Research Article

Received: 9 March 2020 Revised: 6 June 2020 Accepted article published: 6 July 2020 Published online in Wiley Online Library: 3 August 2020

(wileyonlinelibrary.com) DOI 10.1002/jsfa.10643

Proteomic analysis of oilseed cake: a comparative study of species-specific proteins and peptides extracted from ten seed species

Klaudia Kotecka-Majchrzak,[a] Agata Sumara,[b] Emilia Fornal[b] and Magdalena Montowska[a]*

Abstract

Background: In recent years there has been a visible trend among consumers to move away from consuming meat in favor of plant products. Meat producers have therefore been trying to meet the expectations of consumers by introducing new products to the food market with a greater proportion of plant ingredients. Meat products are enriched not only by the addition of vegetable oils but also by ground or whole oilseeds or their preparation. In this study, we present in-solution tryptic digestion and an ultra-performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry (UHPLC-Q-TOF-MS/MS)-based proteomics approach to investigate specific proteins and peptides of ten oilseed cakes, by-products of cold pressing oil from coconut, evening primrose, hemp, flax, milk thistle, nigella, pumpkin, rapeseed, sesame, and sunflower seeds, for authentication purposes.

Results: We identified a total of 229 unique oilseed proteins. The number of specific proteins varied depending on the sample, from 4 to 48 in evening primrose and sesame. Moreover, we identified approximately 440 oilseed unique peptides in the cakes of all the analyzed oilseeds; the largest amounts were found in sesame (107 peptides), sunflower (100), pumpkin, hemp (42), rapeseed (36), and flax cake (35 peptides).

Conclusions: We provide novel information on unique / species-specific peptide markers that will extend the scope of testing the authenticity of a wide range of foods. The results of this peptide discovery experiment may further contribute to the development of targeted methods for the detection and quantification of oilseed proteins in processed foods, and thus to the improvement of food quality.

Supporting information may be found in the online version of this article.

Keywords: oilseed; protein composition; species-specific peptides; mass spectrometry; food authenticity

INTRODUCTION

In recent years, especially in developed countries, a nutritional trend has been observed where the consumption of plant products is preferred to the consumption of animal products. Flexitarians or vegetarians explain their food choices in terms of care for human health, concern for climate change, or a wish to follow a specific lifestyle.[1,2] Producers of meat products have therefore been trying to meet the expectations of consumers by introducing to the food market products such as sausages, pâtés, and ready meals with a reduced meat and fat content by increasing the plant protein and fat fractions or even launching vegetarian versions of meat dishes.[3,4,5]

Meat products have been enriched by the addition of vegetable oils, but ground or whole oilseeds, or their preparations, have also been used. On the European market, products with the addition of sunflower, hemp, and nigella seeds (black seeds/black cumin) are the most common. Various protein raw materials are used, such as oilseed meal or cake, protein concentrates and isolates, which have excellent nutritional value and are relatively easily digestible. In this way, it is possible to reduce the energy value of the product, while enriching it with protein, minerals, vitamins, and dietary fiber of plant origin. Seeds, cakes, and vegetable oils are also rich in bioactive compounds (phytosterols, tocopherols, phenolic compounds, bioactive proteins, and peptides). Their positive effects on human health include lowering blood pressure, reducing blood cholesterol, regulating blood glucose levels, improving the functioning of the digestive and nervous systems, and regulating the endocrine system.[6,7] For example, nigella seeds have been found to contribute to the production of red

* *Correspondence to: M. Montowska, Department of Meat Technology, Poznan University of Life Sciences, Wojska Polskiego 31, Poznan 60-624, Poland. E-mail: magdalena.montowska@up.poznan.pl*

a *Department of Meat Technology, Poznan University of Life Sciences, Poznan, Poland*

b *Department of Pathophysiology, Medical University of Lublin, Lublin, Poland*

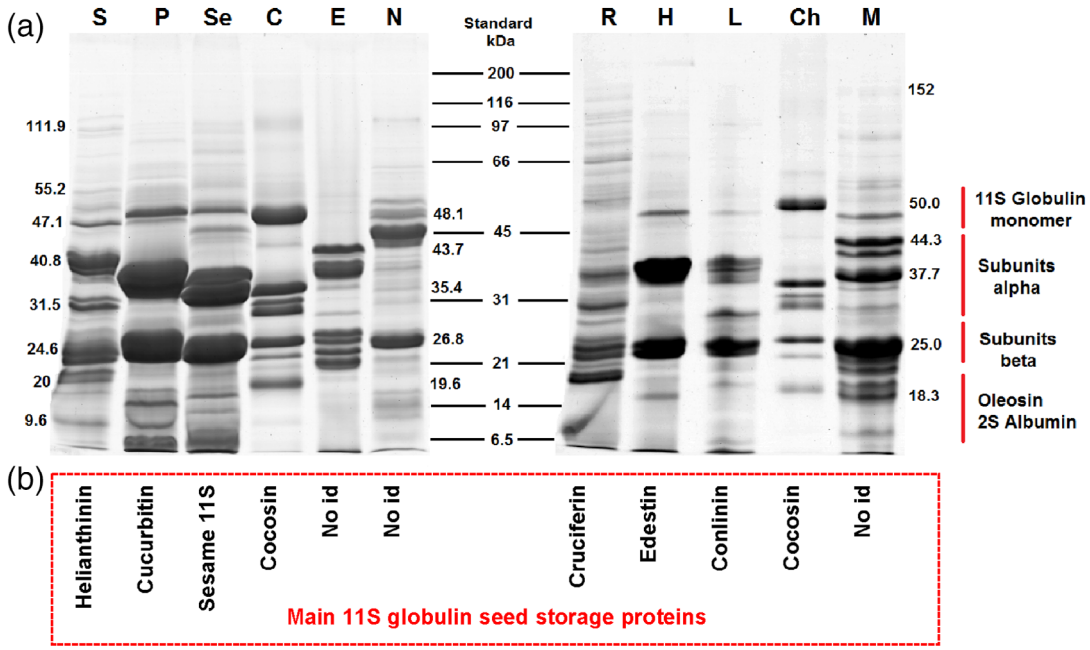

Figure 1 Oilseed cake protein profiles: (a) SDS-PAGE of the extracted protein fractions and molecular weight distribution; (b) Main 11S globulin seed-storage proteins identified using in-solution mass spectrometry analysis. Lanes: S – sunflower; P – pumpkin; Se – sesame; C – coconut; E – evening primrose; N – nigella; R – rapeseed; H – hemp; L – flax (linen); Ch – coconut crisps; M – milk thistle.

blood cells, increasing the speed and effectiveness of anemia treatment,[8] and regular consumption of unrefined rapeseed oil may be used to treat depression.[9]

The growing interest in meat proteins substitutes is translating into an increasing number of studies of oil processing by-products, which are proving to be a good and inexpensive source of vegetable protein. The preparations, apart from their nutritional value, also have very good functional properties. However, properties such as protein solubility, water-holding capacity, fat absorption, emulsification, gelation, and foaming properties will vary significantly depending on the type of raw material, the methods and devices used for its processing, and the method of protein extraction. When protein powders from coconut oil and coconut milk cakes were compared, milk cake presented higher water and oil absorption capacities but oil cake powder exhibited better foaming capacity and emulsifying activity.[10] Differences in solubility, interfacial activity and emulsifying properties of pumpkin seed protein isolate under different environmental conditions were described by Bučko *et al.*[11]

Producers are willingly launching plant matrix products onto the market. Foods of this type are becoming more competitive, which has an economic impact and allows the producer to achieve greater profits. However, frequently, illegal adulterations to the composition of food products are economically motivated,[12] and this involves a risk of allergic reactions among consumers. In addition to food fraud, diets based on vegetable proteins may be the reason for an increase in the occurrence of food allergies in society as a result of cross-reactions between protein homologues.[13,14] Consequently, there is an urgent need for research to determine the exact protein profiles in foods. Identification of specific proteins and peptides leads to control of certain authenticity issues, increased food safety, and care for human health by detecting individual allergens and their homologues in food products.[15]

In this article, we present a comparative proteomic study of ten oilseed cakes to identify and select specific proteins and peptides for the authentication of processed foods using a multiplex strategy. The protein composition of the oilseed cakes (by-products of cold pressing oil from coconut, evening primrose, hemp, flax, milk thistle, nigella, pumpkin, rapeseed, sesame, and sunflower seeds) was analyzed using ultra-performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry (UHPLC-Q-TOF-MS/MS) to detect specific proteins and peptide markers unique to the examined oilseed species. Oilseeds whose proteins had not been widely studied using mass spectrometry were selected for this peptide discovery experiment.

MATERIALS AND METHODS

Samples

The material for the study consisted of ten selected oilseeds, namely coconut (*Cocos nucifera* L.), evening primrose (*Oenothera biennis* L.), hemp (*Cannabis sativa* L.), flax (*Linum usitatissimum* L.), milk thistle (*Silybum marianum* L.), nigella (*Nigella sativa or N. indica*), pumpkin (*Cucurbita pepo* L.), rapeseed (*Brassica napus* L.), sesame (*Sesamum indicum* L.) and sunflower (*Helianthus annuus* L.). The seeds were obtained from the Polish company SemCo Sp. z o.o. (Szamotuły near Poznań, Poland) specializing in the production of oils. In addition, coconut crisps 'Bakalland S.A.' were analyzed to compare changes in sequence coverage due to different processing conditions. Coconut crisps were purchased at a supermarket (Lidl, Poznań, Poland). Seeds were stored at 4 °C for further analysis.

Preparation of oilseed cake and protein profile analysis

The cake was prepared by the cold pressing process using a Yoda oil press YD-ZY-02A (Warsaw, Poland). The oil temperature during the production process did not exceed 40 °C, and the efficiency of the pressing process in relation to the oil content was approximately 85%. The cake obtained was stored at −20 °C until

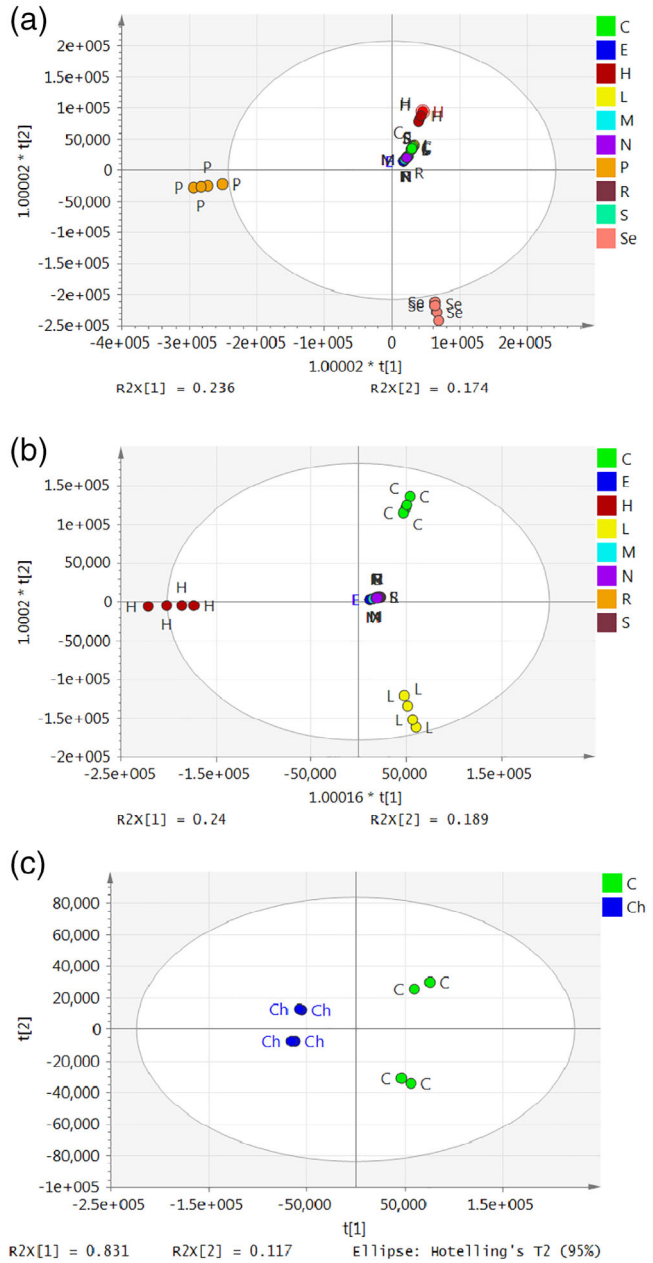

Figure 2 Differentiation between oilseeds: (A) OPLS-DA score plot of mass spectrometry data sets collected from all ten analyzed oilseeds; (B) OPLS-DA score plot after excluding pumpkin and sesame data sets; (C) PCA-X score plot showing differentiation between protein data set collected from coconut cake and coconut crisps. C, coconut; Ch, coconut crisps; E, evening primrose; H, hemp; L, flax (linen); M, milk thistle; N, nigella; P, pumpkin; R, rapeseed; S, sunflower; Se, sesame.

proteomic analysis. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed to compare the protein profiles of oilseed cakes. A sample of 5 mg was solubilized with lysis buffer and SDS-PAGE analysis was performed according to the procedure described previously.[16] The protein concentration was determined using a 2-D Quant kit (GE Healthcare Bio-Sciences, Fairfield, CT, USA). The cakes were analyzed in duplicate.

Trypsin digestion

Oilseed cake samples (3 mg) were rehydrated in 100 µL of 50 mmol L^{-1} ammonium bicarbonate. The proteins were reduced with 200 mmol L^{-1} dithiothreitol (DTT) (56 °C for 1 h) and then alkylated using 200 mmol L^{-1} iodoacetamide for 30 min in darkness at room temperature. The remaining iodoacetamide was quenched by the addition of 200 mmol L^{-1} DTT and incubation at room temperature for 30 min. The samples were digested in an ammonium bicarbonate solution containing 0.083 µg $µL^{-1}$ trypsin (Promega GmbH, Mannheim, Germany) at 37 °C overnight. Trypsin-to-protein ratio was approximately 1:40 (w/w). The digests were purified by reversed-phase extraction using Sep-Pak C18 Plus cartridges (Waters, Milford, MA, USA). The SPE column was equilibrated with solvent A consisting of 98 mL L^{-1} water, 20 mL L^{-1} acetonitrile, 1 mL L^{-1} formic acid, then with solvent B consisted of 65 mL L^{-1} acetonitrile, 35 mL L^{-1} water and 1 mL L^{-1} formic acid. The sample (0.6 mL) was then added to the cartridge and washed with solvent A. The peptides were eluted with solvent B and vacuum-dried in a centrifugal evaporator (miVacDuo Concentrator, Genevac Ltd, Suffolk, UK). Samples were resuspended in 20 mL L^{-1} acetonitrile in Milli-Q water containing 1 mL L^{-1} formic acid (solvent A) before UHPLC-Q-TOF-MS/MS analysis.

Protein and peptide identification

The UHPLC-Q-TOF-MS/MS analysis was performed on an Agilent Technologies (Santa Clara, CA, USA) 1290 Infinity series liquid chromatograph composed of a binary pump, a thermostat, and an autosampler, coupled to a 6550 UHD iFunnel Q-TOF LC/MS. Compounds were ionized by electrospray ionization (ESI) using a JestStream Technology ion source. Chromatographic separation was performed on a 2.1 × 150 mm, 1.8 µm particle-size Agilent RRHD Eclipse Plus C18 column. Instrument control and data acquisition were performed by using Agilent MassHunter Workstation Software. The liquid chromatography (LC) parameters were set as follows: 10 µL injection volume, 0.3 mL min^{-1} mobile phase flow. The mobile phase consisted of 1 mL L^{-1} formic acid in water (solvent A) and 1 mL L^{-1} formic acid in acetonitrile (solvent B). Gradient steps were applied as follows: 0–2 min, 2% B; 2–40 min, to 32% B; 40–45 min, to 37% B; 45–50 min, to 90% B; 50–55 min, 90% B and a 5 min post-run at 2% B. The ion source gas (nitrogen) temperature was 250 °C, the flow rate was 14 L min^{-1}, nebulizer pressure was 35 psig, the sheath gas temperature was 250 °C and the sheath gas flow was 11 L min^{-1}. The capillary voltage was set at 3500 V, nozzle voltage at 1000, and the fragmentor at 400 V. Positive ions formed in an electrospray were acquired in the range of 100–1700 m/z in MS scan mode and in auto MS/MS mode with a scan rate of 5 scan s^{-1} for MS and 3 scan s^{-1} for MS/MS. Internal mass calibration was enabled by using two reference masses at 121.0509 and 922.0098 m/z.

The National Center for Biotechnology Information (NCBI, US National Library of Medicine) protein database search for protein and peptide identification was performed, using the Spectrum Mill MS Proteomics Workbench with >70% precursor peak intensity and 5 ppm precursor mass tolerance, with the following parameters: trypsin enzyme, taxonomy green plants or a given taxonomy genus, two missed cleavages, 50 ppm product ions mass tolerance, carbamidomethylation as fixed modification, methionine oxidation as a variable modification, and peptide charge was set to a maximum of 6+. The matches and Spectrum Mill scores were evaluated at 1% of the false discovery rate (FDR) for identity and homology threshold. A search was made for selected peptides, in FASTA format, against the NCBInr database, using the protein Basic Local Alignment Search Tool

Table 1 Spectrum Mill output scores for coconut shreds and coconut crisps proteins obtained using the UHPLC-QTOF-MS/MS

Sample	Identified protein	Accession no.	Sequence coverage (%)	Number of peptides unique	Unique score
Shreds	11S globulin isoform 2	AKS26849.1	72.9	38	665.06
Crisps			60.0	22	384.32
Shreds	Cocosin	ASQ40963.1	73.8	36	617.12
Crisps			60.7	23	404.72
Shreds	Glyceraldehyde-3-phosphate dehydrogenase	AYJ72172.1	80.0	24	433.65
Crisps		XP_010910405.1	43.5	9	155.75
Shreds	Alpha galactosidase isoform 2	AIL28756.1	69.8	24	350.52
Crisps	Alpha-D-galactosidase, partial	ANF04455.1	45.1	4	62.42
Shreds	Elongation factor 1-alpha	AYJ72170.1	22.8	8	91.1
Crisps		XP_027348473.1	2.4	1	13.52
Shreds	Oleosin	AZZ09171.1	24.2	4	55.15
Crisps			20.0	3	40.03
Shreds	Pyruvate decarboxylase, partial	AFJ91675.1	6.8	1	14.12
Crisps	Pyruvate decarboxylase 2-like	XP_027336763.1	1.8	1	7.55

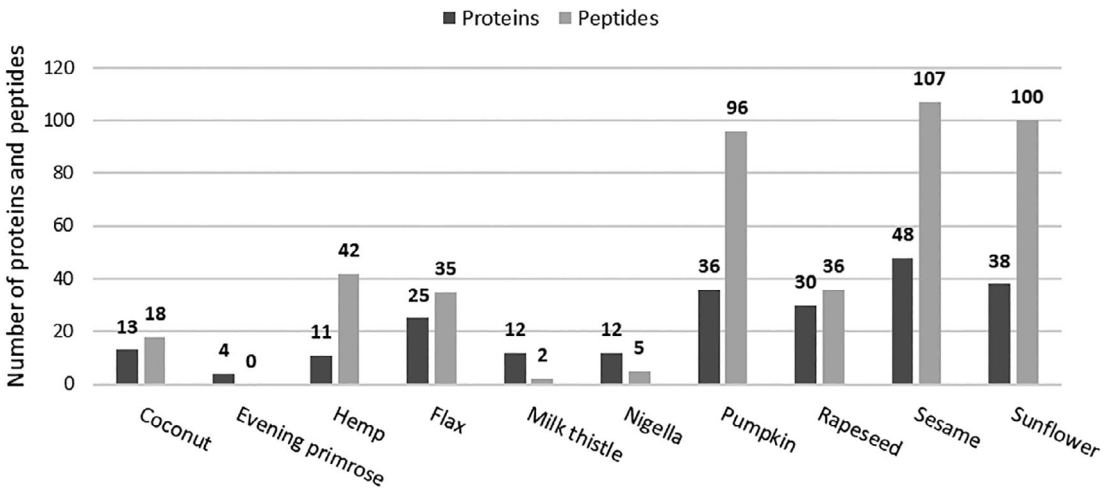

Figure 3 A number *of proteins and peptides identified in oilseed cake samples.*

(BLAST) and blastp algorithm (US National Library of Medicine, Bethesda, MD, USA), for species and protein specificity.

Multivariate data analysis

The raw MS data were processed by Agilent MassHunter Qualitative Analysis software using the Molecular Feature Extractor (MFE) algorithm. The m/z compounds list and their retention times were imported for multivariate data analysis (SIMCA-P version 13.1, Umetrics, MKS Instruments Inc.). The pre-processed data sets using Pareto scaling were initially overviewed using principal component analysis (PCA-X, unsupervised) to detect outliers in the model, and subsequently, to create a model with enhanced interpretability, the data sets were analyzed using a supervised orthogonal partial least-squares discriminant analysis (OPLS-DA).[17,18]

RESULTS AND DISCUSSION

Oilseed cake protein profiles

The profiles of the extracted protein fractions and molecular weight distributions are shown in Fig. 1(a). The examined oilseed cake samples obtained from coconut, evening primrose, hemp, flax, milk thistle, nigella, pumpkin, rapeseed, sesame and sunflower seeds were characterized by major protein bands in the molecular weight range from about 14 to 55 kDa. SDS-PAGE protein profiles showed significant species differences in both the distribution and intensity of protein bands. Globulins were the most abundant, but under reducing conditions the 11S monomers were less visible (molecular weight range from about 45 to 56 kDa) than subunits α and β (MW 30–45 kDa and 20–30 kDa, respectively). Proteins from the albumin family in the MW range of 6–20 kDa were also species-specific, but much less abundant. To date, only a few studies have examined the protein profiles of selected seeds, and protein fractions of evening primrose seeds have not been reported. Moreover, it is difficult to compare results, even within the seeds of one species, if the authors use different raw materials, extraction methods, and electrophoresis conditions. We obtained SDS-PAGE profiles similar to those that have been described previously for coconut meal,[19,20] flaxseed meal,[21] hemp seed meal and hemp protein isolate,[22,23] milk thistle seed protein isolate,[24] black cumin flour obtained from *Nigella*

Table 2 Four selected specific proteins identified in tryptic digests of oilseed cake, obtained from ten seed species (for more results see Tables S1–S10 in the supporting information)

Species	Identified protein	Accession no.	Sequence coverage (%)	Number of peptides unique	Unique score	Calc MW (Da)
Coconut						
Cocos nucifera	11S globulin isoform 2	AKS26849.1	72.9	38	665.06	52 956
Cocos nucifera	Cocosin	ASQ40963.1	73.8	36	617.12	52 656
Cocos nucifera	Glyceraldehyde-3-phosphate dehydrogenase	AYJ72172.1	80.0	24	433.65	36 951
Cocos nucifera	Oleosin	AZZ09171.1	24.2	4	55.15	14 386
Evening primrose						
Oenothera hartwegii	Ribulose-1,5-bisphosphate carboxylase/ oxygenase large subunit, partial (chloroplast)	AAO31550.1	2.0	1	13.51	47 820
Oenothera villaricae	Ycf1 protein (chloroplast)	ANI87061.1	1.0	2	9.8	283 732
Oenothera argillicola	Ribosomal protein L22 (chloroplast)	ABW98743.1	4.3	1	7.79	15 789
Oenothera clelandii	Putative LOV domain-containing protein	AML77629.1	1.2	1	5.47	62 924
Flax						
Linum usitatissimum	Tripeptidyl peptidase II	AFN53692.1	33.06	23	314.01	91 699
Linum usitatissimum	Allene oxide synthase	P48417.1	45.3	21	299.69	59 670
Linum usitatissimum	Conlinin	CAC94011.1	55.9	10	168.13	19 012
Linum usitatissimum	Oleosin high molecular weight isoform	ABB01624.1	49.4	8	130.7	18 712
Hemp						
Cannabis sativa	Edestin 1	CDP79023.1	82.9	53	881.57	58 504
Cannabis sativa	Edestin 2	CDP79028.1	78.4	50	880.86	55 970
Cannabis sativa	ATP synthase F1 subunit 1 (mitochondrion)	ALF04039.1	32.6	17	195.04	55 324
Cannabis sativa	Albumin	SNQ45151.1	42.2	10	148.16	16 742
Milk thistle						
Silybum marianum	Preprosilpepsin 1	AGE15494.1	12.4	5	70.18	54 989
Silybum marianum	Superoxide dismutase	AVN66530.1	36.6	5	58.12	15 340
Silybum marianum	Hypothetical chloroplast RF19 (chloroplast)	ALE29242.1	1.1	3	14.84	213 055
Silybum marianum	Photosystem I assembly protein ycf4 (chloroplast)	ALE29281.1	3.2	1	8.54	21 220
Nigella (black cumin)						
Nigella sativa	Chain A, nigellin-1.1	PDB: 2NB2_A	97.3	4	72.26	4221
Nigella sativa	Thionin NsW2	C0HJI0.1	65.7	4	42.84	3900
Nigella sativa	Non-specific lipid-transfer protein 1	P86527.1	76.0	2	28.16	2643
Nigella damascena	Ribosomal protein L22 (chloroplast)	QBK49549.1	10.6	2	10.86	21 370
Pumpkin						
Cucurbita pepo	11S globulin subunit beta	XP_023515280.1	71.2	40	744.64	54 500
Cucurbita pepo	Vicilin-like	XP_023527143.1	34.7	27	515.27	103 637
Cucurbita pepo	2S albumin	XP_023545481.1	37.5	8	130.06	16 409
Cucurbita pepo	Oleosin 18.2 kDa-like	XP_023550995.1	39.2	7	92.52	19 003
Rapeseed						
Brassica napus	Cruciferin subunit	AAK07609.1	71.5	28	515.01	54 834
Brassica napus	Embryonic protein DC-8-like isoform X1	XP_013687874.1	32.5	19	257.85	79 075
Brassica napus	Oleosin S2-2-like	XP_013677557.1	45.0	9	148.50	20 117
Brassica napus	Napin-3, 1.7S seed storage protein	P80208.1	52.8	7	131.83	14 035
Sesame						
Sesamum indicum	11S globulin seed storage protein 2 precursor	NP_001291336.1	81.4	34	630.82	49 925
Sesamum indicum	Embryonic protein DC-8-like	XP_011098036.1	51.1	25	419.84	64 648
Sesamum indicum	Heat shock 70 kDa protein	XP_011079384.1	46.5	23	356.67	71 459
Sesamum indicum	Oil body-associated protein 2A-like	XP_011102222.1	70.8	14	238.04	27 071
Sunflower						
Helianthus annuus	Putative 11S globulin subunit beta	OTG20713.1	63.2	30	523.90	57 867
Helianthus annuus	Seed biotin-containing protein SBP65-like	XP_022022103.1	50.8	31	517.85	81 287
Helianthus annuus	Seed storage albumin 2 precursor	ALO17641.1	68.7	15	240.66	33 228
Helianthus annuus	Embryonic protein DC-8-like	XP_022038548.1	41.6	16	213.97	40 653

damascena,[25] pumpkin seed protein extract,[11,26] rapeseed meal,[27–29] sesame protein isolate and seed flour,[30,31] and finally, from sunflower flour and protein concentrate.[32,33]

Electrophoretic protein patterns obtained from the analyzed oil seed cake samples under the same technical conditions confirm the specificity of their storage proteins, which may be used in

Table 3 Selected species-specific peptides identified in coconut, hemp, flax milk thistle, nigella, pumpkin, rapeseed, sesame and sunflower oil seed cake (for more results see Tables S11–S19 in the supporting information)

Parent Ion (m/z)	Mr (exp)	Exp z[a]	Total Intensity	RT (min)	Peptide Marker	Protein	Protein Score[b]
Coconut							
991.5071	1982.0050	2	7.97E+06	40.33	GFGTELLAAAFGIDMELAR	11S globulin isoform 2	665.06
635.3315	1269.6470	2	4.42E+07	20.60	AGSEGFQFVSIK	(AKS26849.1)	
561.7771	1122.5430	2	1.64E+07	15.43	GETVFDGELR		
714.3377	2140.9980	3	1.09E+07	27.13	GMVGLVMPGCPETFQSFQR		
1270.0962	5077.3590	4	1.54E+06	37.52	VYQFQEGDVLAVPNGFAYWCYNDGE NPVVAITVLDTSNDANQLDR		
523.7641	2092.0300	4	1.10E+07	14.78	GRVEVADDKGETVFDGELR		
619.6739	1857.0050	3	1.81E+07	29.62	QGQLLIVPQNFAMLER	cocosin (ASQ40963.1)	617.12
681.9973	2043.9760	3	1.81E+06	15.70	AENGLQVLRPSGMEEEER		
432.2241	863.4400	2	7.82E+06	5.82	CAGVSTIR		
1269.8492	5076.3740	4	5.71E+05	37.80	VYQFQEGDVLAVPNGFAYWCYNNGE NPVVAITVLDTSNDANQLDR		
Flax							
658.3107	1315.6059	2	1.31E+08	6.67	DLPGQCGTQPSR	conlinin (CAC94011.1)	168.13
646.9315	1938.7705	3	5.57E+08	7.10	GGGQQSQHFDSCCDDLK		
497.2214	1489.6448	3	1.21E+08	1.43	GGQGGQGQQQQCEK		
857.4158	1713.8151	2	3.26E+08	5.22	QDIQQQGQQQEVER		
596.8050	1192.5957	2	5.72E+07	11.48	QIQEQDYLR		
808.3767	1615.7459	2	2.47E+06	2.58	GGPYHQQGTGSGPSASK	oleosin high molecular weight	130.7
600.2640	1199.5184	2	2.13E+07	8.38	MQDAAGYMGQK	isoform (ABB01624.1)	
862.0914	2584.2552	3	4.63E+07	7.27	TTQPHQVQVHTQHHYPTGGAFGR		
722.0281	2164.0669	3	9.99E+06	19.38	YLQQAGQGVGVGVPDSFDQAK		
Hemp							
795.7362	2385.1920	3	6.01E+06	23.77	NAIYTPHWNVNAHSVMYVLR	edestin 1 (CDP79023.1)	881.57
822.3988	1643.7800	2	4.07E+07	18.77	YLEEAFNVDSETVK		
930.8343	2790.4830	3	2.10E+07	26.95	YTIQQNGLHLPSYTNTPQLVYIVK		
749.6182	2995.4480	4	8.34E+06	24.30	VEAEAGLIESWNPNHNQFQCAGVAVVR		
708.6765	2124.0070	3	3.38E+07	26.80	GILGVTFPGCPETFEESQR		
715.3268	1429.6420	2	1.33E+07	1.53	GQGQGQSQGSQPDR		
745.4300	1489.8370	2	7.61E+07	23.58	ISTVNSYNLPILR		
744.3491	1487.6800	2	4.51E+07	23.48	QASSDGFEWVSFK		
380.5400	1139.5990	3	1.71E+07	4.82	VQVVNHMGQK		
705.3624	1409.7120	2	1.37E+07	12.48	EETVLLTSSTSSR		
473.2215	945.4380	2	6.82E+06	1.82	LQGQNDDR		
535.8177	1070.6200	2	2.44E+07	21.40	GTLDLVSPLR		
595.2760	1189.5380	2	2.85E+07	2.85	QQNQCQIDR		
690.3825	2069.1320	3	1.13E+06	40.62	TLFLPQYLDSELTIFIR	7S vicilin-like protein	366.45
1016.0237	2031.0390	2	9.23E+05	22.58	EILSSQQEGPIVYIPDSR	(SNQ45153.2)	
477.5933	1430.7640	3	1.62E+05	26.82	GPELAAAFGLSLER		
711.3414	2132.0080	3	7.38E+05	32.00	NNYGWSIALDEFSYSPLR		
690.3828	2069.1320	3	1.13E+06	40.62	TLFLPQYLDSELTIFIR		
Milk thistle							
772.361	2315.069	3	6.57E+06	21.75	NVNEEEGGELVFGGVDPNHFR	preprosilpepsin 2 (AGE15495.1)	55.82
764.4089	1527.809	2	1.07E+07	29.00	IFELTPEQYIFK		
Nigella							
517.2484	1549.7280	3	4.64E+05	18.98	ACIGLCAPACLTSR	chain A, nigellin-1.1	72.26
716.2851	1431.5630	2	2.72E+06	6.52	YQDCLSECNSR	(PDB: 2NB2_A)	
568.2364	1702.6910	3	3.55E+05	6.83	DRYQDCLSECNSR		
673.7978	1346.5870	2	6.87E+06	17.32	CTYIPDYAGMR		
521.7097	1042.4110	2	4.77E+05	3.57	TCSGLCGCK	thionin NsW1 (C0HJH9.1)	37.59
Pumpkin							
588.9591	1764.859	3	1.19E+07	19.53	GIAIPGCAETYQTDLR	11S globulin subunit beta	744.64
1034.7792	3102.32	3	1.78E+07	27.88	AEAEAGFTEVWDQDNDEFQCAGVNMIR	(XP_023515280.1)	
697.3734	1393.733	2	3.48E+07	29.25	MLPLGVLSNMYR		
742.9163	1484.822	2	3.67E+07	16.62	ISTANYHTLPVLR		
923.4490	2768.328	3	6.77E+06	19.08	GVLYSNAMVAPHYTVNSHSVMYATR		
1359.6568	2718.305	2	7.07E+06	44.43	SGNLFSGFADEFLEEAFQIDGGLVR		

Table 3. Continued

Parent Ion (m/z)	Mr (exp)	Exp z[a]	Total Intensity	RT (min)	Peptide Marker	Protein	Protein Score[b]
607.3548	1213.694	2	4.85E+07	23.75	GLLLPGFSNAPK		
402.9055	1206.699	3	1.24E+06	25.25	LVFVAQGFGIR		
715.8913	1430.775	2	4.52E+06	14.35	NVANQIDPYLRK		
732.8849	1464.759	2	1.57E+07	12.45	FYLAGRPEQVER		
621.2866	1241.565	2	1.74E+06	10.45	MQEMAGYVGQK	oleosin 18.2 kDa-like	92.52
509.2519	1017.496	2	2.52E+06	2.93	EAGQEIQSR	(XP_023550995.1)	
504.2627	2517.275	5	1.18E+06	11.25	QVQVHHQQQRPSYLQEPTWK		
Rapeseed							
909.1354	2725.388	3	4.96E+06	25.43	GSIHNNAMVLPQWNVNANAALYVTK	cruciferin CRU4 (XP_013585668.1)	538.01
1011.0350	2021.060	2	5.50E+06	34.48	QNNIFNGFAPQILAQAFK		
1252.5595	3755.656	3	9.11E+05	22.92	VTPGCAETFMDSPVFGQGQGQEQGQGQGQGQGQGFR		
1082.2884	4326.125	4	2.05E+07	46.42	CGDTIATPPGVAQWFYNNGNEPLILVAAADIANNLNQLDR		
353.1999	1409.772	4	1.02E+06	15.37	LTFVVHGHALMGK		
1239.2795	3715.816	3	8.57E+06	31.33	CTENLDDPSSADVYKPSLGYISTLNSYNLPILR		
687.3447	1373.677	2	1.69E+07	9.73	SNDNAQINTLAGR		
754.3585	2261.058	3	6.20E+05	24.67	FSTLETTLTQSSGPMGYGMPR		
859.8713	1718.733	2	1.28E+06	15.28	VGYGAGYDYGYGADYK	oleosin S2-2-like (XP_013677557.1)	148.5
784.8692	1568.730	2	8.62E+05	3.25	AHEAHDTSLTTETR		
Sesame							
1208.1400	2415.270	2	1.77E+07	26.20	AMPLQVITNSYQISPNQAQALK	11S globulin seed storage	630.82
717.3678	2150.076	3	7.17E+07	31.55	AFDAELLSEAFNVPQETIR	protein 2 precursor	
927.7964	2781.367	3	3.45E+07	27.05	GNLYSNALVSPDWSMTGHTIVYVTR	(NP_001291336.1)	
769.3709	1537.725	2	4.17E+07	17.17	QEQFQCAGIVAMR		
689.6645	2066.967	3	4.32E+07	13.03	GDAQVQVVDHNGQALMNDR		
663.6326	1988.877	3	3.09E+07	9.93	MTFVRPDEEEGEQEHR		
531.7803	1062.548	2	2.47E+07	12.77	QTFHNIFR		
502.2730	2006.057	4	4.50E+07	15.20	STIRPNGLSLPNYHPSPR		
447.2197	1339.643	3	1.74E+06	23.50	AGNNGFEWVAFK		
670.8246	1340.640	2	7.13E+05	18.62	GFAIEVESTEMK	oil body-associated	238.04
746.3968	2237.174	3	6.12E+05	34.18	LPMGPPALMMSPQELDLGIVK	protein 2A-like (XP_011102222.1)	
803.4257	2408.260	3	3.13E+06	36.88	LIGVEYIISGGIFESLSPEEQK		
Sunflower							
973.4898	1945.973	2	1.14E+07	19.33	VQIVNNQGNSVFDNELR	11S globulin seed storage	523.9
518.2879	1552.848	3	9.36E+06	23.15	FPILEHLQLSAER	protein G3-like (OTG20713.1)	
660.9781	1980.916	3	1.72E+06	12.50	ENIDNPSHADFVNPQAGR		
704.3997	2111.183	3	7.35E+06	19.40	GHIVNVGQDLQIIRPPQAR		
619.6440	1856.915	3	8.74E+06	13.42	FFLAGNPQAQSQQQHR		
1059.8204	3177.443	3	2.20E+06	22.50	GIQGVILSGCPETYEYSQEQQFSGQSER		
418.5560	1253.652	3	2.73E+05	22.45	ETVLFAPSFSR		
928.9621	1856.915	2	1.23E+06	13.42	FFLAGNPQAQSQQQHR		
654.8073	1308.6	2	3.25E+07	13.80	GQFGGQEMDIAR	seed storage albumin 2 precursor	240.66
504.9376	1512.795	3	1.09E+06	18.70	AQILPNVCNLQSR	(ALO17641.1)	
741.0080	2221.008	3	3.17E+06	6.77	SQQCSETEIQRPVSQCQR		
739.8246	1478.636	2	1.96E+07	8.80	ECQCEAVQEVAR		
504.2589	1007.509	2	8.03E+06	10.77	MFLQQGQR		
1223.0804	2445.154	2	6.15E+06	23.65	VIQNLPNQCDLEVQQCNIPY		
1035.9656	4140.834	4	3.60E+06	19.80	GQQHQQQQEQQLLQQCCQELQNIDQQCQCEAVK		
885.4207	1769.835	2	5.90E+05	19.37	YVEQQMQSPMPYIR		
1130.5140	2260.019	2	6.23E+05	12.82	QQEQQGLQQCCNELQNVR		
609.7862	1218.564	2	2.11E+06	1.90	QCSQQVQGQR		

[a]Parent ion charge state.
[b]Spectrum Mill protein score at 1% FDR.

the MS-based investigation of food authenticity. The major 11S globulin seed storage proteins of eight cake species were identified using in-solution mass spectrometry analysis (Fig. 1(b)). No 11S-specific proteins were defined for evening primrose and milk thistle oilseed cake due to insufficient research data and incomplete databases on the protein and peptide composition of oilseeds. However, several specific proteins of a different nature have been identified for these two species (see below). Most of

the research concerns the functional properties and industrial applications of only several commonly grown oil plants, such as olive, rapeseed, sunflower, soy, and peanuts. A need for the availability of high-quality protein sequence information has therefore already been described in connection with MS-based detection of nut allergens.[34]

Differentiation between oilseed cakes using multivariate data analysis

Multivariate data analysis was performed to address the relationships among the selected types of oilseeds and to reveal differentiation or similarity between groups. To differentiate between the types of oilseed cake, multivariate analysis was applied to all peptide compounds extracted from MS data sets collected from in-solution digests of the oilseed cakes. The combined set contained 19 100 peptides. When principal component analysis (PCA-X, unsupervised) was initially performed, the model was not able to discriminate all the cake samples; moreover, distinct outliers that belonged to pumpkin and sesame samples were detected. Subsequently, a supervised OPLS-DA (supervised orthogonal partial least-squares discriminant analysis) model was created for the same data sets; however, no further enhancement of group separation was observed. The OPLS-DA gave an equally weak model with $R^2 = 0.657$ and $Q^2 = 0.313$ for the first four components (Fig. 2(a)).

We attempted to improve data variability by narrowing the analyzed groups and excluding outliers of pumpkin and sesame data sets. In this case, a slightly better model was obtained with $R^2 = 0.709$ and $Q^2 = 0.405$ for the first four components, but a satisfactory separation within the cluster was achieved only for flax, hemp, and coconut sets (Fig. 2(b)). Thus, when comparing the whole peptide sets, it was not possible to separate evening primrose, milk thistle, nigella, rapeseed, and sunflower cake samples.

Next, PCA-X was applied to coconut cake and coconut crisps sets to investigate whether the industrial processing, especially thermal treatment conditions, would have a significant impact on the protein identification and the quantity of identified peptides within the same species of *Cocos nucifera*. The model was able to separate the sample sets with a good spatial distribution and variance score (Fig. 2(c)). The two PCA components separating coconut cake from crisps displayed 94% of the total variance. Thus, the sample differentiation and further identification of proteins were affected by industrial processing. Coconut-specific proteins and unique peptides were identified in both samples but the sequence coverage and the peptide number for cake were higher than for crisps (Table 1).

Identification of specific proteins and peptides

In-solution tryptic digests of ten cold-pressed oilseed cake samples were analyzed by using the UHPLC-Q-TOF-MS/MS method with the Spectrum Mill searching algorithm against the NCBI protein database. At 1% of an FDR for the identity and homology threshold, in total 229 specific proteins were found, depending on the sample from 4 for evening primrose to 48 for sesame. The number of proteins and unique peptides identified in the cake samples is shown in Fig. 3. Seed proteins were identified with good sequence coverage, the number of matched unique peptides and a high score for uniqueness, with the exception of evening primrose, milk thistle and nigella, for which databases are very incomplete. These three species of seeds are good sources of protein, and their extracts had good protein band intensity and distribution (Fig. 1). Table 2 presents four selected proteins identified in each cake sample (all the specific proteins identified in the present study are shown in Tables S1–S10 in the supporting information). Most of the proteins identified were specific to the species investigated, except for evening primrose (*Oenothera biennis* L.), nigella (*Nigella sativa* L.) and rapeseed (*Brassica napus* L.), where the proteins have been assigned to other species, i.e. *O. hartwegii*, *O. clelandii*, *N. damascena*, *B. oleracea*, or *B. rapa*. In these cases, the specificity of the protein for a given genus can be confirmed.

The peptide analysis focused on unique / species-specific peptides. A selection of potentially specific peptides was made based on the Spectrum Mill output scores and peptide intensity. The most abundant peptides, with total intensities in the range of 10^5–10^8, and >70% scored peak intensity (SPI), were searched for species and protein specificity against the NCBInr protein database, using the BLAST alignment search tool. In total, 441 specific peptides were detected in the cake of all analyzed oilseeds. The largest amount was found in sesame (107 peptides), sunflower (100) and pumpkin cake (96 peptides) (Fig. 3). Significant numbers of unique peptides were also found in hemp (42), rapeseed (36), and flax cake (35 peptides). Evening primrose cake was the only matrix for which specific peptides could not be identified, due to incomplete databases as mentioned above. Table 3 presents selected unique peptides derived from two specific proteins identified in coconut, hemp, flax milk thistle, nigella, pumpkin, rapeseed, sesame, and sunflower oil seed cake (all specific peptides identified in the present study are shown in Tables S11–S19 in the supporting information).

To date, only a few studies of oilseeds have been conducted at the peptide level, especially in terms of potential use in food. Sequences of many proteins submitted to the NCBI protein database were obtained by DNA sequencing methods. Most oilseed-cake-specific peptides identified in this study were therefore likely presented for the first time. Previously, several peptides with antioxidant activity were identified by liquid chromatography and tandem mass spectrometry (LC–MS/MS) from the coconut cake globulin and glutelin-2 fractions.[20] Three non-species-specific peptides with ACE-inhibitory and antioxidant activities were detected in coconut cake albumin hydrolysates.[35] These coconut peptides did not coincide with those obtained in our study, whereas the *Nigella sativa* peptide TCSGLCGCK (*m/z* 521.7097^{2+}) identified in our study is part of antimicrobial thionin NsW1, which inhibits the viability of *Bacillus subtilis*, *Staphylococcus aureus*, and *Candida albicans*.[36]

In this paper, unique peptides for milk thistle and nigella oilseed cake are reported for the first time. Two milk-thistle-specific peptides, NVNEEEGGELVFGGVDPNHFR (*m/z* 772.361^{3+}) and IFELTPEQYIFK (*m/z* 764.4089^{2+}) are derived from preprosilpepsin 2, which is an acid protease with aspartic-type endopeptidase activity. Nigellin-1 was a source of four nigella unique peptides ACIGLCAPACLTSR (*m/z* 517.2484^{3+}), YQDCLSECNSR (*m/z* 716.2851^{2+}), DRYQDCLSECNSR (*m/z* 568.2364^{3+}), and CTYIPDYAGMR (*m/z* 449.5346^{3+}) (Table 3).

Regarding flaxseed, hemp, and pumpkin, Silva *et al.* (2017) evaluated the antioxidant activity of flaxseed protein isolate and detected four peptides in the antioxidant protein fraction released with Alcalase®.[37] Several bioactive peptides (antifungal, antimicrobial, anticarcinogenic and ACE inhibitors) were obtained from seed proteins of *Cucurbitaceae* such as pumpkin, squash, and melon.[7] However, the peptides mentioned above differ from those identified in our study due to the use of a different combination of methods and enzymes. Fifteen peptides unique to

hemp cake derived from edestin 1 and edestin 2 (8 and 7 peptides, respectively), identified in our study, have been detected previously in hempseed defatted flour.[38]

The molecular weights of sesame allergenic proteins are in the range of 14–96 kDa.[39] All seven main sesame allergens were identified in the examined cake samples, namely 2S albumin, 7S vicilin-like globulin, oleosin, and 11S globulin, as well as all 12 signature peptides for seven sesame allergens that previously have been selected by Ma *et al.* (2020).[40] However, after performing a BLAST search, we found out that only seven of them turned out to be unique to the sesame species. These seven species-specific markers are 2S albumin peptides QQQQEGGYQEGQSQQVYQR (Ses i 1) and MCGMSYPTECR (Ses i 2), oleosin GVQEGTLYVGEK and ATGQGPLEYAK (Ses i 4), 11S globulin peptides IQSEGGTTELWDER (Ses i 6), FESEAGLTEFWDR (Ses i 7) and EGQLIIVPQNYVVAK (Ses i 7). Three other proteins sharing the food route of allergen exposure were identified in the present study, i.e. pumpkin 11S globulin (Cuc ma 4), 2S albumin (Cuc ma 5), and rapeseed napin-3 (Bra n 1). For other species examined within this study, no allergic reactions have been reported, or allergic responses triggered after exposure by the food route are extremely rare. But when developing new products, food technologists must be aware of potential cross-reactions caused by protein homologues; for instance, peanut allergen homologues have been found in sunflower and pumpkin seeds, and many structural similarities between helianthinin and soy globulin 11S (glycinin) have been observed.[13,14]

Specific proteins and unique peptides presented in this paper are good material for further research into the authenticity of processed foods. Our future direction is to analyze and authenticate meat products manufactured with the addition of these oilseed proteins based on previously detected peptide markers of meat and seeds in parallel.

CONCLUSIONS

The discovery of specific protein and peptide markers is a response to the need for the qualitative and quantitative determination of individual food ingredients. In this work, a set of proteins and peptides specific to seed species and oil processing by-products was identified using a highly sensitive UHPLC-Q-TOF MS/MS method. Having identified a wide selection of specific proteins and peptides, we will likely be able to detect, simultaneously and unambiguously, oilseed protein additives in food products as well as the presence of some allergenic proteins, even if the plant proteins are partially degraded during processing. Thus, the main outcome of our study is that obtaining protein and peptide markers unique to coconut, evening primrose, hemp, flax, milk thistle, nigella, pumpkin, rapeseed, sesame and sunflower seeds could extend the scope of testing for the authenticity of a wide range of foods. This will help to increase food safety, which is crucial for nutritionists, food regulatory agencies, food producers, and, above all, consumers. However, further research is necessary to determine if modifications to the meat product recipe, especially replacement of part of the meat-fat fraction in the product with a vegetable-fat fraction, can be detected using the presented set of proteins and peptides.

ACKNOWLEDGEMENTS

The study was supported by the National Science Centre, Poland, project number 2017/25/B/NZ9/02000.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Božo J, Some aspects of modern nutrition, in *New Technologies, Development and Application (Pp. 610–616)*, ed. by Karabegović BS. E-Publishing Inc., Springer Nature Switzerland AG, Cham (2019).
2 Malek L, Umberger W and Goddard E, Is anti-consumption driving meat consumption changes in Australia? *Br Food J* **121**:123–138 (2019). https://doi.org/10.1108/BFJ-03-2018-0183.
3 Martin D, Ruiz J, Kivikari R and Puolanne E, Partial replacement of pork fat by conjugated linoleic acid and/or olive oil in liver pates: effect on physicochemical characteristics and oxidative stability. *Meat Sci* **80**: 496–504 (2008). https://doi.org/10.1016/j.meatsci.2008.01.014.
4 Choi YS, Choi JH, Han DJ, Kim HY, Lee MA, Jeong JY *et al.*, Effects of replacing pork back fat with vegetable oils and rice bran fiber on the quality of reduced-fat frankfurters. *Meat Sci* **84**:557–563 (2010). https://doi.org/10.1016/j.meatsci.2009.10.012.
5 Bilska A, Waszkowiak K, Błaszyk M, Rudzinska M and Kowalski R, Effect of liver pâté enrichment with flaxseed oil and flaxseed extract on lipid composition and stability. *J Sci Food Agric* **98**:4112–4120 (2017). https://doi.org/10.1002/jsfa.8928.
6 Hidalgo FJ and Zamora R, Peptides and proteins in edible oils: stability, allergenicity, and new processing trends. *Trends Food Sci Technol* **17**: 56–63 (2006). https://doi.org/10.1016/j.tifs.2005.10.006.
7 Ozuna C and León-Galván MF, Cucurbitaceae seed protein hydrolysates as a potential source of bioactive peptides with functional properties. *Biomed Res Int* **2121878**:1–16 (2017). https://doi.org/10.1155/2017/2121878.
8 El-Shanshory M, Hablas NM, Aboonq MS, Fakhreldin AR, Attia M, Arafa W *et al.*, *Nigella sativa* improves anemia, enhances immunity and relieves iron overload-induced oxidative stress as a novel promising treatment in children having beta-thalassemia major. *J Herbal Med* **16**:100245 (2019). https://doi.org/10.1016/j.hermed.2018.11.001.
9 Tayama J, Ogawa S, Nakaya N, Sone T, Hamaguchi T, Takeoka A *et al.*, Omega-3 polyunsaturated fatty acids and psychological intervention for workers with mild to moderate depression: a double-blind randomized controlled trial. *J Affective Disord* **245**:364–370 (2019). https://doi.org/10.1016/j.jad.2018.11.039.
10 Rodsamran P and Sothornvit R, Physicochemical and functional properties of protein concentrate from by-product of coconut processing. *Food Chem* **241**:364–371 (2018). https://doi.org/10.1016/j.foodchem.2017.08.116.
11 Bučko SD, Katona JM, Popović LM, Vaštag ŽG and Petrović LB, Functional properties of pumpkin (*Cucurbita pepo*) seed protein isolate and hydrolysate. *J Serbian Chem Soc* **80**:1–7 (2016). https://doi.org/10.2298/JSC150615081B.
12 Wisniewski A and Buschulte A, How to tackle food fraud in official food control authorities in Germany. *J Consum Prot Food Saf* **14**:319–328 (2019). https://doi.org/10.1007/s00003-019-01228-2.
13 Lakemond CM, Jongh HH, Hessing M, Gruppen H and Voragen AG, Soy glycinin: influence of pH and ionic strength on solubility and molecular structure at ambient temperatures. *J Agric Food Chem* **48**: 1985–1990 (2000). https://doi.org/10.1021/jf9908695.
14 Ozias-Akins P and Breiteneder H, The functional biology of peanut allergens and possible links to their allergenicity. *Allergy* **74**: 888–898 (2019). https://doi.org/10.1111/all.13719.
15 Stoyke M, Becker R, Brockmeyer J, Jira W, Popping W, Uhlig S *et al.*, German government official methods board points the way forward: launch of a new working Group for Mass Spectrometry for protein analysis to detect food fraud and food allergens. *Int J Anal Sci* **102**: 1280–1285 (2019). https://doi.org/10.5740/jaoacint.19-0056.
16 Montowska M, Fornal E, Piątek M and Krzywdzińska-Bartkowiak M, Mass spectrometry detection of protein allergenic additives in emulsion-type pork sausages. *Food Control* **104**:122–131 (2019). https://doi.org/10.1016/j.foodcont.2019.04.022.
17 Montowska M and Fornal E, Label-free quantification of meat proteins for evaluation of species composition of processed meat products. *Food Chem* **237**:1092–1100 (2017). https://doi.org/10.1016/j.foodchem.2017.06.059.

18 Pastor K, Pezo L, Vujić D, Jovanović D and Ačanski M, Discriminating cereal and pseudocereal species using a binary system of GC–MS data – a pattern recognition approach. *J Serb Chem Soc* **83**: 317–329 (2018). https://doi.org/10.2298/JSC170926014P.

19 Kwon K, Park KH and Rhee C, Fractionation and characterization of proteins from coconut (*Cocos nucifera L.*). *J Agric Food Chem* **44**: 1741–1745 (1996). https://doi.org/10.1021/jf9504273.

20 Li Y, Zheng Y, Zhang Y, Xu J and Gao G, Antioxidant activity of coconut (*Cocos nucifera L.)* protein fractions. *Molecules* **23**:707 (2018). https://doi.org/10.3390/molecules23030707.

21 Tirgar M, Silcock P, Carne A and Birch EJ, Effect of extraction method on functional properties of flaxseed protein concentrates. *Food Chem* **215**:417–424 (2017). https://doi.org/10.1016/j.foodchem.2016.08.002.

22 Malomo SA, Rong H and Aluko RE, Structural and functional properties of hemp seed protein products. *J Food Sci* **79**:C1512–C1521 (2014). https://doi.org/10.1111/1750-3841.12537.

23 Mamone G, Picariello G, Ramondo A, Nicolai MA and Ferranti P, Production, digestibility and allergenicity of hemp (*Cannabis sativa L.)* protein isolates. *Food Res Int* **115**:562–571 (2019). https://doi.org/10.1016/j.foodres.2018.09.017.

24 Li F, Wu X, Zhao T, Li F, Zhao J and Yang L, Extraction, physicochemical, and functional properties of proteins from Milk thistle *Silybum marianum* L. Gaernt seeds. *Int J Food Prop* **16**:1750–1763 (2013). https://doi.org/10.1080/10942912.2011.608176.

25 Alu'datt HM, Rababah T, Alhamad NM, Alodat M, Al-Mahasneh AM, Gammoh S *et al.*, Molecular characterization and bio-functional property determination using SDS-PAGE and RP-HPLC of protein fractions from two *Nigella* species. *Food Chem* **230**:125–134 (2017). https://doi.org/10.1016/j.foodchem.2017.03.025.

26 Rezig L, Chibani F, Chouaibi M, Dalgalarrondo M, Hessini K, Guéguen J *et al.*, Pumpkin (*Cucurbita maxima*) seed proteins: sequential extraction processing and fraction characterization. *J Agric Food Chem* **61**: 7715–7721 (2013). https://doi.org/10.1021/jf402323u.

27 Fretzer A, Herfellner T, Stäbler A, Menner M and Eisner P, Influence of process conditions during aqueous protein extraction upon yield from pre-pressed and cold-pressed rapeseed press cake. *Ind Crops Prod* **112**:236–246 (2018). https://doi.org/10.1016/j.indcrop.2017.12.011.

28 Perera PS, McIntosh CT and Wanasundara DPJ, Structural properties of Cruciferin and Napin of *Brassica napus* (canola) show distinct responses to changes in pH and temperature. *Plan Theory* **5**:36 (2016). https://doi.org/10.3390/plants5030036.

29 Raikos V, Neacsu M, Duthie G, Nicol F, Reid M, Cantlay LL *et al.*, Proteomic and glucosinolate profiling of rapeseed isolates from meals produced by different oil extraction processes. *J Food Process Preserv* **41**:e13060 (2017). https://doi.org/10.1111/jfpp.13060.

30 Achouri A, Nail V and Boye JI, Sesame protein isolate: fractionation, secondary structure and functional properties. *Food Res Int* **46**:360–369 (2012). https://doi.org/10.1016/j.foodres.2012.01.001.

31 Hassan BA, Mahmoud SN, Elmamoun K, Adiamo QO and Ahmed MAI, Effects of gamma irradiation on the protein characteristics and functional properties of sesame (*Sesamum indicum L.*) seeds. *Radiat Phys Chem* **144**:85–91 (2018). https://doi.org/10.1016/j.radphyschem.2017.11.020.

32 Salgado RP, Drago SR, Ortiz MES, Petruccelli S, Andrich O, González JR *et al.*, Production and characterization of sunflower (*Helianthus annuus* L.) protein-enriched products obtained at pilot plant scale. *LWT - Food Sci Technol* **45**:65–72 (2012). https://doi.org/10.1016/j.lwt.2011.07.021.

33 Žilić S, Barać M, Pešić M, Crevar M, Stanojević S, Nišavić A *et al.*, Characterization of sunflower seed and kernel proteins. *Helia* **33**:103–114 (2010). https://doi.org/10.2298/hel1052103z.

34 Xiong W, McFarland MA, Pirone C and Parker CH, Selection of tree nut allergen peptide markers: a need for improved protein sequence databases. *J AOAC Int* **102**:1263–1270 (2019). https://doi.org/10.5740/jaoacint.19-0054.

35 Zheng Y, Li Y and Li G, ACE-inhibitory and antioxidant peptides from coconut cake albumin hydrolysates: purification, identification and synthesis. *RSC Adv* **9**:5925–5936 (2019). https://doi.org/10.1039/C8RA10269D.

36 Vasilchenko AS, Smirnov AN, Zavriev SK, Grishin EV, Vasilchenko AV and Rogozhin EA, Novel Thionins from black seed (*Nigella sativa L.*) demonstrate antimicrobial activity. *Int J Pept Res Ther* **23**:171–180 (2017). https://doi.org/10.1007/s10989-016-9549-1.

37 Silva FGD, Hernández-Ledesma B, Amigo L, Netto FM and Miralles B, Identification of peptides released from flaxseed (*Linum usitatissimum*) protein by Alcalase® hydrolysis: antioxidant activity. *LWT - Food Sci Technol* **76**:140–146 (2017). https://doi.org/10.1016/j.lwt.2016.10.049.

38 Aiello G, Fasoli E, Boschin G, Lammi C, Zanoni C, Citterio A *et al.*, Proteomic characterization of hempseed (*Cannabis sativa* L.). *J Proteomics* **147**:187–196 (2016). https://doi.org/10.1016/j.jprot.2016.05.033.

39 Patel A and Bahna SL, Hypersensitivities to sesame and other common edible seeds. *Allergy* **71**:1405–1413 (2016). https://doi.org/10.1111/all.12962.

40 Ma X, Li H, Zhang J, Huang W, Han J, Ge Y *et al.*, Comprehensive quantification of sesame allergens in processed food using liquid chromatography-tandem mass spectrometry. *Food Control* **107**: 106744 (2020). https://doi.org/10.1016/j.foodcont.2019.106744.

