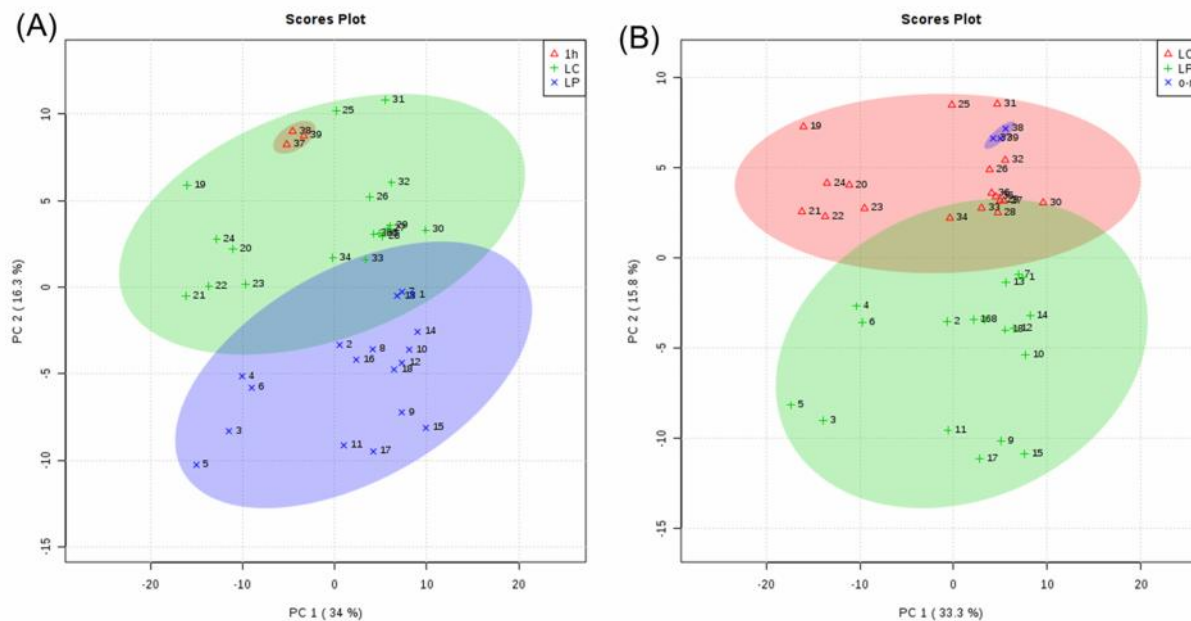


Figure S5. PCA score plot for bio-samples of large yellow croaker pretreated (A) for 1 h, and (B) overnight. The skin samples from 6 large yellow croakers were mixed to form a sample for performing the soaking experiment. The skin sample was soaked in alkaline solution for 1 hour and overnight respectively which results are labeled as 1h and o-n in the inlet. LC and LP

represent the PCA score plots of samples of fish of large yellow croaker and small yellow croaker respectively.

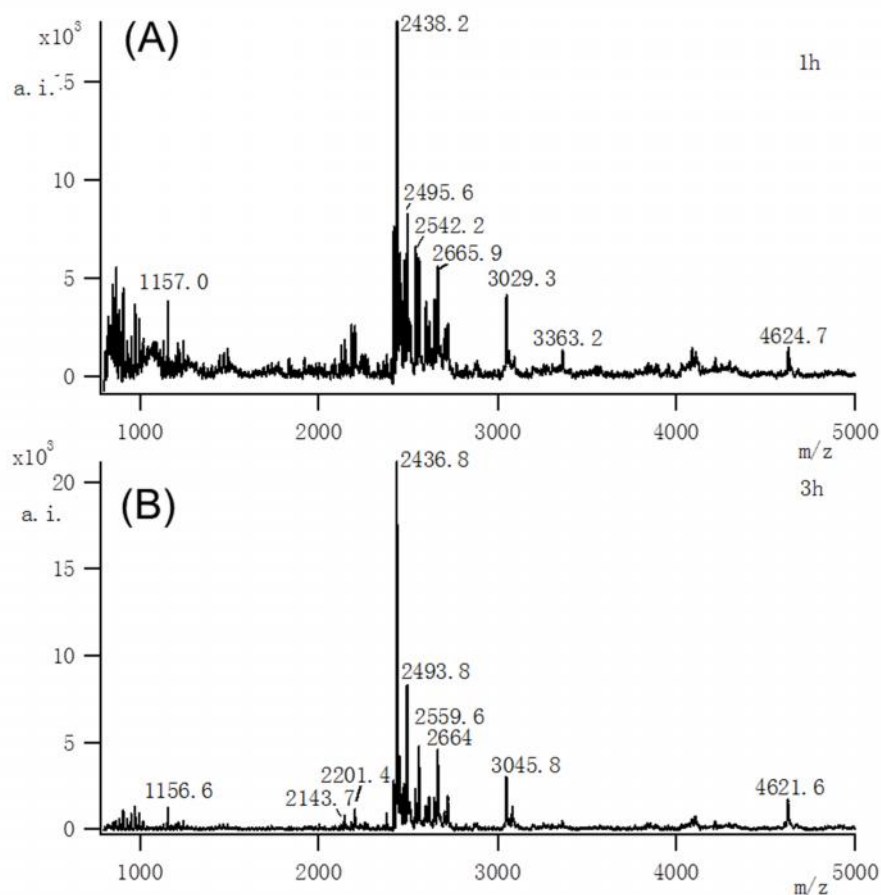


Figure S6. Mass spectra of *Larimichthys Crocea* pretreated for (A) 1 and (B) 3 hours. The skin samples from 6 large yellow croakers were mixed to form a sample for performing the soaking experiment. The skin sample was soaked in alkaline solution for 1 and 3 hours, respectively.

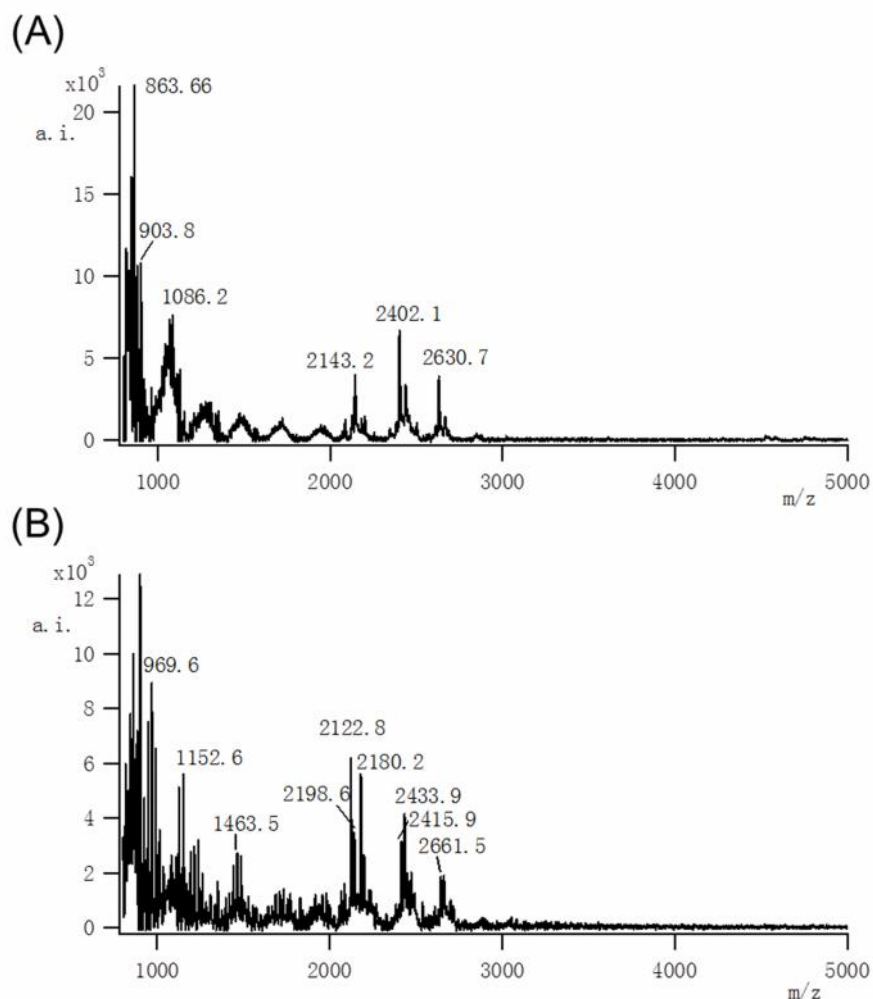


Figure S7. Mass spectra of bio-samples from the fish skin at low concentration. (A) The desalted bio-samples extracted from 6 large yellow croakers were mixed to form a parent solution (v/v=1:1), and then diluted by 100 000 times to obtain the sample for the MS detection. (B) 6 large yellow croaker's skin samples were mixed to form a sample. 0.01 g of the resultant sample was dipped in 10 mL of NaOH solution overnight to extract the bio-samples, and then diluted 10 times by NaOH solution (v/v). The obtained sample was taken for centrifuge column cleaning-up after a neutralization step prior to MS analysis.

