Food Control 137 (2022) 108888

Contents lists available at ScienceDirect

Food Control

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

Shotgun proteomics approaches for authentication, biological analyses, and allergen detection in feed and food-grade insect species

Madhushri S. Varunjikar [a,1], Ikram Belghit [a,1], Jennifer Gjerde [a], Magnus Palmblad [b], Eystein Oveland [a], Josef D. Rasinger [a,*]

[a] Institute of Marine Research, P.O. Box 1870 Nordnes, 5817, Bergen, Norway
[b] Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, the Netherlands

ARTICLE INFO

Keywords:
Tandem mass spectrometry
Edible insect
Feed control
Spectral libraries
Insect allergens
Q exactive orbitrap

ABSTRACT

Untargeted proteomics can contribute to composition and authenticity analyses of highly processed mixed food and feed products. Here, we present the setup of an analytical flow tandem mass spectrometry method (AF-HPLC HR-MS) for analysis of insect meal from five different species. Data acquired were compared with previously published data employing spectra matching and standard bottom-up proteomics bioinformatics analyses. In addition, data were screened for insect species marker peptides and common allergens, respectively. The results obtained indicate that the performance of the newly established AF-HPLC HR-MS workflow is in line with previously published methods for insect species differentiation. Data obtained in the present study, also lead to the discovery of novel markers for the development of targeted MS analyses of insect species in food- and feed-mixes and highlighted that known allergen such as arginine kinase or tropomyosin were consistently detected across all five species tested.

1. Introduction

In 30 years, 9.7 billion people are estimated to live on our planet and the demand for feed and food crops is expected to increase to 25–70% above today's levels (FAO et al., 2018). To ensure food security for the growing population, novel food and feed ingredients such as insects will play an important role as future protein sources in animal feed and human nutrition (IPIFF, 2021). However, in the European Union (EU), their current and future usage in the feed and food sector is and will be regulated by strict legislative texts. To enforce and monitor regulatory guidelines robust and versatile high-throughput analytical tools will be required; in this context mass-spectrometry (MS) based proteomics approaches have shown to hold great promise (Belghit et al., 2021; Lecrenier et al., 2018; Varunjikar et al., 2022).

The most common proteomics workflow takes the bottom-up approach in which proteins in the sample are enzymatically digested by a protease (e.g., trypsin), and the resulting peptides are analysed by high-performance liquid chromatography (HPLC) coupled to a tandem mass spectrometer (HPLC-MS/MS). The data output files including both MS and MS/MS spectra are then analysed using different proteomics bioinformatics tools that allow for peptide identification and protein inference based on different algorithms. The combinations of a quadrupole with a high resolution TOF analyser (QTOF) or with an high resolution orbitrap mass spectrometer (HR-MS) are among the most widely used for shotgun proteomics analyses (Szabó et al., 2021). Untargeted proteomic workflows commonly aim to identify as many peptides and proteins as possible and usually utilise nanoflow HPLC (nano-LC) for chromatographic separation of samples. Nano-LC is more sensitive than normal flow approaches are and hence, the preferred choice in bottom-up proteomics. However, the use of nano-flow LC is technically challenging, and frequent column changes are required due to faster build-up of high back pressure when compared to normal flow HPLC. Normal flow HPLC, also referred to as analytical flow (AF) HPLC, is simpler to set up, more robust to run in routine proteomic analysis settings (Lenčo et al., 2018). Thus, in regulatory laboratories for high-throughput feed or food safety and authenticity analyses, the use of AF-HPLC-MS/MS-based proteomics can contribute to make implementation of proteomic approaches attractive and affordable for control laboratories (Sentandreu & Sentandreu, 2011).

* Corresponding author.
E-mail addresses: madhushri.shrikant.varunjikar@hi.no (M.S. Varunjikar), ikram.belghit@hi.no (I. Belghit), jennifer.gjerde@hi.no (J. Gjerde), n.m.palmblad@lumc.nl (M. Palmblad), eystein.oveland@hi.no (E. Oveland), josef.rasinger@hi.no (J.D. Rasinger).
[1] Equal contribution; shared first authorship.

https://doi.org/10.1016/j.foodcont.2022.108888
Received 10 December 2021; Received in revised form 8 February 2022; Accepted 10 February 2022
Available online 23 February 2022

Abbreviations

AF	Analytical flow
BSF	Black Soldier Fly
HC	House Cricket
HR-MS	High Resolution-Mass Spectrometry
LW	Lesser mealworm
MGF	Mascot Generic Format
mzML	open standard data format for mass spectrometry data
MF	Microflow
MW	Morio worms
PSM	Peptide-Spectrum Match
SLM	Spectral Library Matching
TPP	Trans-Proteomic Pipeline
QTOF	Quadrupole time-of-flight
HPLC-MS/MS	High-Performance Liquid Chromatography coupled to Tandem Mass Spectrometry
YW	Yellow meal Worm

Proteomic-based methods using HPLC-MS/MS were recently identified as promising tools to complement current standard techniques of processed animal protein (PAP) detection in feed in a scientific opinion by the European Food Safety Authority (EFSA) (Aguilera et al., 2018). According to European regulation of animal protein (European Commission, 2013/51, European Commission, 2017/893), insects reared to produce PAP are to be considered farmed animals. In 2017, the European Commission (EC) allowed the use of insect meal processed from seven different black soldier fly (BSF) (*Hermetia illucens*), common housefly (*Musca domestica*), yellow mealworm (*Tenebrio molitor*), lesser mealworm (*Alphitobius diaperinus*), house cricket (*Acheta domesticus*), banded cricket (*Gryllodes sigillatus*) and field cricket (*Gryllus assimilis)* (European Commission, 2017/893). Silkworm (*Bombyx mori*) was recently added to the approved list of insect species in aquaculture (European Comission, 2021/1372), resulting in a total of eight insect species allowed in aquafeed. Recently, in August 2021, the EC adopted the decision to allow the use of insect PAP in formulated pig and poultry feeds (European Comission, 2021/1372). At the time of writing, a draft bill for implementing the regulation to authorise the commercialisation of frozen and dried migratory locust (*Locusta migratoria)* on the EU market was issued (IPIFF, 2021), and following a favourable opinion of the EFSA Panel on Nutrition, Novel Foods and Food Allergens (NDA) (EFSA NDA panel, 2021a), the placing on the market of dried yellow mealworm (*Tenebrio molitor*) larva as a novel food under Regulation (EU) 2015/2283 was authorized (European Comission, 2021/882). Also, a favourable opinion on the draft legal act authorising the placing on the market of frozen, dried and powder forms of house cricket (*Acheta domesticus*) as a novel food was issued (EFSA NDA panel, 2021b). Concerning house cricket, EFSA highlighted that the consumption of the evaluated insect proteins may potentially lead to allergic symptoms and that in addition, allergens present in substrate fed to insects may end up in the insect consumed (EFSA NDA panel, 2021b). Therefore, analytical approaches must be developed which allow for an unambiguous detection and identification of white-listed insect species in insect-protein containing feed or food products. Among the five insect species used in this study, four are white-listed insect species whereas one species, morio worm (*Zophobas morio*), is not officially approved in the EU for use in feed or food but is considered a potential future feed or food ingredient (Rumbos & Athanassiou, 2021). This species was not included in previously published Belghit et al., 2019 but as it might be used as a food and feed in future, we included it in the current study.

For safe use of insects in feed and food real-time polymerase chain reaction (qPCR) assays are being developed (Daniso et al., 2020; Debode et al., 2017; Garino et al., 2021; Köppel et al., 2019). In parallel, targeted and non-targeted HPLC-MS based proteomics methods are being developed by several laboratories. Analyses of MS/MS spectra were shown to be suitable for the identification, quantification and tracing of processed animal protein (PAP) in feed (Belghit et al., 2019, 2021; Marbaix et al., 2016; Rasinger et al., 2016; Steinhilber et al., 2018a,b), the detection of allergens in edible insects (Bose et al., 2021; Francis et al., 2019), and the identification of species origin. When genomic information is scarce (Belghit et al., 2019, 2021; Nessen et al., 2016; Ohana et al., 2016; Rasinger et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013).

The objectives of the present study were to (*i*) set up and optimise an analytical flow LC-MS/MS proteomics assay for insect species authentication, (*ii*) compare data obtained from two different proteomics workflows, microflow HPLC (MF-HPLC) QTOF and the optimized AF-HPLC HR-MS, using spectra matching approaches, and (*iii*) based on both MF-HPLC QTOF data and AF-HPLC HR-MS data, identify common, and unique insect species-specific proteins, and potential allergens.

2. Materials and methods

2.1. Samples

HeLa Protein Digest Standards were purchased from Thermo Fisher Scientific Pierce™ (Thermo Scientific, San Jose, CA) and was used for standardisation of the instrument and optimising the HPLC and MS conditions with the HR-MS orbitrap instrument. Eight samples from species of the Diptera order; black soldier fly larvae (BSF) (*H. illucens*), nine samples from species of the Coleoptera order, including the yellow mealworm (YW) (*T. molitor*) and the lesser mealworm (LW) (*A. diaperinus*), and two samples from the Orthoptera order; house cricket (HC) (*A. domesticus*) were collected from different insect food and feed companies. The eighteen insect meal samples have been reported in more detail elsewhere (Belghit et al., 2019). Additionally, one morio worm (MW) (*Z. morio)* sample was included in the current study (Supplementary Table 1).

2.2. Protein extraction

Insect samples were weighed into a test tube of the One Plus Grinding kit (GE Healthcare Life Science, 80648337, Piscataway, NJ, USA) and lysis buffer (4% SDS, 0.1 M Tris-HCl, pH 7.6). Samples were kept on ice and homogenised in the tube containing resins with a pestle. To this homogenate, freshly prepared, 1 M Dithiothreitol was added to obtain a final concentration of 0.1 M, further, these tubes were centrifuged for 10 min at 15,000 g to remove resin and other debris. The supernatant was collected and heated at 95 °C on the heat-block for 5 min. After this, samples were centrifuged again, and the supernatant was collected in new tubes to store at −20 °C until further processing. The protein concentration of extracted samples was determined by the Pierce 660 assay as described in Rasinger et al. (2016) using BSA for the standard curve (Thermo Scientific, San Jose, CA).

2.3. Protein digestion and purification

Protein extracts from insect samples were digested with filter-aided sample preparation method as described in Belghit et al. (2019), where 150 mg of extracted protein was diluted with 200 μL of 8 M urea solution prepared in Tris-HCl (100 mM, pH 8.5). This solution was transferred to an ultrafiltration spin column (Microcon 30, Millipore, Burlington, MA, USA). Further, these proteins were alkylated as described in Belghit et al. (2019) with 50 mM of iodoacetamide for 20 min for incubation in darkness at room temperature. After incubation, the protein mixture in the column was washed with 200 μL of 8 M urea solution along with 100 μL of 50 mM ammonium bicarbonate solution.

After this step trypsin was added to the filters in 1:50 enzyme to protein ratio and tubes were incubated for 16 h at 37 °C. After incubation filters were centrifuged and washed with 40 μL of 50 mM ammonium bicarbonate solution with the same molarity as mentioned above and later with NaCl (0.5 M). Following centrifugation, the digested tryptic peptides were purified with Pierce™ C18 spin column (ThermoFisher, 89870). The columns were first washed with methanol/water (50/50, v/v), and then equilibrated with wash solvent (acetonitrile/trifluoroacetic acid/water, 5/0.5/94.5, v/v/v). Digested samples were diluted with acetonitrile/trifluoroacetic acid/water (20/2/78, v/v/v) and loaded into the columns. Peptides were eluted with acetonitrile/water (30/70, v/v) and subsequently evaporated in a speed vacuum dryer (LABCONCO CentriVap micro IR). Peptide pellets were dissolved in acetonitrile/formic acid/water (2/0.1/97.9, v/v/v) and kept at −20 °C until mass spectrometric analyses.

2.4. LC-MS/MS analyses

2.4.1. QTOF

For the ESI-MS/MS maXis Impact UHR-TOF (Bruker, Bremen Germany), the method is described in Belghit et al. (2019). Briefly, HPLC analyses were performed using the UltiMate 3000 HPLC system (Thermo Scientific, San Jose, CA). Approximately 5.0 μg samples were separated using 2.0 μm Acclaim PepMap 100 C18, 1 × 150 mm (Thermo Scientific, San Jose, CA). The flow rate was 40 μL/min. Mobile phase A was 95% water, 5% acetonitrile, 0.1% formic acid. Mobile phase B was 20% water, 80% acetonitrile, 0.1% formic acid. The digest was injected, and the organic content of the mobile phase was increased linearly from 4% B to 40% B in 60 min and from 40% B to 90% B in 10 min, and then washed with 90% B for 10 min and with 4% B for 10 min, for a total of 90 min. The column effluent was directly connected to the maXis UHR-TOF coupled with electrospray ionisation (ESI) (Bruker, Billerica, MA, USA). In the survey scan, MS spectra were acquired for 0.5 s in the mass to charge (m/z) range between 50 and 2200. The 10 most intense peptides ions 2+ to 4+ were fragmented during a cycle time of 3 s. The collision-induced dissociation (CID) energy was automatically set according to the m/z ratio and charge state of the precursor ion. The mass spectrometer and HPLC systems are controlled by Compass HyStar 3.2 (Bruker, Billerica, MA, USA). Regarded as micro flow-HPLC QTOF (MF-HPLC QTOF) here onwards in the text.

2.4.2. HR-MS Orbitrap

For the optimisation, HPLC analyses were performed using Vanquish Horizon binary HPLC (Thermo Scientific, San Jose, CA). Separations were performed using 2.2 μm Acclaim Vanquish C18, 2.1 × 250 mm (Thermo Scientific, San Jose, CA). The column temperature was maintained at 50 °C. The solvents A and B were 0.1% (v/v) formic acid in high purity water (18.2 MΩ × cm) and 0.1% formic acid (v/v) in 100% acetonitrile, respectively. Gradient conditions are described in Supplementary Table 2, with different gradient lengths varying from 60 to 80 min. The flow rate varied between 300 and 400 μL/min (Supplementary Table 2). Different amounts of HeLa cells digest were loaded (0.5–40 μg, Supplementary Table 3).

Eluting peptides were analysed on HR-MS Q Exactive Orbitrap (Thermo Scientific, San Jose, CA). MS instrumental tune parameters were set as follows: ESI spray voltage was 3.5 kV, sheath gas flow rate was 40 AU, the auxiliary gas flow rate was 10 AU, the capillary temperature was 320 °C, probe heater temperature was 400 °C and S-lens RF level was set to 50. Data-dependent acquisition (DDA) MS2 method with full MS scans in positive polarity was obtained at resolution settings of 17,500, 35,000, and 70,000 (Supplementary Table 2). Mass range was set at 200–2000 m/z and an AGC target was 5.0×10^5 up to 3.0×10^6 with a maximum injection time of 50 ms. For MS2, the resolution settings were 17,500 and 35,000 at a fixed first mass of 140 m/z with an AGC target value of 5.0×10^5 and an isolation window of 1.2 m/z. The normalised collision energy set was 32 and the top 10 precursors were selected for fragmentation. The signal intensity threshold was 2.0×10^4 with dynamic exclusion of 10, 20 and 30 s (Supplementary Table 2). This is regarded as analytical flow- HPLC HR-MS (AF-HPLC HR-MS) here onwards in the text.

After the optimisation of the HPLC and MS parameters with the HeLa Digest, the developed HR-MS workflow was implemented to analyse the nineteen insect meal samples. Gradient conditions were as follows: 2% B to 35% B in 62 min, hold at 95% B until 5 min and 2% B from 67.1 until 80 min. The flow rate was 400 μL/min flow rate (test number 19 in Supplementary Table 2). MS scans were obtained at a resolution of 70,000. Mass range was set at 350–2000 m/z and an AGC target was 3.0×10^6 with a maximum injection time of 50 ms. For MS2, the resolution was 35,000 at a fixed first mass of 140 m/z with an AGC target value of 3.0×10^6 and an isolation window of 1.2 m/z. The normalised collision energy set was 32 and the top 10 precursors were selected for fragmentation. The signal intensity threshold was 2.0×10^4 with dynamic exclusion of 30 s.

2.5. Bioinformatic analyses

2.5.1. Direct spectral comparison and Spectral library building with SpectraST

Proteomic-based phylogenetic data analysis was performed as described in Varunjikar et al. (2022). In short for direct spectral comparison of tandem mass spectra using compareMS2 (compareMS2, 2021; GUI, 2021; Palmblad & Deelder, 2012) MGF files containing the top 500 most intense tandem mass spectra were created using msConvert (version: 3.0., ProteoWizard). CompareMS2 was used to create distance matrices and phylogenetic trees. Overview of bioinformatics analyses is given in Supplementary Fig. 1.

Using the mzML and pepXML files generated from MF-HPLC QTOF and AF-HPLC HR-MS data and search output, spectral libraries (SLs) were created for each of all the five insect species using SpectraST (version 5.0) as previously described (Belghit et al., 2021). Matching spectra with dot products above 0.8 were considered to be valid matches and the unique identifiers of these spectra were extracted and exported into a text file. Post-processing of the results was done in R (version 4.0.3) Outputs were recorded using tidyverse functions (version 1.3.0) and UpSetR (version 1.4.0).

2.5.2. Protein identification and data analysis

For analyses of acquired spectra from HeLa cell digest MSGFplus (V.1.26.0 (Pedersen, 2021).) search engine was used in R interphase to match the spectra to the UniProt human reference proteome (up000005640). Post analyses were done in R (version 4.0.3).

For identification of PSM, peptides, and proteins and to compare percentage identification from MF-HPLC QTOF and AF-HPLC HR-MS, tandem mass spectra were searched against proteomes of respective species from UniProt databased as described in 2.1 (accessed on June 2021).

For proteome analyses and marker detection, acquired data were matched against reviewed sequences (12, 976) from Arthropoda species (accessed July 2021) using Comet search as implemented in the Trans-Proteomics Pipeline (TPP) (version 5.2.0 (Deutsch et al., 2015)). In all searches, precursor mass tolerance was set to 20 ppm, trypsin was selected as a digestive enzyme (allowing for two non-enzymatic termini), and carbamidomethylation of cysteine and oxidation of methionine were set as fixed and variable modification, respectively. Generated pepXML files were further analysed using PeptideProphet and ProteinProphet using 1% level false discovery rate (FDR) (Keller et al., 2002). Post-processing of the acquired data was done in R (version 4.0.3). Data processing and statistical comparison of proteomics samples were performed in Omics Explorer V 3.6 (Qlucore AB, Lund, Sweden). The data were analysed using two-way ANOVA of the involved insect species (groups were sample species), unsupervised principal component analyses (PCA) and hierarchical cluster analysis (HCA). For

comparing the detected protein Venn diagrams were created using www.biovenn.nl.

2.5.3. Allergen detection

For allergen detection, a list of food allergens was downloaded from (www.allergen.org) along with allergen families and biochemical names (48 sequences) and these allergen sequences were downloaded from UniProt to create a database. The collected data from each instrument were searched against the database using TPP to evaluate allergen detection ability. Data processing and statistical comparison of detected allergenic proteins from samples were performed in Omics Explorer.

3. Results and discussion

3.1. Set up and optimisation of an AF-HPLC HR-MS system for untargeted proteomics

In the present study, the performance of an AF-HPLC coupled to a standard HR-MS was tested by injecting different amounts of HeLa cell digests using different combinations of HPLC and MS and MS/MS2 settings. Since the objective of the present work was to develop a time-efficient method suitable for regulatory use, only three relatively short HPLC run-time lengths (60, 70 and 80 min) with an increasing gradient of 4% (v/v) to 50% (v/v) mobile phase B were tested; run-time lengths of 90 min and longer, which commonly are employed in non-targeted expression proteomics analyses (Kelstrup et al., 2014; Varunjikar et al., 2022), were considered impractical for use in routine regular analyses settings.

As expected, increasing the gradient time resulted in an increased number of tandem mass spectra (Supplementary Table 2). Using a run-time length of 80-min and 20 μg of HeLa digest, yielded a total of 13562 of spectra. When matched against the HeLa cell reference proteome (up000005640) using the MSGFplus search engine this resulted in 8946 peptide-spectrum matches (PSM's) and the identification of 7553 and 1951 unique peptides and proteins, respectively (Supplementary Table 2). Similar results were obtained for the analysis of 5 μg of HeLa digest over a 90 min gradient, with a nanoflow HPLC instrument coupled to Linear Trap Quadrupole (LTQ) Orbitrap Velos mass spectrometer (Michalski et al., 2011) and when analysing 20 μg of HeLa digest using a Standard flow multiplexed Proteomics (SFloMPro) system coupled with a HR-MS Classic using a 90 min gradient (Jenkins & Orsburn, 2020).

Peptide and protein identification on an AF-HPLC HR-MS (as well as on any other HPLC-MS/MS systems), in addition to gradient length, are also dependent on injected sample amounts, which must be optimised for each respective system (Jenkins & Orsburn, 2020). Recently, Lenčo et al. (2018), analysed 0.5 and 2 μg of HeLa digest and observed an increase in protein and peptide identification of up to 14% with 2 μg compared to 0.5 μg of HeLa digest. In the aforementioned study, the authors optimised a standard-flow HPLC-MS system with the aim to identify a sample loading amount that yielded a comparable number of proteins and peptides that usually can be identified when using nano LC-MS systems (Lenčo et al., 2018). In the present study, loading amounts of 0.5–40 μg HeLa digest were analysed. As can be seen in Supplementary Fig. 2 and Supplementary Table 3, the PSM, unique peptide and protein counts increased linearly with increasing amounts of HeLa digest up to 5 μg when a plateau was reached. A 10-fold increase in sample load in the column (in the range of 0.5–5 μg peptide) increased the identification rate of peptides and proteins to 23% and 11%, respectively (Supplementary Fig. 2 and Supplementary Table 3). No further increase in the number of features detected was observed when up to 40 μg peptide (the highest amount of HeLa digest tested in the present study) were injected. Hence, 5 μg were selected for further analyses of the insect meal samples.

Taken together, the data generated here suggest that, given that the sample quantity is not a limiting factor (Jenkins & Orsburn, 2020), using an AF-HPLC HR-MS, could be a viable alternative for use in regulatory laboratories to the more conventional nanoflow HPLC workflow routinely used in MS-based proteomics.

3.2. Quality control and insect species identification

Following setup and optimisation of the AF-HPLC HR-MS setup, the assay settings shown in test 19, Supplementary Table 2, were applied for comparing analysis outputs of insect-based MS data generated in the present study with data previously published by our group (Belghit et al., 2019). Insect MS data acquired previously on an MF-HPLC QTOF instrument (massIVE ID: MSV000083737) were reanalysed using compareMS2 and a TPP-based bioinformatics workflow to compare with data generated here (AF-HPLC HR-MS, massIVE ID: MSV000088034). In both studies (MF-HPLC QTOF and AF-HPLC HR-MS based workflows), similar gradient lengths (varying flow rates) and loading amounts of insect meal samples (80 min and 5.0 μg, respectively) were applied.

Analysis outputs from compareMS2 have previously been found to be useful in the determination of the effects of sample preparation and analysis approaches on the data acquired by mass spectrometry (Van Der Plas-Duivesteijn et al., 2016). Using compareMS2, in the present study distance matrixes were calculated for both insect datasets and two representative dendrograms were constructed. As shown in Fig. 1, mass spectra from both datasets were successfully arranged according to the insect species and molecular phylogeny of insects (Supplementary Fig. 3), respectively. The spectral clustering of insects reflects the relatedness of insect species at the taxonomic level and is in line with data shown previously where insect grouping based on MS data was found to be based on the orders Diptera, Coleoptera, and Orthoptera (Belghit et al. (2019). In other words, overall, all insect species analysed in the present study were well separated using compareMS2, indicating that even with only 500 spectra collected, using AF-HPLC and a routine MS instrument sufficient data can be generated to allow for a species-level differentiation of protein sources in food- and feed samples. The spectral distances obtained by pairwise spectra comparison of data acquired with the HR-MS also were comparable with those obtained using the previously published MF-HPLC QTOF data (Belghit et al., 2019). This was consistent with previous molecular phylogenetic studies conducted using compareMS2 in selected species of interest in relation to food- and feed authenticity and adulteration analyses, respectively (Ohana et al., 2016; Rasinger et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013).

In addition to the compareMS2 analyses, we subjected the previously published data instrument (MF-HPLC QTOF; massIVE ID: MSV000083737) and the data obtained in the present study (AF-HPLC HR-MS; massIVE ID: MSV000088034) to a standard bottom-up proteomics data analysis workflow as described in Belghit et al. (2019), with the exception that the Comet search engine in TPP was used instead of X! Tandem in proteoQC. The spectra identification output of the Comet search engine of both the datasets (MF-HPLC QTOF and AF-HPLC HR-MS) is given in Table 1. The results showed that with the exception of one species (BSF), the number of PSMs, peptides, and proteins were twice as high when running MF-HPLC QTOF based analysis workflow compared to the newly developed AF-HPLC HR-MS -based approach., Approximatively, the same number of PSMs, peptides, and proteins were detected with the MF-HPLC QTOF and AF-HPLC HR-MS for other insect species (YM, HC, LW, and MW, Table 1). Contrary to the raw number of spectra obtained, the percentage of identified spectra was consistently higher using the AF-HPLC HR-MS workflow when compared to the MF-HPLC QTOF workflow. A total of 30% more spectra were identified for BSF, YM, and HC samples when using the AF-HPLC HR-MS workflow (Table 1).

In summary, based on the bioinformatic analysis of the insect samples data published earlier (massIVE ID: MSV000083737) and data generated in the present study (massIVE ID: MSV000088034), the results obtained, indicate that AF-HPLC HR-MS provides data of sufficient

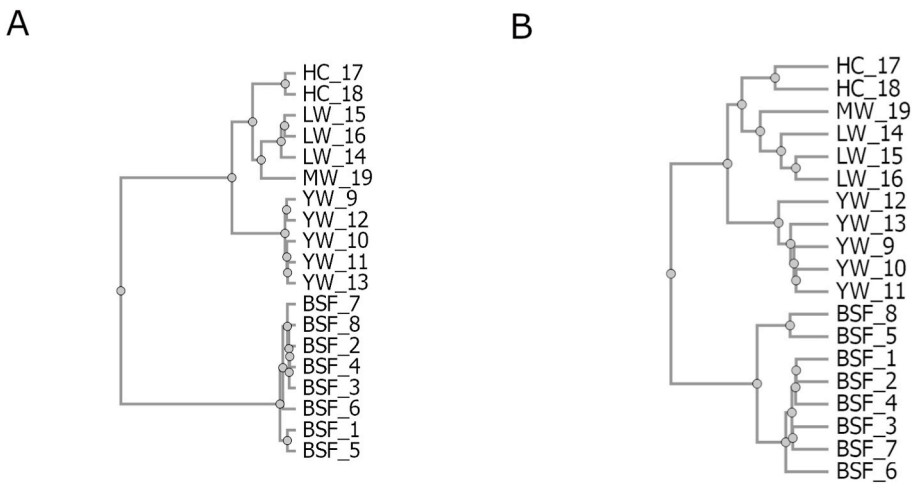

Fig. 1. Species-specific insect meal samples differentiation with direct comparison of spectra obtained by tandem mass spectrometry using compareMS2. Data obtained with (A) MF-HPLC QTOF (described in Belghit et al., 2019) and (B) the developed AF-HPLC HR-MS workflow. BSF = black soldier fly; YM = yellow mealworm; LM = lesser mealworm; HC = house cricket; MW = morio worm.

Table 1
Total numbers of spectra, identified proteins, and peptides using Comet search engine from 19 insect meal samples.

Species	MF-HPLC QTOF Belghit et al. (2019)					AF-HPLC HR-MS (newly developed)				
	tSpectra	PSM	Peptides	Proteins	% id	tSpectra	PSM	Peptides	Proteins	% id
BSF1	28176	16927	16860	11158	60%	9656	8656	8530	5838	90%
BSF2	28497	16857	16761	10817	59%	9089	8101	7960	5314	89%
BSF3	27201	15133	15049	10034	56%	10117	8901	5724	4485	88%
BSF4	28151	17011	16899	10903	60%	10049	9000	8812	5729	90%
BSF5	21910	12672	12616	9010	58%	9272	8484	8427	6047	92%
BSF6	22043	12595	12525	8705	57%	9811	8905	8796	5897	91%
BSF7	25050	12663	12583	8758	51%	10105	9019	8858	5647	89%
BSF8	28171	16283	16199	10677	58%	8749	8041	8015	5993	92%
YW9	30051	6927	6644	899	23%	10171	7228	6989	900	71%
YW10	26590	10490	9944	960	39%	10735	7654	7411	900	71%
YW11	28145	11637	10991	972	41%	10403	7472	7244	910	72%
YW12	26888	9509	9075	938	35%	9190	6263	6110	872	68%
YW13	29566	12470	11770	985	42%	10316	7426	7206	909	72%
LW14	27434	680	570	92	2%	9908	412	387	73	4%
LW15	24166	564	496	83	2%	10283	439	404	72	4%
LW16	25740	566	494	82	2%	10278	408	378	70	4%
HC17	26423	14085	13881	5060	53%	10620	9657	9556	4271	91%
HC18	24762	12507	12358	4851	51%	10121	9274	9172	4196	92%
MW19	24044	416	369	13	2%	10258	368	330	13	4%

* tSpectra - total spectra in the file; PSM– protein spectra matches; Peptides – number of identified peptides; Proteins – number of identified proteins; % id – percentage of number of identified spectra divided by total number of spectra; BSF – black soldier fly; YW – yellow mealworm; LW – lesser mealworm; HC– house cricket; MW – Morio Worms (data used from QTOF instrument Belghit et al. (2019) and HR-MS instrument).

quality to perform non-targeted species identifications of insects intended for use in food and feed. Having established that the performance of the AF-HPLC HR-MS workflow established in this study is in line with previously published assays developed for the untargeted feed- and food authenticity analyses (Belghit et al., 2019), in the next step, we assessed if this approach also is suitable for the targeted identification of insect samples using spectral library matching (SLM) and insect-specific marker peptides and marker proteins, respectively.

For creating the SLs, both MF-HPLC QTOF and AF-HPLC HR-MS data were used; each library contained an average of 12,617 spectra (MF-HPLC QTOF workflow) and 9433 spectra (AF-HPLC HR-MS workflow). Samples from both datasets were matched against these insect spectra reference libraries (cross-matching). After spectra matching of the samples to both libraries, it was found that the best matching spectra originated from samples of the same insect species as the respective library (Supplementary Table 4). As previously shown for mammals and fish (Nessen et al., 2016; Ohana et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013), the spectra library against which the highest number of matching spectra are acquired can be used to determine the identity of the samples (Supplementary Table 4). In both datasets (MF-HPLC QTOF and AF-HPLC HR-MS), BSF libraries yielded the highest number of spectra when matching against spectra from BSF samples (Fig. 2A and B and Supplementary Table 4). Similarly, for HC, LW, and YM libraries, the best matches were obtained from HC, LW and YM samples, respectively in both datasets (Fig. 2A and B and Supplementary Table 4). Surprisingly, all LW samples showed relatively high spectral hits against the YW library and *vice versa*; this could be explained by the relatedness of the insect species belonging to the same order and family (Coleoptera-Tenebrionidae). A single MW sample included in the presented study had relatively low hits against any of the other libraries, with the most hits against the LW library; this could be explained in parts by the "phylogeny" obtained by compareMS2 in which the MW-19 sample clustered closely with other LW samples (Fig. 1A and B). When working with the detection of closely related fish-species, in mixes, we found that using SLM, it was difficult to distinguish cod and haddock, which both are members of the Gadidae family (Varunjikar et al., 2022). We

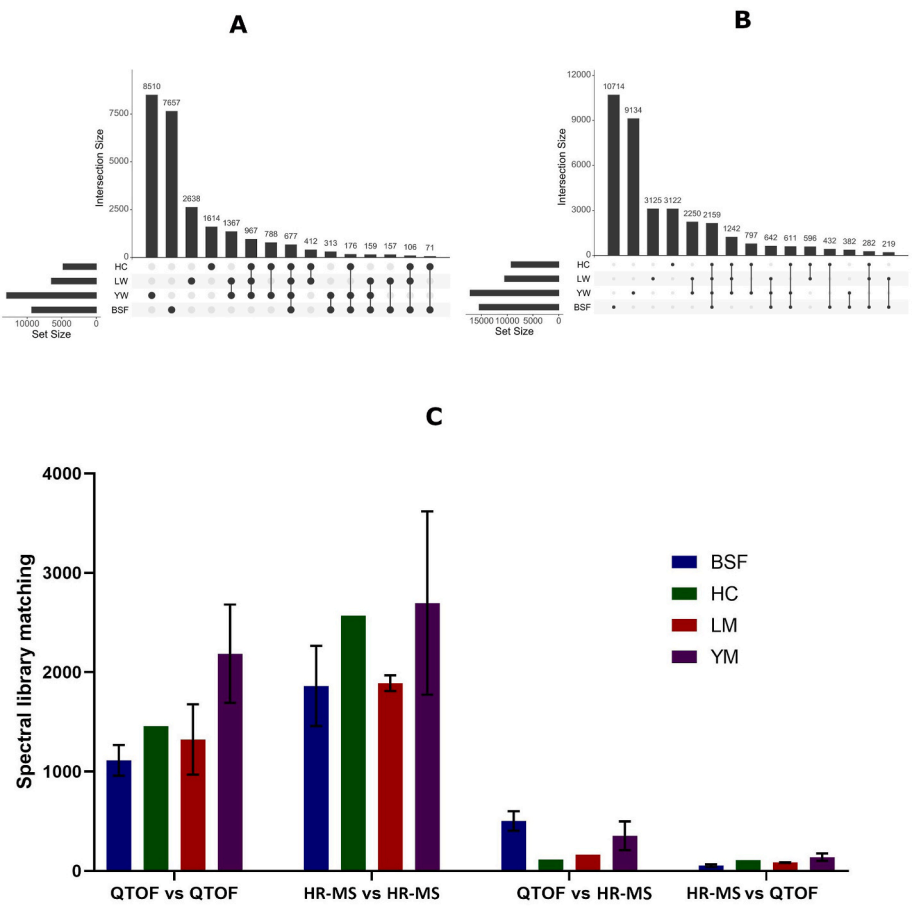

Fig. 2. (A) SpectraST output of library matching indicating MF-HPLC QTOF spectra-specific to insect species BSF, YW, LW, and HC. The detected species-specific spectra can be potential markers for species identification using the untargeted approach. Note: for Morio warm SL was created but to further evaluate species-specific marker additional samples were not available. (B) SpectraST output of library matching indicating HR-MS spectra specific to insect species BSF, YW, LW, and HC. The detected species-specific spectra can be potential markers for species identification using the untargeted approach. Note: for Morio warm SL was created but to further evaluate species-specific marker additional samples were not available. (C) Average number of SLs matching for each analysis; n = 8, 1, 4, and 5 for BSF, HC, LM, and YM, respectively. QTOF vs QTOF: data collected from QTOF and library created on QTOF; HR-MS vs HR-MS: data collected from HR-MS and library created on HR-MS; QTOF vs HR-MS: data collected from QTOF and library created on HR-MS; HR-MS vs QTOF: data collected from HR-MS and library created on QTOF.

therefore speculate that when using SLM, also for closely related insect species from e.g., the Coleoptera-Tenebrionidae family (i.e., YW, LW, and MW), this might be the case in mixed samples.

The compatibility of MF-HPLC QTOF and AF-HPLC HR-MS for building SLs and matching was evaluated by matching acquired data (Fig. 2C). The results of SL matching indicated that the highest number of matches (10–20%) to the SLs were acquired when the libraries were built on the same MS instrument as the query sample. An overview of the spectral matching in Fig. 2C, suggests that the higher number of spectral hits were reported when libraries were built on HR-MS and query data were run on QTOF instrument compared to libraries built on QTOF and queries ran on HR-MS matching. Overall, these findings are consistent with previous work performed on flatfish and other fish species where the highest match was with respective species and closely related species (Nessen et al., 2016; Ohana et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013).

Taken together our data indicate that SL created based on data obtained on MF and AF -HPLC coupled to HR-MS or QTOF instruments can be used for the detection and identification of insect species in food and feed mixtures. The data underlying the analyses presented here were made publicly available on massIVE (massIVE ID: MSV000083737 and massIVE ID: MSV000088034) and can in future be further tested with a larger number of samples for evaluating the robustness of the method.

3.3. Insect protein identification and marker detection

In addition to the spectra matching approaches presented in the previous section, in the present study, we also performed a classic reference proteome dependent bottom-up proteomics data analysis. While this approach is commonly used, is important to note that to date only a few insect-specific reference proteomes exist in public databases and most of the entries are unreviewed (Table 1). Especially for BSF, the UniProtKB database comprises exclusively of unreviewed sequences (1 reviewed and 17,593 unreviewed sequences, accessed on July 2021) and protein identifications at this moment in time might not be very precise. Due to these challenges, the SL-based approach presented above would be beneficial for insect species identification, as previously proposed (Belghit et al., 2019). Regardless, for a comparison with other insect focused studies in the literature, in addition to the SL-based insect identification, we also performed a classic protein identification analysis using both the previously published MF-HPLC QTOF data (massIVE ID: MSV000083737) and the AF-HPLC HR-MS data created in the present study (massIVE ID: MSV000088034).

Originally, a proteoQC based workflow was used to analyse MF-HPLC QTOF data (massIVE ID: MSV000083737). As was the case in the present study, the analyses by (Belghit et al., 2019) also revealed that the rate of protein identification is directly dependent on the size of the UniProtKB database for the insect species in question. For protein identification and species-specific marker detection spectra acquired from both MF-HPLC QTOF and AF-HPLC HR-MS workflows were searched against reviewed sequences from all species of arthropods. A similar approach has been used previously for analyses of non-model species whose reference proteomes are incomplete or not yet available (Francis et al., 2020; Nessen et al., 2016; Varunjikar et al., 2022; Wulff et al., 2013).

Re-analyses of the MF-HPLC QTOF insect dataset (massIVE ID: MSV000083737) generated by Belghit et al. (2019) identified 4745 proteins. The AF-HPLC HR-MS data generated in the present study (massIVE ID: MSV000088034) yielded 4147 protein identifications suggesting that the AF-HPLC HR-MS setup established here, yields result comparable to those obtained previously (Fig. 3A). While comparable in relation to the total number of proteins identified, further analysis of the

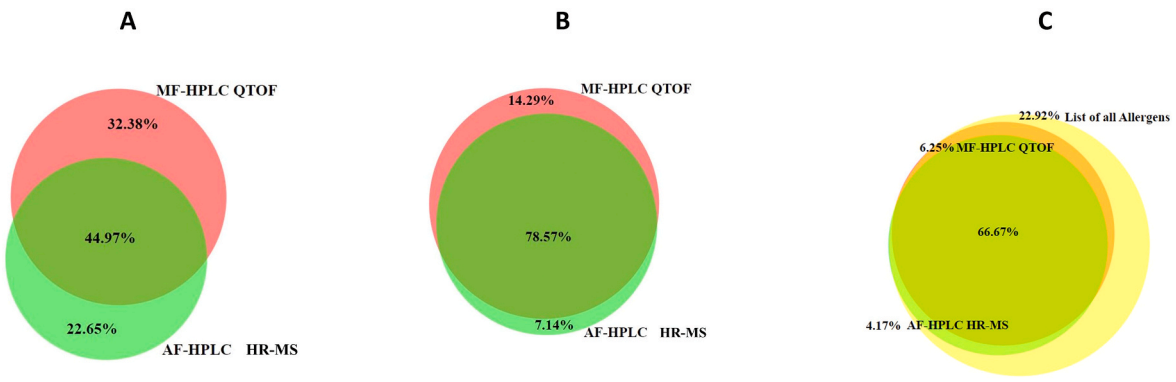

Fig. 3. (A) Insect protein identification; Venn diagrams comparing the percentages of proteins detected in 19 insect meal sample using AF-HPLC HR-MS and MF-HPLC QTOF. (B) Insect marker detection; Venn diagrams comparing the percentages of species-specific proteins detected in 19 insect meal sample using AF-HPLC HR-MS and MF-HPLC TOF (78% proteins were common in both the dataset) (C) Insect allergen detection; Venn diagrams comparing the number of allergens detected in 19 insect meal sample using MF-HPLC QTOF and AF-HPLC HR-MS workflows. Heatmaps illustrating the allergens identification using.

protein data revealed that less than half of the identified proteins (a total of 2758; ~45%) were consistently detected in both datasets; 1986 (32%) and 1389 (22%) proteins were specific to the MF-HPLC QTOF and AF-HPLC HR-MS datasets, respectively (Fig. 3A). A possible reason for the observed difference in protein identification between the two sample analysis workflows can, as was shown previously (Kalli et al., 2014; Rasinger et al., 2016), be the different type of instrument or the different HPLC and MS parameters used. Furthermore, protein extraction protocols also have been shown to affect proteomic profile descriptions (Belghit et al., 2019; Bose et al., 2021; Marbaix et al., 2016; Rasinger et al., 2016). Therefore, to minimize effects of sample preparation, instrument and analysis settings, for future MS-based analyses and differentiation of insects in feed and food, standardized procedures should be established and ideally, be made available in standard operating procedures (SOP) as is the case for example, for the qPCR-based analyses of processed animal proteins (PAP) (European Commission, 2013/51; Olsvik et al., 2017).

Following protein identification, AF-HPLC HR-MS and MF-HPLC QTOF data were compared on species levels (Fig. 4A and B). The results show that samples from the same insect species were grouped together in hierarchical clustering analyses which were performed on MS data passing statistical significance thresholds in a grouped comparison analysis (Qlucore Omics Explorer, $q < 0.1$, Supplementary Table 5). Most of the samples from the Coleoptera family were grouped in the heatmap except for two YW samples analysed using HR-MS workflow; unlike in the compareMS2 output, these were placed on a separate branch of the heatmap (Fig. 4B). Some insect samples used in this study were defatted and processed differently which could have affected protein extractions and protein inference. The heatmap shown in Fig. 4A and B also suggest that there were ~19 proteins with high expression levels in BSF samples when compared to other samples (i.e., LW, YW, MW, and HC). Also, from the Coleoptera family, 21 proteins were displaying different expression levels in YW, LW, and MW compared to BSF. A possible explanation for the overrepresentation of BSF specific proteins could be that the database used for spectra peptide matching and protein inference comprises Arthropoda protein sequences which are dominated by *Drosophila melanogaster* (fruit fly) entries. The latter belongs to the same order as BSF (i.e., Diptera) and therefore the protein matches might be higher; this also was observed in a study by Francis et al., 2020 where proteomics analyses were used for edible insect fingerprinting in novel food. .

To mine for potential species-specific marker proteins for the detection of insects in food and feed, we focused on proteins consistently detected in both AF-HPLC HR-MS and MF-HPLC QTOF data (Fig. 3B). The analysis of the AF-HPLC HR-MS and MF-HPLC QTOF data suggests that for YW, the larval cuticle protein A2B could be a potential marker for species identification (Supplementary Figs. 4A and B; Supplementary Table 6). For HC, cytochrome *c* oxidase (mitochondrial) could be a potential marker protein for species identification given that it was detected only in HC samples (i.e. HC-17 in MF-HPLC QTOF data and both HC-17 and 18 in AF-HPLC HR-MS data) (Supplementary Figs. 4A and B). While further analyses are warranted to confirm that the proteins described here indeed are species-specific, the data provided in this study can be used as the basis to explore the development of quantitative standard reaction monitoring (SRM) assays for the species-specific identification of insects in food and feed as recently demonstrated for PAP identification in animal feed (Lecrenier et al., 2021; Marbaix et al., 2016; Steinhilber et al., 2018a,b; 2019). This work could complement efforts recently reported in a study using a peptidomics approach based on a combination of high-resolution untargeted and targeted species-specific markers for BSF and LM (Leni, Prandi, et al., 2020).

3.4. Detecting allergen in insect species of interest

In addition to the eight insect species permitted to be used as PAP in feed (European Commission, 2017/893), the European Commission recently authorised the marketing of dried yellow mealworms for human consumption (European Comission, 2021/882) and a favourable opinion the placing on the market of house cricket (*Acheta domesticus*) as a novel food was issued by EFSA (EFSA NDA panel, 2021b). Concerning the consumption of the house cricket, EFSA identified no other safety concerns than allergenicity and in a recent review on edible insects and food safety, it was highlighted that extensive allergenic risk assessments would be required before the safe introduction edible insects in the food market were (Ribeiro et al., 2021). In the light of the potential allergenic risk insects mays pose, it was assessed in the present study if untargeted proteomics data acquired from both MF-HPLC QTOF and AF-HPLC HR-MS also can be successfully screened for the presence of relevant known food allergens (Supplementary Table 7).

From the list of 48 allergenic proteins, 37 were detected in both datasets and 32 were consistently detected in both MF-HPLC QTOF and AF-HPLC HR-MS data (Fig. 3C). Using a proteomic and bioinformatic approach, Barre et al., 2021 identified a comparable number of pan-allergens (46 proteins) in house crickets (*Acheta domesticus*) (EFSA NDA panel, 2021b). Among the four families of allergens in silk moth (*Bombyx mori*) which is a close relative of the selected insect species in this study, arginine kinase (Q2F5T5), low molecular mass lipoproteins (Q00802 and Q00801), and tropomyosin 1 and 2 (Q1HPU0 and Q1HPQ0) proteins were detected in the acquired data from both instruments. Tropomyosin is a known IgE-binding protein and cross-reactivity of HC tropomyosin with shrimp tropomyosin was demonstrated with ELISA in a recent study (De Marchi, Wangorsch, &

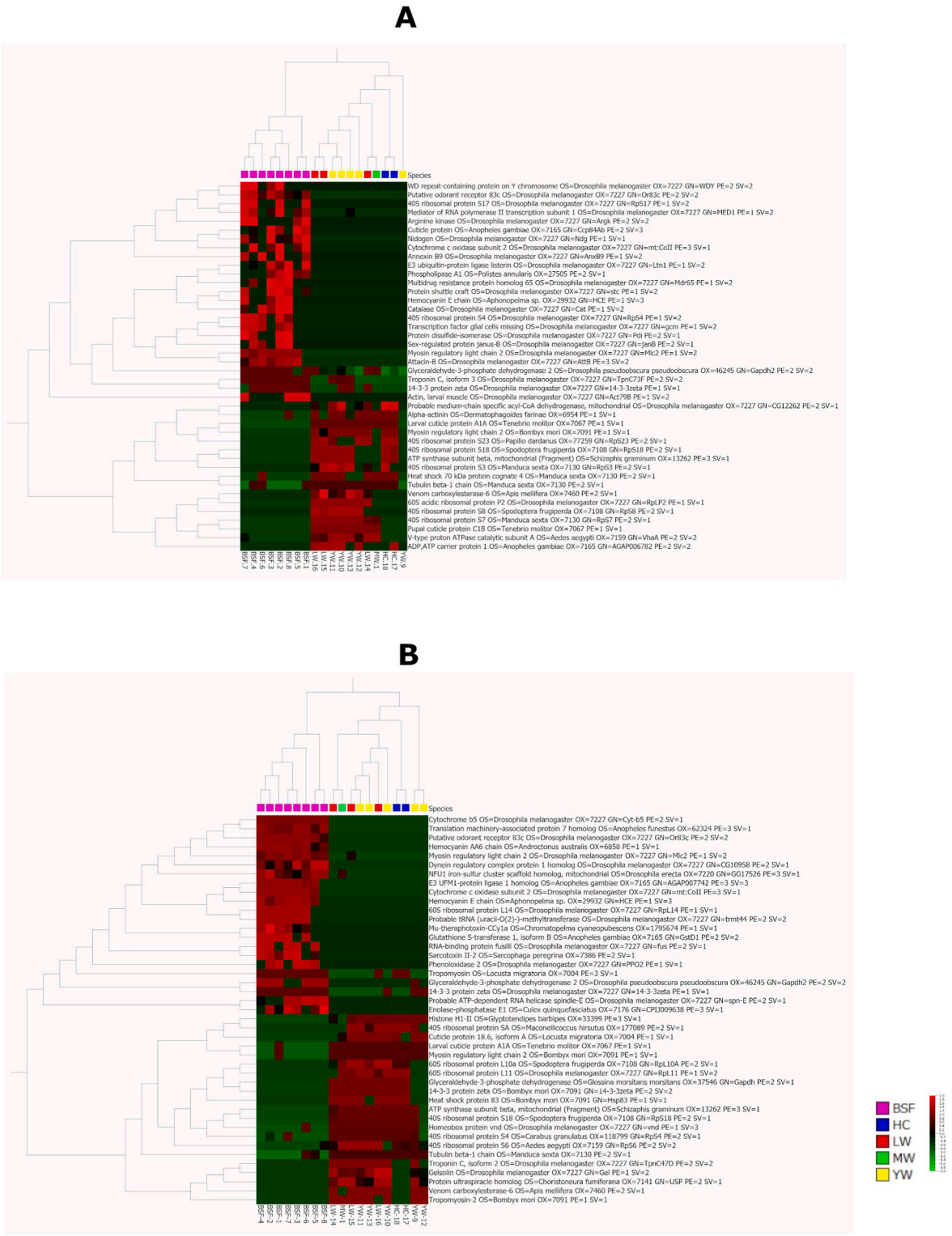

Fig. 4. Insect protein identification; Heatmaps illustrating the protein identification using (A) MF-HPLC QTOF and (B) AF-HPLC HR-MS workflows, based on TPP identification using Comet search engine and Arthropoda reviewed protein as reference database. Hierarchical clustering (HC) of samples and differentially expressed proteins where group comparison was performed using Omics Explorer V3.6. The heatmap represents expressed proteins withing each measured samples; red represents expressed proteins and green represent absent or unexpressed proteins. Note that the proteins might not be from the same species as studied given that most of the proteins for the species of interest in this study were unreviewed. So, they are from different insect species but exhibit similarity to the species BSF, YW, LW, HC, and MW. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

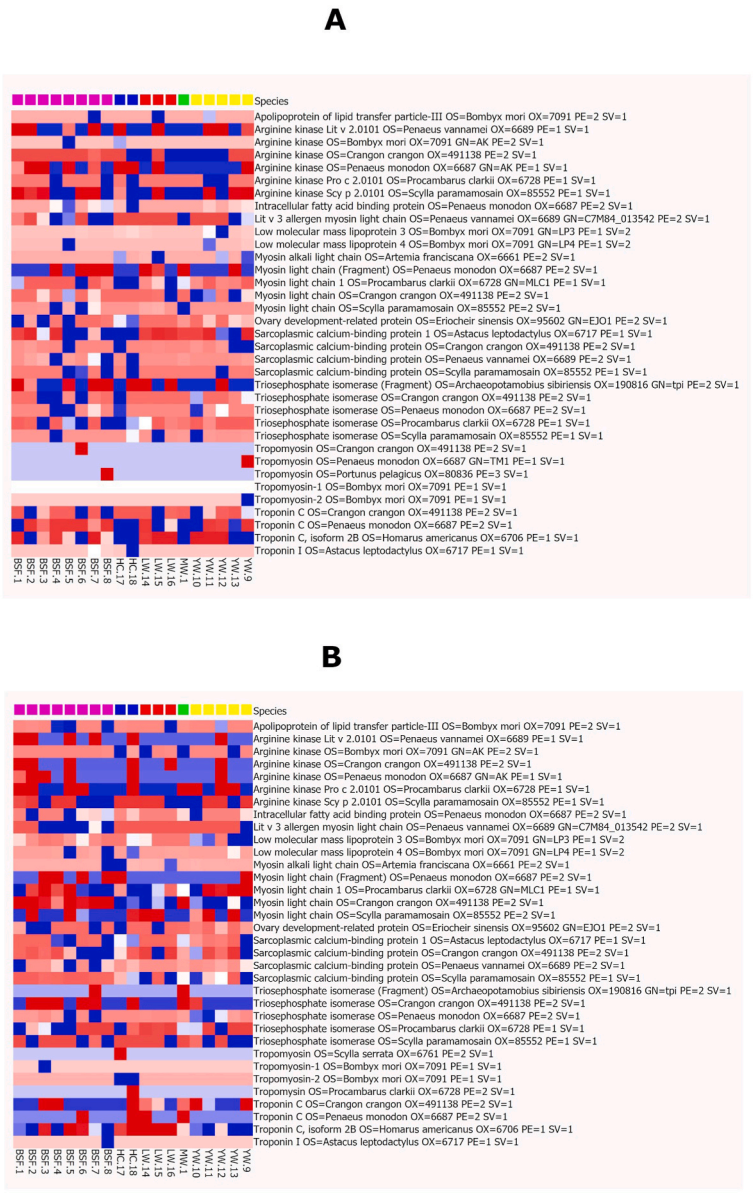

Fig. 5. Insect allergen detection. Heatmaps illustrating the allergens identification using (A) MF-HPLC QTOF and (B) AF-HPLC HR-MS workflows. Heat map representation of 37 allergens across the 19 insect samples. As explained in the insert the pink, blue, red, green, and yellow rectangles represent BSF, HC, LW, MW and YM, respectively. Each line in the heat map represents an allergen. The deeper red colour, the higher is the allergen present in the respective sample; similarly, the deeper the blue colour, the lower is the allergen present in the respective sample as illustrated in the figure insert. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Zoccatelli, 2021). In the current study, tropomyosin-2 from silk moth (Q1HPQ0) was consistently detected in both datasets (Fig. 5A and B) across all species with one exception; in HC and a single replicate of YW (YW-9) samples, Q1HPQ0 was detected only in MF-HPLC QTOF and AF-HPLC HR-MS data, respectively. Interestingly, tropomyosin also was flagged as a key pan-allergen present in the house cricket when highlighting safety concerns related to the consumption of this novel food (EFSA NDA panel, 2021b). Other allergenic proteins detected in the insect samples were arginine kinase and troponin C (from different Arthropoda species) that were present in BSF, HC, LW and YW samples. In a recent study focusing on arginine kinase (Bose et al., 2021), it was shown that protein extraction protocols can affect the quantitation of allergens from cricket samples. It could therefore be possible that the varying profile of allergens detected in the selected insect samples presented here can be attributed to differences in sample processing, instrument selection, and protein extraction protocol (Broekman et al., 2015; De Marchi, Wangorsch, & Zoccatelli, 2021; Pali-Schöll et al., 2019; Van Broekhoven et al., 2016). In other words, like proteomics-based marker detection for insect species differentiation in food and feed, also allergen detection could benefit from standardized procedures summarized in SOPs for the respective purpose (Bose et al., 2021; Marbaix et al., 2016).

The tentative screening for predicted allergens in data obtained from basic MF- and AF- HPLC HR-MS workflows commonly used in regulatory laboratories highlighted the potential of these routine tools for ensuring the safety of novel foods and feeds. What is more, the data created here (massIVE ID: MSV000088034) and by Belghit et al. (2019) (massIVE ID: MSV000083737), lays the foundation for future work focusing on spectra matching in which SL and *in-silico* assessments can be combined for allergen detection as recently exemplified by (FitzGerald et al., 2020; Leni et al., 2020).

4. Conclusion

The combination of standard MS instruments commonly available in regulatory laboratories combined with freely available open-source data analysis approaches allow for implementation of untargeted proteomics assays for food and feed safety research in routine settings. The AF-HPLC HR-MS workflow and associated bioinformatics approaches presented here can be a useful toolset suitable for the detection and differentiation

of insects in feed and food and complement existing methods currently used in the market. The approaches presented and the data generated in the present study and made available in a public repository (massIVE ID: MSV000088034) also were found to be suitable for allergen detection in insect species.

CRediT authorship contribution statement

Madhushri S. Varunjikar: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing –review & original draft. **Ikram Belghit:** Conceptualization, Supervision, Investigation, Methodology, Project administration, Writing – original draft, Writing –review & original draft. **Jennifer Gjerde:** Conceptualization, Investigation, Methodology, review & original draft. **Magnus Palmblad:** Conceptualization, Data curation, Investigation, Methodology, review &original draft. **Eystein Oveland:** Conceptualization, Methodology, Data curation, Investigation, Writing –review & original draft. **Josef D. Rasinger:** Conceptualization, Supervision, Data curation, Investigation, Methodology, Project administration, Software, Writing –review & original draft.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

We are very much thankful for support and guidance of Dr. Marc Dieu from University of Namur, mass spectrometry facility (MaSUN), rue de Bruxelles 61, B-5000 Namur, Belgium and for the compareMS2 graphical user interface by Rob Marissen (CPM). The contents of the graphical abstract were created using BioRender.com.

Funding was provided by the Institute of Marine Research (Programledelse Fiskeernæring, MultiOmicsTools project).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodcont.2022.108888.

References

Aguilera, J., Aguilera-Gomez, M., Barrucci, F., Cocconcelli, P. S., Davies, H., Denslow, N., Lou Dorne, J., Grohmann, L., Herman, L., Hogstrand, C., Kass, G. E. N., Kille, P., Kleter, G., Nogué, F., Plant, N. J., Ramon, M., Schoonjans, R., Waigmann, E., & Wright, M. C. (2018). EFSA scientific colloquium 24 – 'omics in risk assessment: State of the art and next steps. *EFSA Supporting Publications, 15*(11). https://doi.org/10.2903/sp.efsa.2018.EN-1512

Barre, A., Pichereaux, C., Simplicien, M., Burlet-Schiltz, O., Benoist, H., & Rougé, P. (2021). A proteomic-and bioinformatic-based identification of specific allergens from edible insects: Probes for future detection as food ingredients. https://doi.org/10.3390/foods10020280.

Belghit, I., Lock, E. J., Fumière, O., Lecrenier, M. C., Renard, P., Dieu, M., Berntssen, M. H. G., Palmblad, M., & Rasinger, J. D. (2019). Species-specific discrimination of insect meals for aquafeeds by direct comparison of tandem mass spectra. *Animals, 9*(5). https://doi.org/10.3390/ani9050222

Belghit, I., Varunjikar, M., Lecrenier, M.-C., Steinhilber, A. E., Niedzwiecka, A., Wang, Y. V., Dieu, M., Azzollini, D., Lie, K., Lock, E.-J., Berntssen, M. H. G., Renard, P., Zagon, J., Fumière, O., van Loon, J. J. A., Larsen, T., Poetz, O., Braeuning, A., Palmblad, M., & Rasinger, J. D. (2021). Future feed control – tracing banned bovine material in insect meal. *Food Control, 128*, 108183. https://doi.org/10.1016/j.foodcont.2021.108183

Bose, U., Broadbent, J. A., Juhász, A., Karnaneedi, S., Johnston, E. B., Stockwell, S., Byrne, K., Limviphuvadh, V., Maurer-Stroh, S., Lopata, A. L., & Colgrave, M. L. (2021). Protein extraction protocols for optimal proteome measurement and arginine kinase quantitation from cricket Acheta domesticus for food safety assessment. *Food Chemistry, 348*, 129110. https://doi.org/10.1016/J.FOODCHEM.2021.129110

Broekman, H., Knulst, A., den Hartog Jager, S., Monteleone, F., Gaspari, M., de Jong, G., Houben, G., & Verhoeckx, K. (2015). Effect of thermal processing on mealworm allergenicity. *Molecular Nutrition & Food Research, 59*, 1855–1864. https://doi.org/10.1002/mnfr.201500138

European Commission Regulation 2021/1372. (2021a). Commission Regulation (EU) 2021/1372 of 17 August 2021 amending Annex IV to Regulation (EC) No 999/2001 of the European Parliament and of the Council as regards the prohibition to feed non-ruminant farmed animals, other than Fur animals, with protein deri. *Official Journal of the European Union, 64*, 1–21. http://data.europa.eu/eli/reg/2021/1372/oj.

European Commission Regulation 2021/882. (2021b). Commission implementing regulation (EU) 2021/882 of 1 June 2021 authorising the placing on the market of dried Tenebrio molitor larva as a novel food under regulation (EU) 2015/2283 of the European parliament and of the council, and amending commission im. *Official Journal of the European Union, 64*(L194), 16–20. http://data.europa.eu/eli/reg_impl/2021/882/oj.

compareMS2 GUI. https://github.com/524D/compareMS2, (2021).

Daniso, E., Tulli, F., Cardinaletti, G., Cerri, R., & Tibaldi, E. (2020). Molecular approach for insect detection in feed and food: The case of gryllodes sigillatus. *European Food Research and Technology, 246*(12), 2373–2381. https://doi.org/10.1007/S00217-020-03573-1

De Marchi, L., Mainente, F., Leonardi, M., Scheurer, S., Wangorsch, A., Mahler, V., Pilolli, R., Sorio, D., & Zoccatelli, G. (2021). Allergenicity assessment of the edible cricket Acheta domesticus in terms of thermal and gastrointestinal processing and IgE cross-reactivity with shrimp. *Food Chemistry, 359*, 129878. https://doi.org/10.1016/J.FOODCHEM.2021.129878

De Marchi, L., Wangorsch, A., & Zoccatelli, G. (2021). Allergens from edible insects: Cross-reactivity and effects of processing. https://doi.org/10.1007/s11882-021-01012-z/.

Debode, F., Marien, A., Gérard, A., Francis, F., Fumière, O., & Berben, G. (2017). Development of real-time PCR tests for the detection of Tenebrio molitor in food and feed. *Food Additives & Contaminants: Part A, 34*(8), 1421–1426. https://doi.org/10.1080/19440049.2017.1320811

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

EFSA NDA panel. (2021). Safety of dried yellow mealworm (Tenebrio molitor larva) as a novel food pursuant to Regulation (EU) 2015/2283. *EFSA Journal, 19*(1), 1–29. https://doi.org/10.2903/j.efsa.2021.6343

EFSA NDA panel. (2021b). Safety of frozen and dried formulations from whole house crickets (Acheta domesticus) as a Novel food pursuant to Regulation (EU) 2015/2283. *EFSA Journal, 19*(8). https://doi.org/10.2903/j.efsa.2021.6779

European Commission. (2013). COMMISSION REGULATION (EU) No 51/2013 of 16 January 2013 amending Regulation (EC) No 152/2009 as regards the methods of analysis for the determination of constituents of animal origin for the official control of feed. *Official Journal of the European Union*, 33–43, 2013, L 20/33) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R0051&from=EN.

European Commission. (2017). *Commission regulation (EU) 2017/893* (pp. 1–25). Official Journal of the European Union, 2017 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0893&rid=1.

Fao, I. F. A. D., Unicef, W. F. P., & Who. (2018). The state of food security and nutrition in the world 2018. Building climate resilience for food security and nutrition. Rome, FAO. Licence: CC BY-NC-SA 3.0 IGO. In *Building climate resilience for food security and nutrition*. https://doi.org/10.1093/cjres/rst006

FitzGerald, R. J., Cermeño, M., Khalesi, M., Kleekayai, T., & Amigo-Benavent, M. (2020). Application of in silico approaches for the generation of milk protein-derived bioactive peptides. *Journal of Functional Foods, 64*, 103636. https://doi.org/10.1016/J.JFF.2019.103636

Francis, F., Doyen, V., Debaugnies, F., Mazzucchelli, G., Caparros, R., Alabi, T., Blecker, C., Haubruge, E., & Corazza, F. (2019). Limited cross reactivity among arginine kinase allergens from mealworm and cricket edible insects. *Food Chemistry, 276*, 714–718. https://doi.org/10.1016/J.FOODCHEM.2018.10.082

Francis, F., Mazzucchelli, G., Baiwir, D., Debode, F., Berben, G., & Caparros Megido, R. (2020). Proteomics based approach for edible insect fingerprinting in novel food: Differential efficiency according to selected model species. *Food Control, 112*, 107135. https://doi.org/10.1016/J.FOODCONT.2020.107135

Garino, C., Zagon, J., & Nesic, K. (2021). Novel real-time PCR protocol for the detection of house cricket (Acheta domesticus) in feed. *Animal Feed Science and Technology, 280*, 115057. https://doi.org/10.1016/J.ANIFEEDSCI.2021.115057

IPIFF. (2021). An overview of the European market of insects as feed. 2020–2021. https://ipiff.org/.

Jenkins, C., & Orsburn, B. (2020). *Standard flow multiplexed proteomics (SFloMPro) – an accessible and cost-effective alternative to NanoLC workflows*. BioRxiv. https://doi.org/10.1101/2020.02.25.964379, 2020, 02.25.964379.

Kalli, A., Smith, G. T., Sweredoski, M. J., & Hess, S. (2014). Evaluation and optimization of mass spectrometric mode: Focus on LTQ-Orbitrap Mass analyzers. *Journal of Proteome Research, 12*(7), 3071–3086. https://doi.org/10.1021/pr3011588.Evaluation

Keller, A., Nesvizhskii, A. I., Kolker, E., & Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry, 74*(20), 5383–5392. https://doi.org/10.1021/ac025747h

Kelstrup, C. D., Jersie-Christensen, R. R., Batth, T. S., Arrey, T. N., Kuehn, A., Kellmann, M., & Olsen, J. V. (2014). Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field orbitrap mass spectrometer. *Journal of Proteome Research, 13*(12), 6187–6195. https://doi.org/10.1021/pr500985w

Köppel, R., Schum, · R.afael, Habermacher, · M., Sester, C., Lucia, ·, Piller, E., Meissner, · S., & Pietsch, K. (2019). Multiplex real-time PCR for the detection of insect DNA and

determination of contents of Tenebrio molitor, Locusta migratoria and Achaeta domestica in food. https://doi.org/10.1007/s00217-018-03225-5, 245, 559-567.
Lecrenier, Caroline, M., Marien, A., Veys, P., Belghit, I., Dieu, M., Gillard, N., Henrottin, J., Herfurth, U. M., Marchis, D., Morello, S., Oveland, E., Poetz, O., Rasinger, J. D., Steinhilber, A., Baeten, V., Berben, G., & Fumière, O. (2021). Inter-laboratory study on the detection of bovine processed animal protein in feed by LC-MS/MS-based proteomics. *Food Control, 125*(November 2020), 1–7. https://doi.org/10.1016/j.foodcont.2021.107944
Lecrenier, M. C., Planque, M., Dieu, M., Veys, P., Saegerman, C., Gillard, N., & Baeten, V. (2018). A mass spectrometry method for sensitive, specific and simultaneous detection of bovine blood meal, blood products and milk products in compound feed. *Food Chemistry, 245*(September), 981–988. https://doi.org/10.1016/j.foodchem.2017.11.074
Lenčo, J., Vajrychová, M., Pimková, K., Prokšová, M., Benková, M., Klimentová, J., Tambor, V., & Soukup, O. (2018). Conventional-flow liquid chromatography-mass spectrometry for exploratory bottom-up proteomic analyses. *Analytical Chemistry, 90*(8), 5381–5389. https://doi.org/10.1021/acs.analchem.8b00525
Leni, G., Prandi, B., Varani, M., Faccini, A., Caligiani, A., & Sforza, S. (2020). Peptide fingerprinting of Hermetia illucens and Alphitobius diaperinus: Identification of insect species-specific marker peptides for authentication in food and feed. *Food Chemistry, 320*, 126681. https://doi.org/10.1016/J.FOODCHEM.2020.126681
Leni, G., tedeschi, tullia, faccini, A., pratesi, federico, folli, claudia, puxeddu, ilaria, Migliorini, paola, Gianotten, natasja, Jacobs, J., Depraetere, S., caligiani, A., & Sforza, S. (2020). Shotgun proteomics, in-silico evaluation and immunoblotting assays for allergenicity assessment of lesser mealworm, black soldier fly and their protein hydrolysates. https://doi.org/10.1038/s41598-020-57863-5.
Marbaix, H., Budinger, D., Dieu, M., Fumière, O., Gillard, N., Delahaut, P., Mauro, S., & Raes, M. (2016). Identification of proteins and peptide biomarkers for detecting banned processed animal proteins (PAPs) in meat and bone meal by mass spectrometry. *Journal of Agricultural and Food Chemistry, 64*(11), 2405–2414. https://doi.org/10.1021/acs.jafc.6b00064
Michalski, A., Damoc, E., Hauschild, J. P., Lange, O., Wieghaus, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., & Horning, S. (2011). Mass spectrometry-based proteomics using Q exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer. *Molecular & Cellular Proteomics, 10*(9), 1–11. https://doi.org/10.1074/mcp.M111.011015
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
Olsvik, P. A., Fumière, O., Margry, R. J. C. F., Berben, G., Larsen, N., Alm, M., & Berntssen, M. H. G. (2017). Multi-laboratory evaluation of a PCR method for detection of ruminant DNA in commercial processed animal proteins. *Food Control, 73*, 140–146. https://doi.org/10.1016/j.foodcont.2016.07.041
Pali-Schöll, I., Meinlschmidt, P., Larenas-Linnemann, D., Purschke, B., Hofstetter, G., Rodríguez-Monroy, F. A., Einhorn, L., Mothes-Luksch, N., Jensen-Jarolim, E., & Jäger, H. (2019). Edible insects: Cross-recognition of IgE from crustacean- and house dust mite allergic patients, and reduction of allergenicity by food processing. *World Allergy Organization Journal, 12*(1), 100006. https://doi.org/10.1016/J.WAOJOU.2018.10.001
Palmblad, M., & Deelder, A. M. (2012). Molecular phylogenetics by direct comparison of tandem mass spectra. *Rapid Communications in Mass Spectrometry, 26*(7), 728–732. https://doi.org/10.1002/rcm.6162
Pedersen. (2021). MSGFplus: An interface between R and MS-GF+. https://doi.org/10.18129/B9.bioc.MSGFplus.
Rasinger, J. D., Marbaix, H., Dieu, M., Fumière, O., Mauro, S., Palmblad, M., Raes, M., & Berntssen, M. H. G. (2016). Species and tissues specific differentiation of processed animal proteins in aquafeeds using proteomics tools. *Journal of Proteomics, 147*, 125–131. https://doi.org/10.1016/j.jprot.2016.05.036
Ribeiro, J. C., Sousa-Pinto, B., Fonseca, J., Fonseca, S. C., & Cunha, L. M. (2021). Edible insects and food safety: Allergy. *Journal of Insects as Food and Feed, 7*(5), 833–847. https://doi.org/10.3920/jiff2020.0065
Rumbos, C. I., & Athanassiou, C. G. (2021). The superworm, zophobas morio (Coleoptera:tenebrionidae): A 'sleeping giant' in nutrient sources. *Journal of Insect Science, 21*(2). https://doi.org/10.1093/jisesa/ieab014
Sentandreu, M. A., & Sentandreu, E. (2011). Peptide biomarkers as a way to determine meat authenticity. *Meat Science, 89*(3), 280–285. https://doi.org/10.1016/J.MEATSCI.2011.04.028
Steinhilber, A. E., Schmidt, F. F., Naboulsi, W., Planatscher, H., Niedzwiecka, A., Zagon, J., Braeuning, A., Lampen, A., Joos, T. O., & Poetz, O. (2018a). Mass spectrometry-based immunoassay for the quantification of banned ruminant processed animal proteins in vegetal feeds. *Analytical Chemistry, 90*(6), 4135–4143. https://doi.org/10.1021/acs.analchem.8b00120
Steinhilber, A. E., Schmidt, F. F., Naboulsi, W., Planatscher, H., Niedzwiecka, A., Zagon, J., Braeuning, A., Lampen, A., Joos, T. O., & Poetz, O. (2018b). Species differentiation and quantification of processed animal proteins and blood products in fish feed using an 8-plex mass spectrometry-based immunoassay. *Journal of Agricultural and Food Chemistry, 66*(39), 10327–10335. https://doi.org/10.1021/acs.jafc.8b03934
Steinhilber, A. E., Schmidt, F. F., Naboulsi, W., Planatscher, H., Niedzwiecka, A., Zagon, J., Braeuning, A., Lampen, A., Joos, T. O., & Poetz, O. (2019). Application of mass spectrometry-based immunoassays for the species- A nd tissue-specific quantification of banned processed animal proteins in feeds [Research-article]. *Analytical Chemistry, 91*(6), 3902–3911. https://doi.org/10.1021/acs.analchem.8b04652
Szabó, D., Schlosser, G., Vékey, K., Drahos, L., & Révész, Á. (2021). Collision energies on QTof and Orbitrap instruments: How to make proteomics measurements comparable? *Journal of Mass Spectrometry, 56*(1), 1–12. https://doi.org/10.1002/jms.4693
Van Broekhoven, S., Bastiaan-Net, S., De Jong, N. W., & Wichers, H. J. (2016). Influence of processing and in vitro digestion on the allergic cross-reactivity of three mealworm species. *Food Chemistry, 196*, 1075–1083. https://doi.org/10.1016/J.FOODCHEM.2015.10.033
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*. https://doi.org/10.1002/rcm.7494
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*, 108417. https://doi.org/10.1016/j.foodcont.2021.108417
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