Food Chemistry 203 (2016) 28–34

Contents lists available at ScienceDirect

Food Chemistry

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

Identification of meat products by shotgun spectral matching

D. Ohana [a,*], H. Dalebout [a], R.J. Marissen [a], T. Wulff [b], J. Bergquist [c], A.M. Deelder [a], M. Palmblad [a]

[a] Center for Proteomics and Metabolomics, Leiden University Medical Center, Albinusdreef 2, 2300 ZC Leiden, The Netherlands
[b] The Novo Nordisk Foundation Center for Biosustainability (DTU Biosustain), Kogle Alle 6, DK-2970 Hørsholm, Denmark
[c] Analytical Chemistry, Department of Chemistry – Biomedical Centre and SciLifeLab, Uppsala University, Uppsala, Husargatan 3, 751 24 Uppsala, Sweden

ARTICLE INFO

Article history:
Received 4 May 2015
Received in revised form 27 January 2016
Accepted 29 January 2016
Available online 2 February 2016

Keywords:
Food products
Shotgun proteomics
Identification
Spectral libraries
Quantitation

ABSTRACT

A new method, based on shotgun spectral matching of peptide tandem mass spectra, was successfully applied to the identification of different food species. The method was demonstrated to work on raw as well as processed samples from 16 mammalian and 10 bird species by counting spectral matches to spectral libraries in a reference database with one spectral library per species. A phylogenetic tree could also be constructed directly from the spectra. Nearly all samples could be correctly identified at the species level, and 100% at the genus level. The method does not use any genomic information and unlike targeted methods, no prior knowledge of genetic variation within a genus or species is necessary.

1. Introduction

In recent years, methods for species identification and food authentication have garnered more interest throughout the food industry and among regulatory agencies. Increasing numbers of cases of adulteration in food labeling, mostly done for profit by substituting expensive ingredients with cheaper alternatives, are revealed. Particularly in the meat industry, fraudulent product labeling can greatly improve profits, leading to concern amongst consumers. Consumers rely on proper labeling so that informed choices can be made according to their health concerns, lifestyle, ethical stance, religion or other personal reasons. The existence of fraudulent labeling of food is remarkable since there are strict regulations by law for food labeling, stressing the need for robust analytical tests for continuous control to ensure regulations are followed. Several techniques to detect meat adulteration via DNA, lipid and protein analysis are developed and currently used for authentication. Detection methods include immunosassays (Ayaz, Ayaz, & Erol, 2006; Rao & Hsieh, 2007), Spectroscopy (Al-Jowder, Kemsley, & Wilson, 2002; Boyacı et al., 2014; Ellis, Broadhurst, Clarke, & Goodacre, 2005), mass spectrometry (MS) (Mamone, Picariello, Caira, Addeo, & Ferranti, 2009; Taylor, Linforth, Weir, Hutton, & Green, 1993) and PCR (Jonker, Tilburg, Hagele, & de Boer, 2008; Ulca, Balta, Cagin, & Senyuva, 2013). The DNA-based techniques using polymerase chain reaction (PCR) with subsequent readout or sequencing, and the antibody-based technique enzyme-linked immunosorbent assay (ELISA) are most frequently used for this identification. In the last decade, MS-based methods for sample analysis have gained interest and multiple techniques have been developed (Bouley, Chambon, & Picard, 2004; Montowska, Rao, Alexander, Tucker, & Barrett, 2014; Vaclavik et al., 2011; von Bargen, Dojahn, Waidelich, Humpf, & Brockmeyer, 2013). Most MS-approaches for identification and authentication are based on protein pattern profiling or on the identification of specific peptides or proteins to identify the meat species. Recently, an alternative mass spectrometry method based on spectral matching using spectral libraries in a bottom up proteomics approach was demonstrated on fish species. (Wulff, Nielsen, Deelder, Jessen, & Palmblad, 2013). The use of spectral matching with spectral libraries, an untargeted approach, has several advantages over targeted approaches since the comparison, and identification is based on the matching or non-matching of hundreds or thousands of peptides. This means that the identification is robust against single nucleotide polymorphisms (SNP's). Furthermore the use of spectral matching with spectral libraries is an open analysis method with no need for prior genomic information of any species, and is able to distinguish between a large number of species in a single, parallel, search.

Spectral libraries can not only be used for the identification but also for classification of species based on phylogenetic similarity. Such molecular phylogenetics is more commonly done by comparing DNA sequences (Hackett et al., 2008; Thomson & Shaffer, 2010), but in 2012, a new method and software were described, com-

* Corresponding author.
E-mail address: d.ohana@lumc.nl (D. Ohana).

http://dx.doi.org/10.1016/j.foodchem.2016.01.138
0308-8146/

pareMS2, based on the direct comparison of tandem mass spectra (Palmblad & Deelder, 2012). This method does $N \times N$ comparisons for N species to construct a distance matrix for molecular phylogenetic analysis, whereas for Identification, N spectral libraries are searched using the SpectraST software (Lam et al., 2007) for each query (unknown), requiring $1 \times N$ comparisons. The dot product metric used in compareMS2 is essentially the same as in SpectraST, but the two approaches address different questions, the former describing the relationships between all species and the second returning the closest match to a particular species or sample. Both methods have the advantages of using all acquired tandem mass spectra and the fact that no genomic information is needed.

2. Materials and methods

2.1. Samples

Meat samples were purchased from local butchers and supermarkets in the Netherlands and Sweden. A total of 26 different meat types were collected, of which 16 were mammalian and 10 birds. Three muscle tissue samples of ~50 mg per species were taken from each species, labeled and preserved at −80 °C. For the analysis of the raw samples, 3 samples per species were used. For 13 different species another 3 muscle tissue samples of ~50 mg were collected and heated in a frying pan, without the use of any butter or oil, until the appropriate temperature for consumption of the species was reached (80 °C for the birds species and 60 °C for the mammal species). Surface chemistry effects of cooking the meat were not taken into account as the samples were homogenized afterwards. For the quantitation of mixtures the same procedure was followed as for the raw meat samples, with the only differences that the samples contained two different meat species and that the samples were weighed. These samples contained 150 mg of raw horse and cow meat, with the ratios of 100%, 90%, 67%, 50%, 33%, 10% and 0% horse meat. An overview of the workflow is depicted in Fig. 1.

2.2. Protein extraction

Extraction of the meat proteins was done in 150 µl trifluoroethanol and 150 µl 50 mM ammonium bicarbonate. Samples were homogenized for 9 min using 0.5 mm zirconium oxide beads in an air-cooled Bullet Blender® (Next Advance Inc., Averill Park, NY) and were placed for 60 min at 37 °C to incubate. Supernatant was collected after centrifugation at 16,000g for 30 min at room temperature. The final protein concentration was determined using a micro bicinchoninic acid (BCA) protein assay kit (Thermo Fischer Scientific, product #23235).

2.3. Digestion

Three cooked and three raw samples per species, each containing ~150 µg protein in a volume of 20 µl 50 mM ammonium bicarbonate, were used for a tryptic digestion. Proteins were first reduced with 6 µl 60 mM dithiothreitol for 5 min at 95 °C and afterwards alkylated using 10 µl 100 mM iodoacetamide for an hour at room temperature in the dark. Samples were diluted a 5-fold with 50 mM ammonium bicarbonate to reduce the TFE-concentration. The proteins were digested at 37 °C overnight by Trypsin (sequencing grade, Promega, Madison, WI, USA) at an enzyme:protein ratio of 1:50. After digestion, the samples were snap frozen, lyophilized and dissolved in 20 µl mobile phase A (0.05% formic acid in Milli-Q water).

2.4. Liquid chromatography–tandem mass spectrometry

For each analysis, 2 µl of sample was loaded and desalted on a C18 PepMap 300 µm, 5 mm-i.d., 300 Å precolumn (Thermo Scientific) and separated by reversed-phase liquid chromatography using two identical 150 mm 0.3 mm-i.d. ChromXP C18CL, 120 Å columns (Eksigent) coupled parallel and connected to a splitless NanoLC-Ultra 2D plus system (Eksigent) with a linear 45-min gradient increasing from 4% to 35% acetonitrile in 0.05% formic acid and with a constant flow rate of 4 µl/min. The LC system was coupled to an amaZon speed ETD ion trap (Bruker Daltonics) configured with an Apollo II ESI source. After each MS scan, up to 10 abundant multiply charged species in the mass range of 300–1300 m/z were selected for tandem mass spectrometry and actively excluded for one minute after having been selected twice. The LC system was controlled by HyStar 3.2 and the ion trap by trapControl 7.1.

2.5. Molecular phylogenetics

Data Analysis version 4.2 (Bruker) was used to generate a 4000-compound MGF-file from the raw data of each run containing one type of meat sample (http://www.matrixscience.com/help/data_file_help.html). Each compound can be a sum of two or more tandem mass spectra of the same precursor acquired during the elution of the same chromatographic peak (peptide). These MGF files were used to analyze the meat samples and the earlier created MGF-files from fish samples (Wulff et al., 2013) with compareMS2 exactly (Palmblad & Deelder, 2012; Wulff et al., 2013). The output files, containing the number of shared tandem mass spectra or peptides between each pair of species, were loaded into MEGA6 build 6140226 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013). MEGA6 was used to generate a distance matrix and UPGMA tree using all default settings.

2.6. Building and searching spectral libraries

Spectral libraries were built for 26 different species, with two additional libraries for biological replicates of two species, including a sample of wild boar and roe deer. The raw LC–MS/MS data was first converted to mzXML (Pedrioli et al., 2004) using compassXport version 3.0.4. These mzXML files contain all MS/MS data from each acquired spectrum, and were searched with X!Tandem (Craig & Beavis, 2004) against random sequences as described in Wulff et al. (2013), using the K-score and with an allowed mass measurement error of 2.5 Dalton plus isotope error to ensure each spectrum would find at least one match in the random library, and converted to pepXML (Keller, Eng, Zhang, Li, & Aebersold, 2005). This is the standard workflow in the Trans-proteomic Pipeline, hijacked for the purpose of combining unidentified tandem mass spectra into SpectraST spectral libraries using common and well tested software. The spectral libraries created using SpectraST include all tandem mass spectra that pass the default SpectraST criteria like the minimum signal intensity, and number of peaks (Lam et al., 2007). Query samples in mzXML format were searched against the created spectral libraries using SpectraST with a dot product of over 0.7 considered as a match. The dot product cut-off of 0.7 leads to a false discovery rate of about 0.05 for each spectrum. The reliability is quantified as a false-discovery rate, measured by testing a number of samples and tallying the correct and incorrect answers.

3. Results and discussion

3.1. Spectral matching

All 28 spectral libraries, created from the mammal and bird samples, used for spectral matching were validated first. The validation