Fish species authentication in commercial fish products using mass spectrometry and spectral library matching approach

Varunjikar, M.S.; Pineda-Pampliega, J.; Belghit, I.; Palmblad, M.; Grosvik, B.E.; Meier, S.; ... ; Rasinger, J.D.

Food Research International 192 (2024) 114785

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Fish species authentication in commercial fish products using mass spectrometry and spectral library matching approach

Madhushri S. Varunjikar [a,1,*], Javier Pineda-Pampliega [a,1,*], Ikram Belghit [a], Magnus Palmblad [b], Bjørn Einar Grøsvik [a], Sonnich Meier [a], Pål Asgeir Olsvik [c], Kai K. Lie [a], Josef D. Rasinger [a]

[a] Institute of Marine Research, P.O. Box 1870 Nordnes, 5817 Bergen, Norway
[b] Leiden University Medical Center, 2300 RC Leiden, the Netherlands
[c] Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway

ARTICLE INFO

Keywords:
Authentication
Seafood
Proteomics
Tandem mass spectra
Spectral libraries

ABSTRACT

Seafood fraud has become a global issue, threatening food security and safety. Adulteration, substitution, dilution, and incorrect labeling of seafood products are fraudulent practices that violate consumer safety. In this context, developing sensitive, robust, and high-throughput molecular tools for food and feed authentication is becoming crucial for regulatory purposes. Analytical approaches such as proteomics mass spectrometry have shown promise in detecting incorrectly labeled products. For the application of these tools, genome information is crucial, but currently, for many marine species of commercial importance, such information is unavailable. However, when combining proteomic analysis with spectral library matching, commercially important fish species were successfully identified, differentiated, and quantified in pure muscle samples and mixtures, even when genome information was scarce. This study further tested the previously developed spectral library matching approach to differentiate between 29 fish species from the North Sea and examined samples including individual fish, laboratory-prepared mixtures and commercial products. For authenticating libraries generated from 29 fish species, fresh muscle samples from the fish samples were matched against the reference spectral libraries. Species of the fresh fish samples were correctly authenticated using the spectral library approach. The same result was obtained when evaluating the laboratory-prepared mixtures. Furthermore, processed commercial products containing mixtures of two or three fish species were matched against these reference spectral libraries to test the accuracy and robustness of this method for authentication of fish species. The results indicated that the method is suitable for the authentication of fish species from highly processed samples such as fish cakes and burgers. The study shows that current and future challenges in food and feed authentication can efficiently be tackled by reference spectral libraries method when prospecting new resources in the Arctic.

1. Introduction

Seafood mislabeling is a type of food fraud where fish products are mislabeled or substituted with another species (Bouzembrak et al., 2018; Chien et al., 2022). In studies from more than 30 countries, rates of mislabeling seafood were found to be around 36 %, indicating a vast scale of global fraud (Leahy, 2021). According to European regulations (European Commission 1169/2011), consumers should be properly informed about the origin of seafood and whether it is produced by aquaculture or caught in the wild. Mislabeling of seafood species can cause severe consequences for sensitive consumers (e.g., by the presence of allergens), posing a health hazard (Visciano & Schirone, 2021). Reliable analytical methods must be in place to authenticate seafood products, ensure the integrity of the global food chain, and protect

Abbreviations: AGC, Automatic Gain Control; MS, Mass Spectrometry; HCD, Higher-Energy Collision Dissociation; DDA, Data Dependent Acquisition; HPLC-MS/MS, High-Performance Liquid Chromatography coupled to tandem Mass Spectrometry; IT, maximum Injection Time; NCE, Normalized Collision energy; TPP, Trans-Proteomic Pipeline; MGF, Mascot Generic Format.

* Corresponding authors.
E-mail addresses: madhushri.shrikant.varunjikar@hi.no (M.S. Varunjikar), javier.pineda.pampliega@hi.no (J. Pineda-Pampliega), ikram.belghit@hi.no (I. Belghit), n.m.palmblad@lumc.nl (M. Palmblad), bjorn.grosvik@hi.no (B. Einar Grøsvik), sonnich@hi.no (S. Meier), pal.a.olsvik@nord.no (P. Asgeir Olsvik), kaikristoffer.lie@hi.no (K.K. Lie), josef.rasinger@hi.no (J.D. Rasinger).

[1] Equal contributions.

https://doi.org/10.1016/j.foodres.2024.114785
Received 3 May 2024; Received in revised form 9 July 2024; Accepted 15 July 2024
Available online 16 July 2024

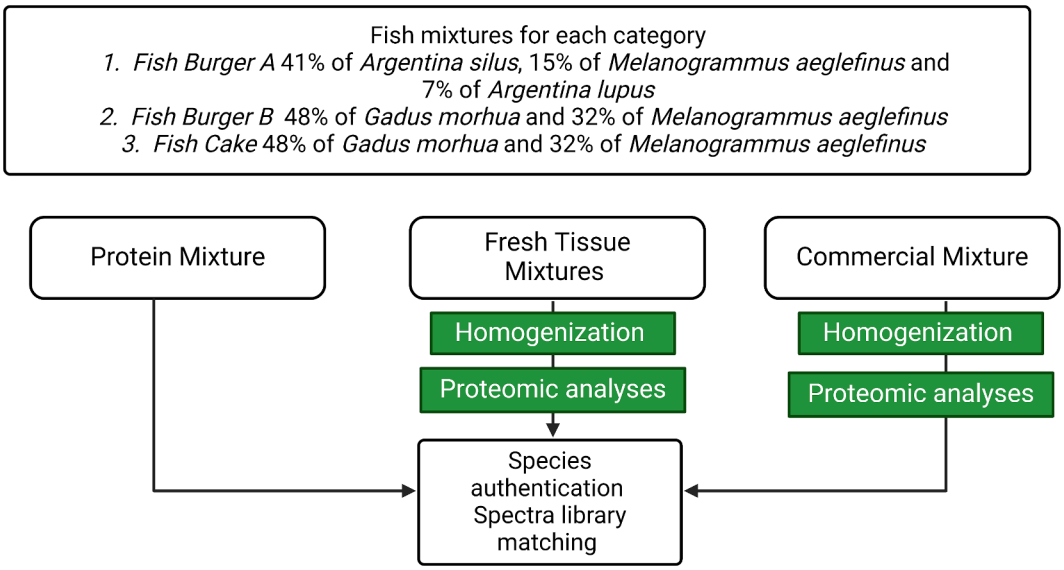

Fig. 1. Experimental Design for fish mixtures. 1. Fish Burgers A.silus41/M.aeglefinus15/A.lupus7 where *A. silus* was 41%, *M. aeglefinus* was 15%, and *A. lupus* was 7% of total weight; 2. Fish cake G. morhua48/M. eglefinus32 where G. morhua was 48% and *M. aeglefinus* was 32% of total weight 3. Fish Burgers G. morhua48/M. eglefinus32 where G. morhua was 48% and *M. aeglefinus* was 32% of total weight. Commercial mixtures were store bought samples with other food materials such as flour, oil, spices, vegetables, and starch; fresh tissue mixtures were prepared in laboratory with frozen fish fillet using same percentage as commercial mixtures; protein mixtures were prepared in laboratory using protein extracts from fish fillet of respective species using same percentage as commercial mixtures.

consumer safety.

Fraudulent practices include the replacement of high-value fish species with similar fish species with low commercial values. The motivation behind such fraud is to gain economic benefit, which severely affects consumer trust (Deconinck et al., 2020). Commercially most relevant fish species susceptible to fraud include tuna (several species), Atlantic cod (*Gadus morhua*), and sole (*Solea solea*), which are replaced or adulterated with alternative low-value species (Chien et al., 2022; Klapper et al., 2023). For example, the sole was replaced by other white fish species, such as pangasius (*Pangasianodon hypophthalmus*) (Deconinck et al., 2020). Also, cod is highly susceptible to fraud due to global trade and increasing demands worldwide (Chien et al., 2022). For instance, Atlantic cod (*G. morhua*) was replaced with shark species in catering, indicating fraudulent commercial market practices (Pardo et al., 2018). Fish products in the food market include a wide variety of products such as fish balls, fish burgers, and canned fish products, where the visual identification of the species is much more complicated, or even impossible, widening the possibilities of fraud.

Along with consumer safety, the mislabeling of seafood species threatens environmental protection policies, as there have been incidents in which illegally fished species entered food markets. For example, endangered shark species were sold in the food market and added to pet food, indicating that product mislabeling can also influence red-listed species (French & Wainwright, 2022; Zuccolo et al., 2023). As these species are at risk of global extinction, monitoring the food and feed chain is crucial to have greater control over protection policies (French & Wainwright, 2022). Therefore, from the environmental conservation perspective, it is crucial to authenticate fish species entering food and feed chains.

Molecular and chemical methods can benefit the authentication of highly processed fish species from complex mixtures (Nessen et al., 2016). Recently, DNA-based methods such as DNA-barcoding have increasingly been used for fish species authentication and identification (Chien et al., 2022; Shokralla et al., 2015). Furthermore, DNA metabarcoding combined with next-generation sequencing (NGS) were crucial methods for the authentication of food species and complex mixtures in recently discovered frauds (French & Wainwright, 2022; Klapper et al., 2023). For the quantification of fish mixtures, next-generation sequencing was found to be effective in separating closely related species (Varunjikar et al., 2022). Even though DNA-based methods are highly robust and accurate, processing procedures during food preparation might breakdown DNA affecting detection rates (Carrera et al., 2013a). Proteomics methods offer an alternative to DNA for authenticating fish species in both pure and mixed seafood samples (Carrera et al., 2013a; Chien et al., 2022; Hu et al., 2022; Nessen et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013).

In bottom-up proteomics, both targeted and untargeted approaches can be considered for the authentication of species origin of food samples (Wulff et al., 2013). Targeted proteomic methods are based on species-specific peptide markers, which help in the detection and quantification of species in given samples (Carrera et al., 2013a; Hu et al., 2022). Recent studies show that these methods were effective in identifying adulteration of Atlantic salmon (*Salmo salar*) with rainbow trout (*Oncorhynchus mykiss*) (Gu et al., 2020). Similarly, for differentiation of fish species from the Merlucciidae family, protein factions of parvalbumin were used (Carrera et al., 2013b). While targeted proteomic methods are effective in species identification and authentication, species-specific peptide markers must be identified using reference proteome information from targeted fish species. Such reference proteomes are seldomly available for marine fish species of commercial interest (Varunjikar, 2023); therefore, to overcome their limited availability, the implementation of mass spectrometry spectral library matching is deemed beneficial (Nessen et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013). In addition, the spectral library matching utilizes information collected from the proteome of species, instead of focusing on a limited number of target peptides. This approach was previously used for other processed food products such as meat and sausages (Ohana et al., 2016). In addition, we previously have tested spectral library matching on mixed unprocessed samples containing a mixture of three different fish species and shown to be effective (Varunjikar et al., 2022).

This study aims to demonstrate the effectiveness of the spectral library matching method in detecting fish species in feed and food. To evaluate the strengths and weakness of this approach, we texted pure fish tissue, binary mixtures of closely related species, and complex mixtures of commercial food products containing extra components in addition to fish and undergoing several treatments. All these samples were searched against 29 reference library spectra created for this study,

creating a reference spectral library collection that could be used in future studies.

2. Material and methods

2.1. Sampling

In this study muscle tissue samples from 29 different species were collected. Names of these species are provided in Supplementary Table 1 along with common names. These samples were collected by the Institute of Marine Research, from the North Sea in 2010–2013 for research purposes. All applicable international, national, and/or institutional guidelines for the care and ethical use of animals were followed. All the species were assigned by a trained ichthyologist for validation purposes. Prior to protein extraction, fish were frozen and stored at −80 °C. Fish fillet samples were used to build the proteomic spectral library. For authentication of each reference spectral library, biological replicates from the same species were used (i.e., tissue samples of the same individual but not the same as the used for creating the reference spectral library, called *quality-control samples*).

For commercial product samples containing fish mixtures, one type of fish cake and two types of fish burgers were purchased from Norwegian supermarkets. Selected commercial fish cake and fish burgers contained a mixture of two or three different fish species (a detailed overview of the composition is shown in Supplementary Table 2). Prior to protein extraction, commercial fish mixtures were frozen and stored at −80 °C.

2.2. Experimental design

Binary mixtures were prepared in the laboratory for *G. morhua* and *Melanogrammus aeglefinus* using fresh fish muscle tissues with the ratios of 100 %, 90 %, 67 %, 50 %, 33 %, 10 %, 1 %, and 0 % of *G. morhua.* All these samples were processed in triplicates to assess the effects of sample preparation. Species were selected based on being closely related commercial fish species of different value.

In addition to different fish species, commercial samples contain more ingredients and suffer different processing conditions that could affect proteomic results. Two similar mixtures for each commercial sample were prepared in the laboratory to check these potential differences, one using frozen tissue from the same fish species and another using protein extracts.

The mixtures were commercially available samples and laboratory-made mixtures from fresh fish muscles and protein extracts, as shown in Fig. 1. One fish burger containing 41 % of *Argentina silus*, 15 % of *M. aeglefinus* and 7 % of *Argentina lupus* (to simplify, from no on Burger A*)*. The second fish burger (from now on *Burger B*) and the fish cake (from now on *Fish Cake*) contains 48 % *G. morhua* and 32 % *M. aeglefinus*). All samples were prepared as triplicates. A description of mixtures is given in the Supplementary Table 2.

2.3. Protein and peptide extraction

Protein extraction was performed on all fish muscle tissues, commercial mixtures, and laboratory-prepared samples. 100 mg of samples were weighted into the PlusOne Sample Grinding kit (GE Healthcare Life Science) and 1 mL of lysis buffer was added (4 % SBS, 0.1 M Tris-HCl, pH 7.6) as described before by Varunjikar et al., (2022). Samples were kept on ice and homogenized, later these samples were centrifuged for 10 min at 15,000 g to remove resin and other debris. Supernatants from the samples were collected and heated at 95 °C for 5 min; protein concentration of extracts was determined using the Pierce 660 assay (ThermoFisher Scientific), and samples were stored at −20 °C until further processing.

Peptide extractions were performed using Filter Aided Sample Preparation (FASP) as described in Varunjikar et al., (2022). Protein extracts (with 100 µg of protein) were diluted using 200 µL of urea solution (8 M in Tris-HCl 100 mM pH 8.5) and transferred to spin columns (Microcon 30, Millipore). Samples were alkylated with 50 mL of iodoacetamide and incubated in the dark at room temperature. Following the incubation step samples were washed with 200 µL of urea solution (8 M in Tris-HCl 100, mM pH 8.5) and 100 µL of ammonium bicarbonate (50 mM). For the digestion of peptides, trypsin was added to the filter in 1:50 enzyme-to-protein ratio and incubated for 16 h at 37 °C. Subsequently, incubation filters were washed with 40 µL of ammonium bicarbonate (50 mM) and 40 µL of sodium chloride (0.5 M). After the final centrifugation step, the peptide concentration in the solution was determined using a Nanodrop (Thermo Fisher Scientific).

After peptide extraction, peptides were further cleaned using Pierce™ C18 spin columns (Thermo Fisher Scientific). Columns were first equilibrated with methanol and water (50/50, vol/vol) and then with wash solution containing acetonitrile, trifluoroacetic acid, and water (5/0.5/94.5, vol/vol/vol). Peptide samples were mixed with acetonitrile, trifluoroacetic acid, and water (20/2/78, vol/vol/vol) and loaded on the columns. The peptides were eluted from Pierce™ C18 spin column using acetonitrile and water (30/70, vol/vol). Samples were evaporated in a speed vacuum dryer (LABCONCO CentriVap micro IR) to obtain peptide pellets which were dissolved in acetonitrile, formic acid, and water (2/0.1/97.9, vol/vol/vol). Samples were stored at –20 °C until mass spectrometric analyses.

2.4. HPLC-MS/MS

2.4.1. NanoLC Exploris 480 mass spectrometry

For all samples about 0.5 µg of tryptic peptides dissolved in 2 % acetonitrile, 0.5 % formic acid, were injected into an Ultimate 3000 RSLC system (Thermo Fisher Scientific) connected online to a Exploris 480 mass spectrometer (Thermo Fisher Scientific) equipped with EASY-spray nano-electrospray ion source (Thermo Fisher Scientific).

2.4.2. Trapping and desalting

Samples were loaded and desalted on a pre-column (Thermo Fisher Scientific Acclaim PepMap 100, 2 cm x 75 µm ID nanoViper column, packed with 3 µm C18 beads) at a flow rate of 5 µl/min for 5 min with 0.1 % trifluoroacetic acid.

2.4.3. HPLC analyses

Peptides were separated during a biphasic acetonitrile gradient from two nanoflow UPLC pumps (flow rate of 250 nl/min) on a 25 cm analytical column (PepMap RSLC, 25 cm x 75 µm ID. EASY-spray column, packed with 2 µm C18 beads). Solvent A and B were 0.1 % formic acid (vol/vol) in water and 100 % ACN respectively. The gradient composition was 5 % B during trapping (5 min) followed by 5–8 % B over 1 min, 8–26 % B for the next 39 min, 26–35 % B over 8 min, and 35–90 % B over 2 min. Elution of very hydrophobic peptides and conditioning of the column were performed for 5 min isocratic elution with 90 %B and 10 min isocratic conditioning with 5 % B respectively. Instrument control was through Thermo Fisher Scientific SII for Xcalibur 1.6.

2.4.4. Orbitrap Exploris 480 data dependent acquisition (DDA)

Peptides eluting from the LC-column were ionized in electrospray and analyzed by Orbitrap Exploris 480. The mass spectrometer was operated in DDA-mode (data-dependent acquisition) to automatically switch between full scan MS and MS/MS acquisition with a 2.4 s cycle time. Instrument control was through Orbitrap Exploris 480 Tune 3.1 and Xcalibur 4.4.

MS spectra were acquired in the scan range 375–1500 m/z with RF lens at 40 %, resolution R = 120 000 at m/z 200, automatic gain control (AGC) target of 3e6 and a maximum injection time (IT) at “Auto”. The most intense eluting peptides above an intensity threshold of 50,000 counts, and charge states 2 to 6, were sequentially isolated to a target

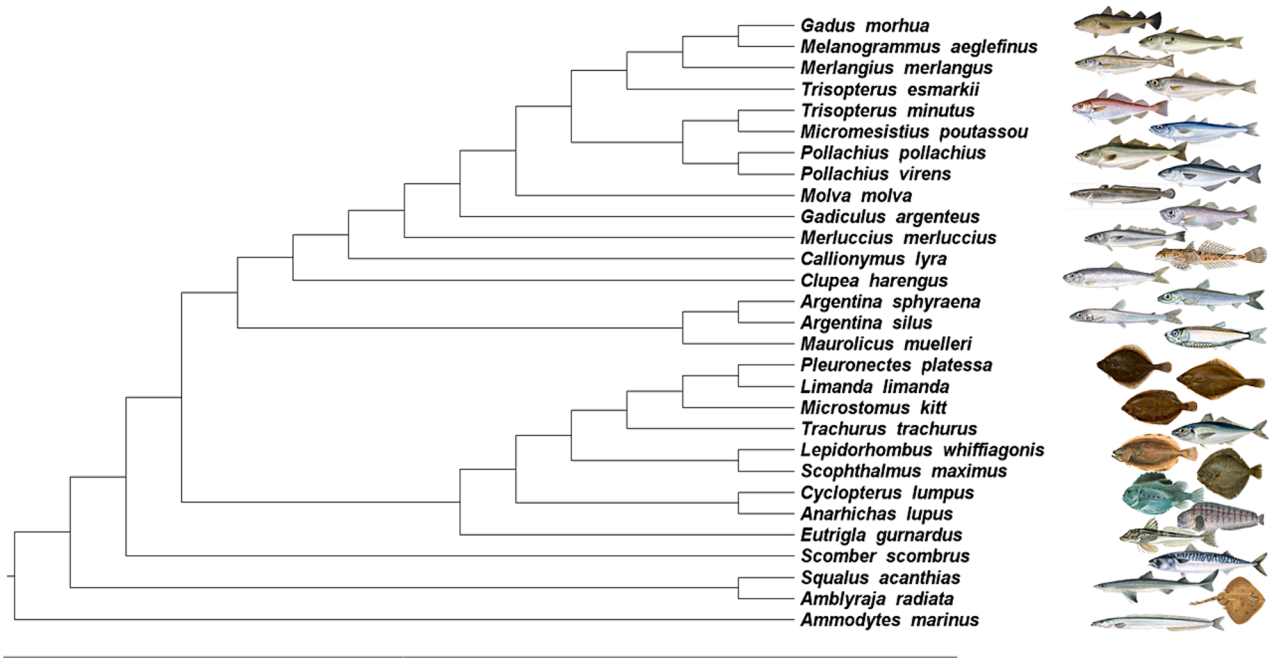

Fig. 2. Phylogenetic tree built from compareMS2 with top 500 spectra. Fish species differentiation using compareMS2 Version 2.0.0-alpha for direct comparison of HPLC-MS/MS tandem mass spectrometric data. Molecular phylogenetic tree indicate that fish species were separated as per the taxonomic classification. 29 fish species from the North sea were differentiated as per the molecular phylogeny using only tandem mass spectrometry-based proteomics data.

value (AGC) of 1e5 and a maximum IT of 70 ms in the C-trap, and isolation width maintained at 1.6 *m/z*, before fragmentation in the HCD (Higher-Energy Collision Dissociation) cell. Fragmentation was performed with a normalized collision energy (NCE) of 30 %, and fragments were detected in the Orbitrap at a resolution of 15 000 at *m/z* 200, with the first mass fixed at *m/z* 110. One MS/MS spectrum of a precursor mass was allowed before dynamic exclusion for 30 s with "exclude isotopes" on. Lock-mass internal calibration was not used.

2.4.5. Ionsource parameter

The spray and ion-source parameters were as follows. Ion spray voltage = 1800 V, no sheath and auxiliary gas flow, and capillary temperature = 275 °C.

Settings for both the HPLC and the mass spectrometer, along with the conditions in which samples are treated, are detailed the Supplementary Material 1.

2.5. Proteomics Bioinformatics

The Thermo.raw files obtained from mass spectrometry were converted to mzML format using msConvert (version: 3.0.23233., ProteoWizard (Kessner et al., 2008)). compareMS2 (version: 2.0 (Marissen et al., 2023)) was used to perform molecular phylogenetic analyses and quality control using proteomics data (top- 500 spectra). The phylogenetic tree created with compareMS2 was compared against a phylogenetic tree generated using the NCBI Taxonomy Browser.

For peptide identification, tandem mass spectra were in mzML format were matched against Zebra fish (*Danio rerio*) reference proteome (UP000000437) accessed on November 2022 using Comet search (Eng et al., 2013) in the Trans-Proteomics Pipeline (TPP) (version 5.2.0 (Deutsch et al., 2015)). For the search, precursor mass tolerance was set to 20 ppm, trypsin was selected as a digestive enzyme allowing two non-enzymatic termini and carbamidomethylating of carbon and oxidation of methionine were set as a fixed and variable modification. Generated pepXML files were used for creating spectral libraries. A total of 29 spectral libraries were created, each corresponding to a different fish species, utilizing a single sample from each species with SpectraST (version 5.0 (Lam, 2011)) in Windows PowerShell. This process generates the corresponding.splib files for each library.

Quality control of each reference spectral library was performed by cross matching the spectra of 24 species-specific samples, being all included in the list of 29 species, but not using the same sample as the one previously used for creating the spectral libraries, against all 29 spectral libraries created. This process was carried out using SpectraST implemented in the Trans-Proteomic Pipeline (TPP) (version 6.3.3 (Deutsch et al., 2023)). Matched spectra have dot product values ranging from zero to one: a dot product of zero indicates mismatched spectra, while a dot product of one indicates identical spectra (Varunjikar et al., 2022). R script "1_Extract_matching_spectra " was developed to extract all matching spectra with dot products above 0.8, which were considered valid matches. These matches were then combined into a single.xlsx file using R script "2_Combine_matching_spectra" (both scripts are available at: https://github.com/imr-marinetox/Fish_species_authentication_spectra_library_matching). The number of matching spectra per sample in each library and the percentage of matching were calculated also using R (version 4.3.2). Processed mass spectrometry data and spectral libraries were deposited in an online repository massive.ucsd.edu and ProteomeXchange (MassIVE id: MSV000094602 and ProteomeXchange id: PXD051671).

After the quality control, all the query samples (25 corresponding to different concentrations of *G. morhua* and *M. aeglefinus*, and 25 corresponding to commercial samples, as described in section 2.2) were analyzed in the same manner as the samples used for the quality control. This involved searching against each reference library and using R to extract and calculate the data regarding the number and percentage of matching spectra.

3. Results and discussion

The need for tools to authenticate species in food and feed material is increasing, driven by a growing demand for quality products and possibilities of potential fraud. In the global seafood market, the substitution of similar species due to mislabeling or intentional modification could have serious economic and health implications (Bi et al., 2019).

Table 1
Validation of the spectral libraries using specie-specific samples different that the used for creating the libraries. First, second and third best matching Spectral Library for each sample. * Indicates query samples without best matching with their corresponding reference libraries.

Query sample	Matching spectra (best match)	% matching spectra (best match)	Reference library	Matching spectra (2nd best match)	% matching spectra (2nd best match)	Reference library	Matching spectra (3rd best match)	% matching spectra (3rd best match)	Reference library
Amblyraja radiata	10,474	49,43	A. radiata	3319	15,67	S. acanthias	1770	8,36	A. silus
Ammodytes marinus*	2668	15,03	E. gurnardus	2620	14,79	T. trachurus	2562	14,44	L. whiffiagonis
Anarhichas lupus	7869	44,81	A. lupus	3584	20,43	E. gurnardus	3492	19,90	C. lumpus
Argentina silus	10,763	50,64	A. silus	9854	46,32	A. sphyraena	3325	15,66	M. molva
Argentina sphyraena	7758	44,48	A. sphyraena	6418	36,93	A. silus	2430	13,99	M. muelleri
Callionymus lyra	7999	46,33	C. lyra	3081	17,89	T. trachurus	3048	17,67	L. whiffiagonis
Clupea harengus	9421	44,40	C. harengus	2720	12,84	A. sphyraena	2557	12,12	M. aeglefinus
Cyclopterus lumpus	9287	42,70	C. lumpus	4183	19,27	A. lupus	3789	17,45	E. gurnardus
Eutrigla gurnardus	37,325	99,40	E. gurnardus	8115	21,70	A. lupus	7296	19,51	T. trachurus
Gadiculus argenteus	7041	33,00	G. argenteus	5873	27,34	T. minutus	5859	27,30	M. poutassou
Gadus morhua	11,438	54,88	G. morhua	8159	39,14	M. merlangus	8131	39,03	P.virens
Lepidorhombus Whiffiagonis	9163	42,99	L. whiffiagonis	5394	25,32	S. maximus	3733	17,54	T. trachurus
Limanda limanda	10,558	48,10	L. limanda	6657	30,35	P. platessa	5321	24,28	M. kitt
Maurolicus muelleri	6108	27,86	M. muelleri	1924	8,78	A. sphyraena	1826	8,34	A. silus
Merlangius merlangus	36,775	99,43	M. merlangus	15,351	41,61	P. virens	14,449	39,23	M. poutassou
Melanogrammus aeglefinus	10,162	50,14	M. aeglefinus	7827	38,62	M. merlangus	7494	37,00	P. virens
Merluccius merluccius	5734	29,37	M. merluccius	3432	17,58	M. aeglefinus	3417	17,49	M. molva
Micromesistius poutassou	8286	38,56	M. poutassou	6134	28,53	T. minutus	5904	27,54	G. morhua
Microstomus kitt	9858	45,62	M. kitt	6072	28,12	P. platessa	5394	24,97	L. limanda
Molva molva	7695	36,05	M. molva	4631	21,80	M. aeglefinus	4566	21,53	G.morhua
Pleuronectes platessa*	3041	13,60	A. marinus	1165	5,21	L. limanda	1075	4,81	P. platessa
Pollachius pollachius	6732	31,44	P. pollachius	6555	30,64	P. virens	6073	28,41	G. morhua
Pollachius virens	7167	32,97	P. virens	5938	27,35	G. morhua	5915	27,17	M. aeglefinus
Scomber scombrus	8689	41,08	S. scombrus	3640	17,27	T. trachurus	3292	15,61	E. gurnardus
Scophthalmus maximus	37,190	99,43	S. maximus	12,371	33,17	L. whiffiagonis	7847	21,07	P. platessa
Squalus acanthias	10,611	49,66	S. acanthias	3323	15,59	A. radiata	1696	7,95	E. gurnardus
Trachurus trachurus	7157	32,52	T. trachurus	3106	14,11	E. gurnardus	3013	13,69	L. whiffiagonis
Trisopterus esmarkii*	7694	36,29	T. minutus	7502	35,90	T. esmarkii	6711	31,70	M. poutassou
Trisopterus minutus	9093	41,89	T. minutus	6769	31,57	T. esmarkii	6467	29,83	M. poutassou

Table 2
First, second and third best matching Spectral Library for the samples G. morhua and M. aeglefinus.

Query sample	Matching spectra (best match)	% matching spectra (best match)	Reference library	Matching spectra (2nd best match)	% matching spectra (2nd best match)	Reference library	Matching spectra (3rd best match)	% matching spectra (3rd best match)	Reference library
G.morhua and M. eglefinus (100—0)	10735 ± 614	52,71 ± 1,88	G. morhua	7958 ± 224	39,08 ± 0,52	M. merlangus	7893 ± 230	38,8 ± 0,36	P. virens
G.morhua and M. eglefinus (90—10)	11026 ± 123	52,08 ± 1,01	G. morhua	8297 ± 122	39,19 ± 0,9	M. aeglefinus	8179 ± 68	38,65 ± 0,01	P. virens
G.morhua and M. eglefinus (67—33)	9648 ± 310	47,68 ± 1,06	G. morhua	8267 ± 402	40,87 ± 2,22	M. aeglefinus	7729 ± 515	38,15 ± 1,71	M. merlangus
G.morhua and M. eglefinus (50—50)	8766 ± 319	43,24 ± 1,53	G. morhua	8245 ± 575	40,68 ± 3,1	M. aeglefinus	7628 ± 314	37,61 ± 1,68	M. merlangus
G.morhua and M. eglefinus (33—67)	9257 ± 150	45,17 ± 0,61	M. aeglefinus	8838 ± 551	43,12 ± 1,9	G. morhua	7730 ± 277	37,71 ± 0,71	M. merlangus
G.morhua and M. eglefinu s (10—90)	10542 ± 158	50,25 ± 0,88	M. aeglefinus	8916 ± 484	42,53 ± 2,2	G. morhua	8575 ± 72	40,86 ± 0,25	M. merlangus
G.morhua and M. eglefinus (1—99)	10665 ± 674	51,35 ± 1,58	M. aeglefinus	8165 ± 669	39,36 ± 2,04	G. morhua	7929 ± 680	38,17 ± 2,14	M. merlangus
G.morhua and M. eglefinus (0—100)	10470 ± 537	51,52 ± 2,01	M. aeglefinus	8053 ± 332	39,62 ± 1,19	M. merlangus	7680 ± 367	37,83 ± 1,37	G. morhua

Fig. 3. Best matched libraries of binary mixtures of G. morhua and M. eglefinus. In the Y-axis are indicated the different mixtures of both species. In axis X are represented the first, second and third best-matched libraries for each mixture.

For authentication purposes, molecular methods such as DNA barcoding are used as the standard method. However, DNA barcoding suffers from limitations concerning the various processes that commercial food products undergo, such as thermal processing (Hird et al., 2006). Proteomics approaches emerged as a complementary set of tools for species authentication of highly processed food products. Nevertheless, targeted proteomic techniques rely on protein identification (Nessen et al., 2016), which could be negatively affected by the lack of a well annotated genome for fish species of commercial or environmental interest (Lin et al., 2023).

In this study, we implemented the spectral library matching approach, a tandem mass spectrometry-based spectral library method, which does not rely on genetic information as protein identification has been bypassed saving the time of analyses. This method has been used previously for fish identification (Klapper et al., 2023; Nessen et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013). In this study, we expanded on previous works and demonstrated the applicability of this technique for the authentication of a wider array of fish species of from the North Sea and commercial products, respectively, and to investigate species specificity when closely related species are tested.

As a first step, using compareMS2, the quality of spectra acquired from samples was assessed and spectral libraries were created. Subsequently, we applied this method to laboratory-made mixtures of two of the most used fish species in fish-derived products namely, *G. morhua* (cod) and *M. aeglefinus* (haddock), two closely related gadoid species. The final step in our analysis was to apply the method to commercial products obtained from the local Norwegian stores and to laboratory-made mixtures imitating commercial products to evaluate the reliability of the method for fish species identification.

3.1. Molecular phylogeny of fish species using compareMS2

Proteomes of muscle tissue samples from 29 species were compared using a direct spectra comparison tool, compareMS2 2.0 for quality assurance prior to building spectral libraries from these samples (Marissen et al., 2023). The phylogenetic tree was built using the distance matrix generated based on spectra similarly and differences which arranged 29 fish species based on relatedness (Fig. 2). Most of the species from the same family were arranged according to taxonomic tree (Supplementary Fig. 1), where members from the Gadidae family were grouped. Similarly, all flatfish species were grouped together on a different distance branch from the Gadidae family.

3.2. Validation of spectral libraries

All 29 reference spectral libraries used for spectral matching were validated using the quality-control samples). Q*uality-control samples* corresponding to *Scophthalmus maximus*, *Merlangius merlangus,* and *Eutrigla gurnardus* were the same as the ones used for creating their libraries due to unavailability of fish individual in the sample archives.

As seen in Table 1, 26 of the 29 *quality-control samples* found the highest number of matching spectra (dot values over 0.8) against the spectral library corresponding to their specie. The second best-matching species corresponds to a close relative to the best-matching species as they have similar proteomes (Table 1, Fig. 2, Supplementary Fig. 1). The second-best match is a close relative to the best matching species due to similarities in proteome (Wulff et al., 2013). We found three *quality-control samples Trisopterus esmarkii*, *Pleuronectes platessa,* and *Ammodytes marinus* that did not match with their respective reference spectral library. In the case of *T. esmarkii* and *P. platessa*, their best match was *Trisopterus minutus* and *A. marinus* reference spectral library,

Table 3
First, second and third best matching Spectral Library for the commercial samples. Fish Burgers A: *A. silus* was 41%, *M. aeglefinus* was 15%, and *A. lupus* was 7% of total weight; Fish cake: G. morhua was 48% and *M. aeglefinus* was 32% of total weight; Fish Burgers B: G. morhua was 48% and *M. aeglefinus* was 32% of total weight.

Query sample	Matching spectra (best match)	% matching spectra (best match)	Reference library	Matching spectra (2nd best match)	% matching spectra (2nd best match)	Reference library	Matching spectra (3rd best match)	% matching spectra (3rd best match)	Reference library
Fish Burger A	6190 ± 333	35,74 ± 1,37	A. silus	5607 ± 196	32,38 ± 0,68	A. sphyraena	4197 ± 116	24,32 ± 0,27	M. aeglefinus
Fish Burger A (fish fillet)	6227 ± 464	36,44 ± 2,78	A. silus	5652 ± 339	33,04 ± 2,06	A. sphyraena	4264 ± 323	25,04 ± 1,77	M. aeglefinus
Fish Burger A (protein extract)	5857 ± 875	35,57 ± 3,73	A. silus	5223 ± 774	31,69 ± 3,29	A. sphyraena	3889 ± 148	23,81 ± 0,56	M. aeglefinus
Fish Cake	6628 ± 418	38,53 ± 2,06	G. mordua	6102 ± 328	35,46 ± 1,58	M. aeglefinus	5807 ± 131	33,75 ± 0,94	M. merlangus
Fish Cake (fish fillet)	6877 ± 182	39,98 ± 0,85	G. mordua	6741 ± 93	39,16 ± 0,16	M. aeglefinus	5866 ± 202	34,12 ± 1,42	M. merlangus
Fish Cake (protein extract)	6595 ± 360	39,59 ± 1,23	G. mordua	6230 ± 262	37,37 ± 0,76	M. aeglefinus	5280 ± 361	31,69 ± 1,39	M. merlangus
Fish Burger B	6752 ± 1020	39,78 ± 4,14	G. mordua	6277 ± 995	36,96 ± 4,11	M. aeglefinus	5699 ± 466	33,63 ± 1,13	M. merlangus
Fish Burger B (fish fillet)	6393 ± 127	37,93 ± 0,59	G. mordua	6360 ± 126	37,72 ± 0,76	M. aeglefinus	5570 ± 68	33,05 ± 0,81	M. merlangus
Fish Burger B (protein extract)	6909 ± 271	42,14 ± 0,36	G. mordua	6327 ± 138	38,57 ± 0,35	M. aeglefinus	5768 ± 471	35,12 ± 1,86	M. merlangus

respectively due to the phylogenetic relationship of this species (Fig. 2, Supplementary Fig. 1), as their proteome is similar.

Concerning *A. marinus* the *quality-control sample* matches with the corresponding library in a low degree. In Fig. 2, we can observe the phylogenetic tree created with compareMS2 based on mass spectra data, where *A. marinus* appears at the bottom of the graph, almost as an outgroup. However, when observing the genetic phylogenetic tree (Supplementary Fig. 1), this species is located between *A. sphyraena* and *Maurolicus muelleri*. These results indicate that the poor quality of the sample used to create the spectral library of *A. marinus,* due to either the sample preparation process or storage conditions indicating that spectral library was inaccurate. The example demonstrates the use of compareMS2 as a quality control tool for proteomic samples prior to spectral library building (Marissen et al., 2023; Varunjikar et al., 2022) due to its capability to rank various factors in sample preparation and find outliers from the dataset (Van Der Plas-Duivesteijn et al., 2016).

The binary mixtures and commercial products tested in the present work comprised four different species (*G. morhua, M. aeglefinus, A. lupus* and *A. silus*). Here we created 29 fish spectra reference libraries to showcase that, despite numerous potential options, the method described in principle typically ensures samples are consistently matched best with their corresponding library.

Regarding the *quality-control samples* of these four species, these always displayed the best match with their corresponding library with somewhat different values regarding the percentage of matching spectra. The best and second-best percentages of matching spectra for each species are 54,88 and 39,14 for *G. morhua*, 50,14 and 38,62 for *M. aeglefinus*, 44,81 and 20,43 for *A. lupus,* and 50,64 and 46,32 for *A. silus* (Table 1).

3.3. Spectral matching of binary mixtures of G. morhua and M. aeglefinus

G. morhua and *M. aeglefinus* are both species commonly used in commercial fish-based products in Europe. Some of these products could contain one or the other species, such as the famous “fish and chips” from the United Kingdom, where sometimes *G. morhua* is replaced by *M. aeglefinus; regardless* the product continues being sold as if containing the more expensive species (Wulff et al., 2013). In addition, different commercial products can contain different proportions of both species. This is because both species are phylogenetically closely related (Supplementary Fig. 1) and have a similar proteinic composition (Fig. 2). These similarities can make it challenging to distinguish which species is present in a commercial product. Additionally, if the product contains a mixture of species, it becomes difficult to determine which one is the most abundant. Before testing this in commercial products, we designed an experiment to determine our method’s sensitivity to distinguish between *G. morhua* and *M. aeglefinus*. For this, the muscle tissues of both species were mixed in different amounts and the resulting samples were analyzed. Table 2 and Fig. 3 show the matching results. In both species, when *G. morhua* and *M. aeglefinus* are the only species in the sample, the best matching library showed the correct result, displaying a percentage of matching at least ten points above the second-best library matching. It is interesting to notice that, as soon as there is a small amount of one of the species in the mixture, it is easily detected, even when it is only one percent, as happens with *G. morhua*. The sensitivity looks higher for the presence of *G. morhua*, because in a mixture at 50 % its corresponding library has the best match, but with a small difference.

Taken together, these results obtained in the present work show that the MS-based approach described here can with satisfactory sensitivity determine if *G. morhua* or *M. aeglefinus* are present in a mixture, and even more importantly, also determine which one of these two species is present whit a higher abundance.

Fig. 4. Best matched libraries of commercial products and their counterparts made with fillets and extracted proteins. In the Y-axis are indicated the different samples. In axis X are represented the first, second and third best-matched libraries for each mixture.

3.4. Spectral matching of commercial products

Once filleted and processed, the identification of the species present in most commercial fish products is difficult or hardly possible through visual inspection, highlighting the need for molecular identification. In addition, certain products undergo increasingly extensive processing, incorporating various compounds such as flour, preservatives, colorants, and other compounds, possibly followed by diverse cooking treatments. In this study, we wanted to assess if the spectral library matching approach described above could be used to determine the presence of fish species in processed commercial products.

To test the reliability of our approach in commercial products, we selected two different products commonly available in supermarkets in Norway namely, fish burgers and fish cakes. We selected two types of fish burgers (*Burger A* and *Burger B*) and one type of fish cake (*Fish Cake*). According to the food labels, the selected products contain 80 % fish; *Burger A*: 52 % *A. silus*, 19 % *M. aeglefinus* and 9 % of *A. lupus, Burger B*: 58 % *G. morhua* and 32 % *M. aeglefinus, Fish Cake*: 58 % *G. morhua* and 32 % *M. aeglefinus*. The remaining percentage corresponds to other types of ingredients such as onion, sugar, dextrose, and different preservatives (detailed in Supplementary Table 2).

Burger A, comprised of three distinct fish species. By comparison to the binary mixtures prepared in the laboratory and described above, this sample was a more complex mixture, and the composition of fish species was inaccurate (Table 3, Fig. 4). According to the label, the primary component is *A. silus*, with its corresponding library providing the best match. However, instead of *M. aeglefinus*, which should be the second-best match, we observed a significant number of matches in the library corresponding to *A. sphyraena*. Furthermore, the third component, *A. lupus*, did not rank among the top three matched libraries. This discrepancy could be attributed to the lower proportion of *A. lupus* in the burger, accounting for less than ten percent. The presence of *A. Sphyraena* can be explained by its similarity to *A. silus*, as evidenced by genetic and proteomic phylogenetic trees (Fig. 2 and Fig. 3). As previously observed for closely related species such as *G. morhua* and *M. aeglefinus,* due to the proteome level similarities spectra overlap is higher affecting quantification of species in mixtures when applying spectral library matching (Varunjikar et al., 2022). Whereas the accuracy in detecting both *M. aeglefinus* remains high; even with only 15 % tissue corresponding to this mixture, its library ranks third in terms of the proportion of matches. This indicate that phylogenetically distantly related species can be accurately identified and relatively quantified from mixtures using spectral library method as previously shown (Nessen et al., 2016; Ohana et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013).

Burger B and *Fish Cake* comprise the same fish species present in the same proportions; however, processing conditions and the remainder of the compounds present in these two food types differ. Regardless, in both samples we found the results obtained to be similar with the best-matching library being the one of *G. morhua*, followed by the one of *M. aeglefinus* (Fig. 4).

When analyzing mixtures created in the laboratory comprising the same compositions and proportions of the fish species as were present in the commercial products tested, we obtained the same results as were observed in the analysis of their commercial counterparts. This finding indicates that, despite the commercial products having added other edible materials such as flour, oil, vegetable, and spices than the targeted fish species, the spectral library matching approach described in the present work shows identified species and relatively quantified mixtures. As previously shown for product fried dish sold as “kibbeling” in Dutch markets, fried fish was correctly identified as cod even though the sample mostly consists of batter (Nessen et al., 2016). Moreover, commercial samples also might have been exposed to different food processing treatments. Obtaining comparable data from such samples and laboratory produced analogues thus corroborates that MS-based approaches are promising methods to include in the molecular toolsets

for the detection of food and feed fraud even in heavily processed samples. Additional work, however, is needed to increase the specificity and sensitivity of the spectral library method.

4. Conclusions

DNA-based methods are commonly used for species identification in food products. Proteomics methods can be employed in situations where DNA-based methods are not suitable, such as for heavily processed food products or adulterated food products containing very low amounts of DNA. In this study, we demonstrated that spectral library tandem mass spectrometry-based method is a reliable tool for determining the presence of fish species in commercial mixtures in most of the cases, without the need for protein identification, which is particularly valuable for species lacking genetic and / or proteomic information. This approach also offers the advantage of being less time-consuming compared to other bioinformatic approaches, yielding consistent results with both clean laboratory samples and commercial mixtures. In line with FAIR data practices (Pineda-Pampliega et al., 2022), the reference spectral libraries for 29 fish species of commercial and environmental importance in the North Sea and the scripts used were made available in a public repository and can in future be used to help tackle current and emerging challenges in feed and food authentication analyses when prospecting new resources in the Arctic, reducing the time needed.

CRediT authorship contribution statement

Madhushri S. Varunjikar: Conceptualization, Experimental design, Laboratory work, Methodology, Bioinformatics, Data analyses, Writing original draft, and editing. **Javier Pineda-Pampliega:** Writing – review & editing, Writing – original draft, Methodology, Bioinformatics, Data analyses, Formal analysis, Data curation. **Ikram Belghit:** Writing – review & editing, Writing – original draft, Methodology. **Magnus Palmblad:** Writing – review & editing, Writing – original draft, Methodology. **Bjørn Einar Grøsvik:** Writing – review & editing, Writing – original draft, Methodology. **Sonnich Meier:** Writing – review & editing, Writing – original draft, Methodology. **Pål Asgeir Olsvik:** Writing – review & editing, Writing – original draft. **Kai K. Lie:** Writing – review & editing, Writing – original draft, Supervision, Methodology. **Josef D. Rasinger:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The author, Dr. J. D. Rasinger, is currently employed with the European Food Safety Authority (EFSA) in the Food Ingredients and Packaging Unit (FIP). However, the present article is published under the sole responsibility of the author Dr. J. D. Rasinger and may not be considered as an EFSA scientific output. The positions and opinions presented in this article are those of the author alone and are not intended to/do not necessarily represent the views/any official position of EFSA. The author's main contributions to this article were made before joining EFSA when still employed at the Institute of Marine Research. The other authors declare no conflict of interest.

Data availability

Processed mass spectrometry data and spectra libraries were deposited in an online repository massive.ucsd.edu and ProteomeXchange (MassIVE id: MSV000094602 and ProteomeXchange id: PXD051671).

Acknowledgment

The present work was financed by the Norwegian Research Council (NFR) project "*ClimSeaFood*" (grant number 15804), "skriveprosjekt fremtidens havbruk- risiko basert overvåkning og regelverk i bærekraftige fôr" (grant number 16053), and the "MultiOmicsTools" project (Ministry of Trade, Industry and Fisheries, 15470, IMR).

HPLC-MS/MS proteomic analyses were carried out at the Proteomics Unit at University of Bergen (PROBE). Some figures were generated using app.biorender.com with laboratory license for publications. For figure generation fish images were downloaded from www.artsdatabanken.no database and cited in respective figure with credits.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2024.114785.

References

Bi, H., Zhong, C., Shao, M., Wang, C., Yi, J., Qiao, L., & Zhang, J. (2019). 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 Communications in Mass Spectrometry, 33*(16), 1336–1343. https://doi.org/10.1002/RCM.8474

Bouzembrak, Y., Steen, B., Neslo, R., Linge, J., Mojtahed, V., & Marvin, H. J. P. (2018). Development of food fraud media monitoring system based on text mining. *Food Control*. https://doi.org/10.1016/j.foodcont.2018.06.003

Carrera, M., Cañas, B., & Gallardo, J. M. (2013a). Fish authentication. In *Proteomics in foods (Issue November* (pp. 205–222). US: Springer. https://doi.org/10.1007/978-1-4614-5626-1_12.

Carrera, M., Cañas, B., & Gallardo, J. M. (2013b). Proteomics for the assessment of quality and safety of fishery products. *Food Research International, 54*(1), 972–979. https://doi.org/10.1016/j.foodres.2012.10.027

Chien, H. J., Huang, Y. H., Zheng, Y. F., Wang, W. C., Kuo, C. Y., Wei, G. J., & Lai, C. C. (2022). Proteomics for species authentication of cod and corresponding fishery products. *Food Chemistry, 374*, Article 131631. https://doi.org/10.1016/J.FOODCHEM.2021.131631

Deconinck, D., Volckaert, F. A. M., Hostens, K., Panicz, R., Eljasik, P., Faria, M., Monteiro, C. S., Robbens, J., & Derycke, S. (2020). A high-quality genetic reference database for European commercial fishes reveals substitution fraud of processed Atlantic cod (Gadus morhua) and common sole (Solea solea) at different steps in the Belgian supply chain. *Food and Chemical Toxicology, 141*(May), Article 111417. https://doi.org/10.1016/j.fct.2020.111417

Deutsch, E. W., Mendoza, L., Shteynberg, D. D., Hoopmann, M. R., Sun, Z., Eng, J. K., & Moritz, R. L. (2023). Trans-proteomic pipeline: robust mass spectrometry-based proteomics data analysis suite. *Journal of Proteome Research, 22*(2), 615–624. https://doi.org/10.1021/acs.jproteome.2c00624

Deutsch, E. W., Mendoza, L., Shteynberg, D., Slagel, J., Sun, Z., & Moritz, R. L. (2015). Trans-proteomic pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *Proteomics - Clinical Applications, 9*(7–8), 745–754. https://doi.org/10.1002/prca.201400164

Eng, J. K., Jahan, T. A., & Hoopmann, M. R. (2013). Comet: An open-source MS/MS sequence database search tool. *Proteomics, 13*(1), 22–24. https://doi.org/10.1002/pmic.201200439

French, I., & Wainwright, B. J. (2022). DNA barcoding identifies endangered sharks in pet food sold in Singapore. *Frontiers in Marine Science, 9*, Article 836941. https://doi.org/10.3389/FMARS.2022.836941/BIBTEX

Gu, S., Deng, X., Shi, Y., Cai, Y., Huo, Y., Guo, D., & Han, F. (2020). Identification of peptide biomarkers for authentication of Atlantic salmon and rainbow trout with untargeted and targeted proteomics approaches and quantitative detection of adulteration. *Journal of Chromatography B, 1155*, Article 122194. https://doi.org/10.1016/J.JCHROMB.2020.122194

Hird, H., Chisholm, J., Sanchez, A., Hernandez, M., Goodier, R., Schneede, K., Boltz, C., & Popping, B. (2006). Effect of heat and pressure processing on DNA fragmentation and implications for the detection of meat using a real-time polymerase chain reaction. *Food Additives and Contaminants, 23*(7), 645–650. https://doi.org/10.1080/02652030600603041

Hu, L., Zhang, H., Hu, Z., Chin, Y., Zhang, X., Chen, J., & Hu, Y. (2022). Comparative proteomics analysis of three commercial tuna species through SWATH-MS based mass spectrometry and chemometrics. *Food Control, 141*, Article 109162. https://doi.org/10.1016/J.FOODCONT.2022.109162

Kessner, D., Chambers, M., Burke, R., Agus, D., & Mallick, P. (2008). ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics, 24*(21), 2534–2536. https://doi.org/10.1093/bioinformatics/btn323

Klapper, R., Velasco, A., Döring, M., Schröder, U., Sotelo, C. G., Brinks, E., & Muñoz-Colmenero, M. (2023). A next-generation sequencing approach for the detection of mixed species in canned tuna. *Food Chemistry: X, 17*, Article 100560. https://doi.org/10.1016/J.FOCHX.2023.100560

Lam, H. (2011). Building and searching tandem mass spectral libraries for peptide identification. *Molecular and Cellular Proteomics, 10*(12), 1–10. https://doi.org/10.1074/mcp.R111.008565

Lin, M., Varunjikar, M., Lie, K., Søfteland, L., Dellafiora, L., Ørnsrud, R., Sanden, M., Berntssen, M., Dorne, J., Bafna, V., & Rasinger, J. (2023). Multi-tissue proteogenomic analysis for mechanistic toxicology studies in non-model species. *Environment International, 108309*. https://doi.org/10.1016/J.ENVINT.2023.108309

Marissen, R., Varunjikar, M. S., Laros, J. F. J., Rasinger, J. D., Neely, B. A., & Palmblad, M. (2023). compareMS2 2.0: An improved software for comparing tandem mass spectrometry datasets. *Journal of Proteome Research, 22*(2), 514–519. https://doi.org/10.1021/acs.jproteome.2c00457

Nessen, M. A., van der Zwaan, D. J., Grevers, S., Dalebout, H., Staats, M., Kok, E., & Palmblad, M. (2016). Authentication of closely related fish and derived fish products using tandem mass spectrometry and spectral library matching. *Journal of Agricultural and Food Chemistry, 64*(18), 3669–3677. https://doi.org/10.1021/acs.jafc.5b05322

Ohana, D., Dalebout, H., Marissen, R. J., Wulff, T., Bergquist, J., Deelder, A. M., & Palmblad, M. (2016). Identification of meat products by shotgun spectral matching. *Food Chemistry, 203*, 28–34. https://doi.org/10.1016/j.foodchem.2016.01.138

Pardo, M.Á., Jiménez, E., Viðarsson, J. R., Ólafsson, K., Ólafsdóttir, G., Daníelsdóttir, A. K., & Pérez-Villareal, B. (2018). DNA barcoding revealing mislabeling of seafood in European mass caterings. *Food Control, 92*, 7–16. https://doi.org/10.1016/J.FOODCONT.2018.04.044

Pineda-Pampliega, J., Bernhard, A., Hannisdal, R., Ørnsrud, R., Mathisen, G. H., Solstad, G., & Rasinger, J. D. (2022). Developing a framework for open and FAIR data management practices for next generation risk- and benefit assessment of fish and seafood. *EFSA Journal, 20*(S2). https://doi.org/10.2903/J.EFSA.2022.E200917

REGULATION (EU) No 1169/2011, (testimony of REGULATION (EU) No 1169/2011). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:304:0018:0063:en:PDF.

Shokralla, S., Hellberg, R. S., Handy, S. M., King, I., & Hajibabaei, M. (2015). A DNA mini-barcoding system for authentication of processed fish products. *Scientific Reports, 5*, 1–11. https://doi.org/10.1038/srep15894

Leahy, S. (2021). Revealed: Seafood fraud happening on a vast global scale | Fish | The Guardian. *The Guardian*, 1–7. https://www.theguardian.com/environment/2021/mar/15/revealed-seafood-happening-on-a-vast-global-scale.

Van Der Plas-Duivesteijn, S. J., Klychnikov, O., Ohana, D., Dalebout, H., Van Veelen, P. A., De Keijzer, J., Nessen, M. A., Van Der Burgt, Y. E. M., Deelder, A. M., & Palmblad, M. (2016). *Differentiating samples and experimental protocols by direct comparison of tandem mass spectra*. Doi: 10.1002/rcm.7494.

Varunjikar. (2023). *Proteomic Tools for Food and Feed Authentication, University of Bergen*. https://hdl.handle.net/11250/3053481.

Varunjikar, M. S., Moreno-Ibarguen, C., Andrade-Martinez, J. S., Tung, H.-S., Belghit, I., Palmblad, M., Olsvik, P. A., Reyes, A., Rasinger, J. D., & Lie, K. K. (2022). Comparing novel shotgun DNA sequencing and state-of-the-art proteomics approaches for authentication of fish species in mixed samples. *Food Control, 131*, Article 108417. https://doi.org/10.1016/j.foodcont.2021.108417

Visciano, P., & Schirone, M. (2021). Food frauds: Global incidents and misleading situations. *Trends in Food Science & Technology, 114*, 424–442. https://doi.org/10.1016/J.TIFS.2021.06.010

Wulff, T., Nielsen, M. E., Deelder, A. M., Jessen, F., & Palmblad, M. (2013). Authentication of fish products by large-scale comparison of tandem mass spectra. *Journal of Proteome Research, 12*(11), 5253–5259. https://doi.org/10.1021/pr4006525

Zuccolo, V., Rego, F. M., Hughes, E., & Griffiths, A. M. (2023). Endangered shark species traded as "cação" in São Paulo during the COVID-19 lockdown: DNA-barcoding a snapshot of products. *Molecular Biology Reports, 50*(12), 9985–9992. https://doi.org/10.1007/S11033-023-08876-6/FIGURES/4