

Accepted Manuscript

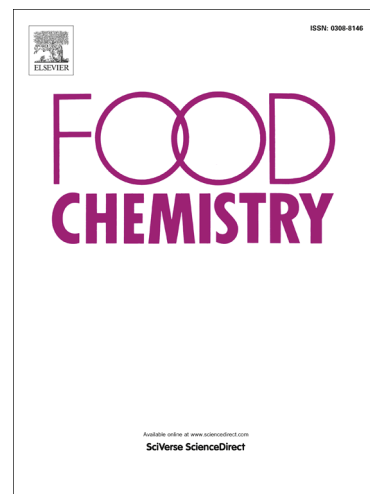
Discovery of Highly Conserved Unique Peanut and Tree Nut Peptides by LC-MS/MS for Multi-Allergen Detection

Jennifer Sealey-Voyksner, Jerry Zweigenbaum, Robert Voyksner

PII: S0308-8146(15)01053-5
DOI: <http://dx.doi.org/10.1016/j.foodchem.2015.07.043>
Reference: FOCH 17838

To appear in: *Food Chemistry*

Received Date: 27 March 2015
Revised Date: 18 June 2015
Accepted Date: 9 July 2015



Please cite this article as: Sealey-Voyksner, J., Zweigenbaum, J., Voyksner, R., Discovery of Highly Conserved Unique Peanut and Tree Nut Peptides by LC-MS/MS for Multi-Allergen Detection, *Food Chemistry* (2015), doi: <http://dx.doi.org/10.1016/j.foodchem.2015.07.043>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Discovery of Highly Conserved Unique Peanut and Tree Nut Peptides by LC- MS/MS for Multi-Allergen Detection.

Jennifer Sealey-Voyksner¹, Jerry Zweigenbaum³, Robert Voyksner^{2*}

¹ImmunogenX, 1502 West Hwy NC-54, Suite 504, Durham, North Carolina, USA

²LCMS Limited, 1502 West Hwy NC-54, Suite 504, Durham, North Carolina, USA

³Agilent Technologies, Inc., 2850 Centerville Road, Wilmington, Delaware, USA

* Corresponding Author: robert_voyksner@lcmslimited.com (919-201-0047)

Abstract:

Proteins unique to peanuts and various tree nuts have been extracted, subjected to trypsin digestion and analysis by liquid chromatography/quadrupole time-of-flight mass spectrometry, in order to find highly conserved peptides that can be used as markers to detect peanuts and tree nuts in food. The marker peptide sequences chosen were those found to be present in both native (unroasted) and thermally processed (roasted) forms of peanuts and tree nuts. Each peptide was selected by assuring its presence in food that was processed or unprocessed, its abundance for sensitivity, sequence size, and uniqueness for peanut and each specific variety of tree nut. At least two peptides were selected to represent peanut, almond, pecan, cashew, walnut, hazelnut, pine nut, Brazil nut, macadamia nut, pistachio nut, chestnut and coconut; to determine the presence of trace levels of peanut and tree nuts in food by a novel multiplexed LC/MS method.

Key words:

LC-MS/MS analysis of nuts, discovery of tree nut and peanut peptide markers, allergen detection of nuts in foods, nut specific peptide markers

24

25 **1.0 Introduction:**

26 Peanut and tree nut consumption is prevalent in many societies and has been recognized to
27 provide a coronary health benefit (Estruch et al., 2013). Approximately 1% of the US and UK
28 population is allergic to peanut, tree nuts and other nuts; the numbers of individuals allergic to
29 these foods in other countries varies because of diet and the ways in which nuts are processed
30 (Crespo, James, Fernandez-Rodriguez, & Rodriguez, 2007). Typically, food allergies develop
31 within the first two years of life but can also be lost as children grow. Allergies to milk can
32 disappear; however, tree nut and peanut allergies are rarely lost (Wood, 2003). Peanut allergies
33 are responsible for the most fatalities from anaphylactic shock (Bock, Muñoz-Furlong, &
34 Sampson, 2007; Hourihane, Kilburn, Nordlee, Hefle, Taylor, & Warner, 1997).

35 The primary treatment for food allergy is avoidance of the allergen. In order to aid consumers,
36 both the UK and the US have developed labelling requirements for food ("2003/89/EC," 2003;
37 "2007/68/EC," 2007; "Food Allergen Labeling and Consumer Protection Act," 2004). Accidental
38 exposure to tree nuts and peanuts can easily occur when children share food. Hidden allergens,
39 cross-contact during manufacturing, and school craft projects using nuts (e.g., peanut butter)
40 can result in serious health issues (Sicherer, Burks, & Sampson, 1998). The presence of hidden
41 allergens, resulting from manufacturing errors or unintentional contamination, necessitates a
42 multi-nut allergen detection method, capable of detecting and quantifying trace levels of such
43 allergens.

It is generally understood that certain allergenic proteins, typically those stable to heat and digestion, are the primary cause of the immune response (Sicherer & Sampson, 2010). Nut allergens are most commonly seed storage proteins (Crespo et al., 2007) and some of these have been studied extensively (Alcocer, Rundqvist, & Larsson, 2012; Burks, Shin, Cockrell, Stanley, Helm, & Bannon, 1997; De Leon, Drew, Glaspole, Suphioglu, O'Hehir, & Rolland, 2007; Nordlee, Taylor, Townsend, Thomas, & Bush, 1996, Teuber & Peterson, 1999). Given that some of these allergenic proteins have been well characterised, it is reasonable to assume that a method for their detection is needed. However, there is some debate on whether this is necessary, or whether simply the detection of a marker(s) representing the specific food generating the allergic response is needed.

The many challenges to allergen detection have been described (Van Hengel, 2007). Various methods have been used for food allergen detection. Both immunoassay and PCR (polymerase chain reaction) methods have been routinely employed. While immunoassays suffer from false positives due to protein cross-reactivity, PCR analyses detect the presence of DNA, not protein, and, thus, may not be reliable for allergens with little or no DNA present (Poms, Klein, & Anklam, 2004). ELISA and real-time PCR are the most commonly used quantitative methods for food allergens, with ELISA being the most prominently used analytical technique for hidden food allergen detection. Recent concern has focused on the effects of different food processing methods and the reliability of the analytical test used, whereby changes in proteins induced by the Maillard reaction were found to affect detection of some proteins (Cucu, Devreese, Trashin, Kerkaert, Rogge, & De Meulenaer, 2012; Platteau, Cucu, Taverniers, Devreese, De Loose, & De Meulenaer, 2013). It is difficult to detect multiple allergens in one ELISA assay, although at

least one ELISA method has been developed for multi-allergen detection (Ben Rejeb, Abbott, Davies, Cleroux, & Delahaut, 2005).

There is continued interest in methods that are more specific, more quantitative and more reliable for the detection of hidden and trace levels of allergens in food. Through the extensive study of food proteins using LC/MS, many researchers have discovered relevant peptide-based markers for food allergens. Two reviews in 2011 describe the field of “allergenomics” that would include both the top-down (intact protein analysis) and the bottom-up (peptide mapping) study of proteins, identifying the epitopes (protein components responsible for the allergic response) or identifying proteins that can be used for specific determination of the presence of an allergenic food (Faeste, Ronning, Christians, & Granum, 2011; Picariello, Mamone, Addeo, & Ferranti, 2011). We have previously utilized LC/MS to study gluten proteins in wheat as well as barley and rye grains (Sealey-Voyksner, Khosla, Voyksner, & Jorgenson, 2010; Sealey-Voyksner et al., manuscript in preparation). Our allergenomic approach allowed the identification of a suite of marker peptides, considered to be immunogenic to people with coeliac disease. MS can provide unambiguous identification of proteins while overcoming the problems of ELISA (cross-reactivity) and PCR (indirect protein identification) (Costa, Mafra, Carrapatoso, & Oliveira, 2012; Ansari, Stoppacher, & Baumgartner, 2012; Careri, Costa, et al., 2007; Shefcheck & Musser, 2004; Careri, Elviri, Lagos, Mangia, Speroni, & Terenghi, 2008) using LC-MS/MS with at least 2 peptides to represent each protein. As such, one research group has reported the use of LC-MS/MS for the detection of multiple allergens in four nuts (Heick, Fischer, & Popping, 2011).

To date, we believe there are no reports of a comprehensive method for the detection of most tree nuts and peanut in one analysis. Our proteomic approach, described here, uses LC-MS/MS to specifically detect 12 tree nut allergens and peanut in one analysis. This multiplexed approach provides the means to test food for hidden allergenic compounds with accuracy and sensitivity to satisfy both inspection and labelling purposes. Our methodology studied extracted nut proteins, digested with trypsin, then analysed by LC-MS/MS, using a quadrupole time-of-flight (q-TOF) mass spectrometer to identify peptides that represent known nut proteins. Other research has focused only on those peptides that are the most sensitive for MS analysis. A unique aspect of our research was the goal to find conserved peptides, those that would be present in both native (raw) and processed states.

Specificity is another important analytical parameter. To achieve sequence specificity, we chose peptides with a minimum of 6 amino acids. Larger peptides can offer more uniqueness, but result in reduced MS sensitivity. Each peptide sequence was searched against the NCBI nr protein database for plants, in order to verify that the sequences were unique, not only to the respective nut, but other common plant-based sources (such as barley, corn, rice, soy and wheat etc.), which are common ingredients in food products. As tryptic digestions may result in missed cleavage sites that can result in interfering peptides, we also confirmed the specificity of each nut peptide with respect to all the other nuts, as well as numerous plant-based ingredients used in food products to identify potential interferences.

2 Materials and methods

2.1 Chemicals and solvents

Water used for the sample preparation and LC separation was from an in-house Milli Q system.

Acetonitrile was from Fisher Scientific (Waltham, MA); trifluoroacetic acid (TFA) was from Sigma-Aldrich Corp. (St. Louis, MO); sodium phosphate dibasic was Fluka brand from Sigma-Aldrich; Hydrochloric acid (HCl) was purchased from Fisher Scientific; and tris(hydroxymethyl)aminomethane (TRIS) was from Sigma-Aldrich.

2.2 Sample preparation:

Peanut, almond, pecan, cashew, walnut, hazelnut, pine nut, Brazil, macadamia, pistachio, chestnut and coconut were purchased from local grocery stores in raw (native) form and roasted (thermally processed) form when available. Roasting of the raw nuts was also performed in a conventional oven at 176.7 °C for 30 min. The nut samples were ground to a fine power. The proteins from 30 mg of each sample were extracted for 2 hours at 50 °C using 1 mL of 50 mM TRIS/HCl (pH 7.5). Other extraction solvents were evaluated including: 0.01 M HCl (pH 2.5–3) and 50 mM dibasic sodium phosphate (pH 8).

These sample extracts were both analysed for intact proteins and digested for peptide mapping to identify markers. Extracts were centrifuged at 8000 *g* for 10 min and the supernatant was collected for analysis by LC-MS for the presence of proteins. Extracts that were subjected to enzymatic digestion were not centrifuged.

Enzymatic digestion was performed using trypsin (# T8003-500 mg, Sigma) on the 50 mM TRIS/HCl (pH 7.5) extracted sample. A volume of 25 μ L of a 12-mg/mL trypsin solution prepared in 50 mM sodium phosphate was transferred to each sample and the samples were incubated at 38 °C for 2 hours, vortexing each sample every 15 minutes. After 2 hours the reaction was stopped by either heating to 95 °C for 10 min or by the addition of 200 μ L of acetonitrile with 0.2% TFA. The digested extracts were centrifuged at 8000 *g* for 10 minutes and the upper 200 μ L were aliquoted to a 96-well plate for LC/MS analysis.

Packaged and finished foods were purchased in local stores. These foods were extracted and digested exactly as described for the raw and roasted nuts.

2.3 LC/MS analysis:

LC/MS analysis of the nut proteins and trypsin digest products was performed using an Agilent 1290 LC system and an Agilent 6530 high definition q-TOF mass spectrometer, equipped with a Jet Stream technology dual electrospray source. The specific conditions for the protein and peptide analysis are as follows:

2.3.1 Protein analysis

The nut proteins were separated on an Agilent Poroshell 300 column (2.1 x 75 mm, 2.7 μ m, 300 Å; Agilent Technologies, Santa Clara, CA), using a gradient of 0 to 60% **B** in 30 min (**A** = 95% water with 5% acetonitrile with 0.025% TFA and **B** = 5% water with 95% acetonitrile plus 0.025% TFA) then to 90% **B**, at a flow rate of 0.3 mL/min. The post-run column equilibration was 6 minutes. The column was held at 30 °C for the analysis. The injection volume was 5.0 μ L.

The q-TOF MS was operated in electrospray positive ion detection mode, specifically in 2 GHz extended dynamic range mode, scanning from m/z 300–2800 in 0.5 s. The instrument was calibrated over the mass range using the Agilent tuning solution (part no G1969-85000). The electrospray source conditions were as follows: capillary 4000 V, nozzle 1000 V, fragmentor 160 V, drying gas 350 °C at 9 L/min, nebulisation gas 30 psi and sheath gas 350 °C at 11 L/min. The first 2.5 minutes of the run were diverted to waste. The reference mass standards were delivered to the second nebuliser on an electrospray source by an isocratic pump operating at 0.7 mL/min with a 1:100 split resulting in 7 μ L/min flow rate into the electrospray source. The reference ions used were m/z 322.04812 and 2421.9140 from the ESI-TOF biopolymer analysis reference mass standards (Agilent part no G1969-85003). Agilent Mass Hunter software was used to acquire and process the data (acquisition version B.05.01 and qualitative data analysis version B.06.00).

2.3.2 Peptide analysis

The peptides released from the trypsin digestion of the nut proteins were separated on an Agilent Poroshell 120 column (2.1 x 50mm, 2.7 μ m; part no 699775-902) using a gradient of 0 to 40% **B** in 70 minutes (**A** = 95% water with 5% acetonitrile plus 0.025% TFA and **B** = 5% water with 95% acetonitrile plus 0.025% TFA) then from 40% **B** to 60% **B** at 80 min at a flow rate of 0.25 mL/min. The post-run column equilibration was 6 min. The column was held at 30 °C for the analysis. The injection volume was 20 μ L.

The q-TOF MS was operated in electrospray positive ion detection in 2 GHz extended dynamic range mode and auto MS/MS mode. For full single MS spectra the TOF covered the mass range

of m/z 300–2800 in 0.5 s. Auto MS/MS allowed the acquisition of MS/MS spectra of up to 3 of the most intense precursor ions selected in the single MS analysis at above 15,000 counts in intensity and in the m/z range 600–2000. The MS/MS spectra were generated using nitrogen as the collision gas and the collision energies (CE) were based on the precursor m/z as shown by the following equations:

Single charge precursor ions: $CE = m/z * 0.04$

Double charge precursor ions: $CE = m/z * 0.025$

Triply charged precursor ions: $CE = m/z * 0.022$

The MS/MS spectra were acquired at 3 spectra per second over the m/z range of 60–2000.

The instrument was calibrated over the m/z range using the Agilent tuning solution (Agilent part no G1969-85000). The electrospray source conditions were as follows: capillary 4000 V, nozzle 1000 V, fragmentor 160 V, drying gas 350 °C at 9 L/min, nebulisation gas 30 psi and sheath gas 350 °C at 11 L/min. The first 3 min of the run was diverted to waste. The reference mass standards were delivered to the second nebuliser on the dual Jet Stream electrospray source by an isocratic pump operating at 0.7 mL/min with a 1:100 split resulting in 7 µL/min flow rate into the electrospray source. The reference masses used were m/z 322.04812 and 2421.9140 from the ESI-TOF biopolymer analysis reference mass standards (Agilent part no G1969-85003). Agilent Mass Hunter software was used to acquire and process the data (acquisition version B.05.01 and qualitative data analysis version B.06.00).

2.4 Database searching

The Spectrum Mill MS Proteomics workbench Rev B.04.00.127 (Agilent Technologies) was used to search a green plant database containing 194,606 entries, generated from the full NCBI nr database. Searches were based on a mass tolerance window of 20 ppm for the precursor ion and 50 ppm window for the product ions around the measured m/z . The search used trypsin digestion with up to 2 missed cleavages with no modifications. The LC-MS/MS data from the tryptic digests of nut proteins were searched to identify peptide sequences from the major peaks in the LC/MS run and to match the identified peptide to a specific nut protein. An MS/MS scoring algorithm was used to match ions present in the spectrum with the sequence based upon factors described by Kapp et al. (2003).

Scores of 15 and above are well assigned fragmentation and correct interpretation. Scores greater than 10 are very good and usually correct. In addition, Spectrum Mill was used to confirm the amino acid sequences based upon the acquired MS/MS spectra as a secondary confirmation of the postulated sequence. Spectrum Mill was also used to verify that these sequences were unique to the nut analysed (based on a search of that sequence in the green plant NCBI nr database using MS Edman program) and would not be found in other nuts or green plants.

2.5 Peptide standards

Peptide standards and isotopically labelled peptides to confirm the identity of the peptide and to estimate quantity of the level of nut in a food were purchased from Thermo Scientific Pierce Protein Research (Rockford, IL). Each synthesised peptide was analysed by high resolution

accurate mass MS/MS and the resulting accurate mass and product ion spectra matched the theoretically expected sequence for each of the peptides. The LC/MS analysis indicated greater than 95% purity for the synthesised peptide. Peptide peanut 1 (p1) and peanut 3 (p3) shown in Table 1 were synthesised with isotopically labelled arginine ($^{13}\text{C}-6$ $^{15}\text{N}-4$) on the C-terminal end to serve as an internal standard for the nut analysis. The peptides were initially dissolved in 80/20 water/acetonitrile to make a 1 mg/mL stock solution, then diluted to make the threshold standard of 0.1 ppm.

2.6 Total protein assay

The Bradford assay was used to determine total protein extracted. The reagent was obtained from Sigma Aldrich (product # B6919) and the procedure supplied with the reagent was followed to obtain protein concentrations. Bovine serum albumin was used as the protein standard to calibrate the assay over the concentration range of 25 to 2000 $\mu\text{g/mL}$ (Kruger, 2009).

2.7 Detection of nuts in foods

The analysis of foods was performed qualitatively to determine if marker peptides above 0.1 ppm threshold could be detected. Detection was based upon the detection of 2 marker peptides and two product ions from each peptide within a 30% relative ion intensity ratio. For peanut, the 2 most intense marker peptides were used and the other 2 peptides served as confirmatory peptides. Food samples were spiked at 0.1 ppm and 1 ppm of the target peptides prior to extraction and digestion to demonstrate a signal could be observed in that food matrix. The LC/MS analysis used conditions described in section 2.3, except the q-TOF was operated in

target mode using the MRM conditions listed in Table 2 and the LC separation was shortened using a gradient from 0 to 40% **B** in 15 minutes.

3.0 Results and Discussion

3.1 Protein extraction

Three protein extraction solvents were evaluated based on yield of nut protein for peanuts, almonds, hazelnuts and walnuts. These solvents included 50 mM TRIS/HCl (pH 7.5), 0.01 M HCl (pH 2.5–3) and 50 mM dibasic sodium phosphate (pH 8). Based on the integrated total ion current (TIC) chromatogram for the extracted protein profile (data not shown), the TRIS/HCl extraction solvent resulted in the best yield of protein. The normalised protein LC/MS TIC intensities were 1: 0.87: 0.12 for 50 mM TRIS/HCl: 50 mM sodium phosphate: 0.01M HCl. There were no differences in the protein profiles between TRIS and phosphate, but HCl did not extract the higher molecular weight proteins, which comprise the majority of the nut proteins observed in TRIS and phosphate extracts. Based on this evaluation, TRIS/HCl was used for all the nut extracts. The Bradford assay indicated the TRIS/HCL extract typically contained in the range of 4 to 8 mg of protein for a 30 mg nut sample. Based upon the protein levels in selected nuts obtained from the USDA National Nutrient Database for Standard Reference Release 27, the protein extraction yield for the nuts evaluated were as follows: peanuts 83%, almonds 80%, pecans 96%, hazelnuts 92%, cashews 79%, Brazil nuts 91% and walnuts 85%. Triplicate extraction recoveries varied by less than 15% for the nuts discussed.

3.2 Protein profile in raw and roasted nuts

The peptides markers representing each specific nut need to be present in raw (native) and cooked or processed forms of that nut. Thermal processing of food most commonly results in glycosylation on arginine and lysine, resulting in a mass shift for tryptic fragment peptides (Poulsen et al., 2013). This makes peptide selection potentially difficult because it would be reasonable to expect protein changes with processing and therefore peptides released by enzymatic cleavage with trypsin could vary. To determine the extent of protein variation both raw and cooked nut proteins were analysed by LC/MS to determine molecular weight changes in the protein. The molecular weight of the unroasted peanut extract ranged from 12000 to 62000 Da with the majority of the ion current in the 50–62 kDa range. The roasting process changed the profile, reducing the 50-62 kDa protein intensity. Examination of the mass spectra for the roasted peanuts showed numerous peaks, indicating great heterogeneity. These data are not shown but the same phenomenon is observed in the mass spectra of the unroasted and roasted nuts. This could indicate glycosylation of the protein, resulting in numerous peaks in electrospray, which cannot be deconvoluted. The unroasted peanut mass spectra at the same retention times were more homogeneous, showing distinctive multiply charged ions that could be used in deconvolution to determine the molecular weight. For peanuts, the major proteins of the roasted nuts increased in molecular weight by over 1000 Da compared to the unroasted nut. Other tree nuts showed the same pattern, increasing in molecular weight upon roasting. Almonds, macadamia nuts, pecans, pine nuts and walnuts all showed more than a 5000 Da increase for the roasted nuts relative to the unroasted nuts. This qualitative investigation indicated proteins are modified; therefore, it is likely that roasting will result in changes in some

peptides in the protein and some of these changes may be due to glycosylation. Thus, sequences of any potential marker peptides chosen from the trypsin digestion of the nut proteins need to be verified that they are unaltered by cooking or processing (Blom, Sicheritz-Pontén, Gupta, Gammeltoft, & Brunak, 2004).

The protein level in roasted and unroasted nuts were measured by the Bradford assay to determine if potential changes in the peptide intensities could be attributed to differences in solubility between the roasted and unroasted nut proteins. The results for the Bradford assay indicated roasting did result in some loss in measurable protein relative to the unroasted nut, due to reduced solubility or interferences caused by glycosylation in the Bradford assay. The reductions of protein levels in the roasted nut relative to the unroasted nut were measured as follows: peanuts 14%, almonds 9%, pecans 2%, hazelnuts 6%, cashews 15%, Brazil nuts 14% and walnuts 11%.

3.3 Discovering peptide markers for nuts

The selection of tryptic digest peptide markers that uniquely represent each nut protein required a systematic evaluation of peptides generated. The workflow for the selection of marker peptides that represent each nut is outlined in Figure 1. The initial peptide selection was limited to those formed from trypsin as identified by Spectrum Mill search of the LC-MS/MS data against the NCBI nr database for plants, that resulted in scores that indicated confidence in the sequence and that sequence represented the target nut (Table 1). The q-TOF offered a significant advantage in the initial database searching due to its mass accuracy. This

was evident in the database search for peanuts, where the total number of hits was reduced from 1190 (22% were peanut protein related peptides) to 339 (72% were peanut protein related peptides) when the search window was reduced from a 1 Da window to a 20 ppm window. The measured mass accuracies for the parent ions were within 2 mmu of the exact masses as shown in Table 1.

The initial step in peptide selection began by selecting peptides with the best MS intensities that represent the nut protein based upon the NCBI nr database search. Selected peptides were also screened for potential problems with stability, including hydrolysis of aspartic acid, oxidation of methionine and deamidation of glutamine under basic conditions. The next step compared the selected peptide intensities in the roasted and unroasted nuts to verify that they showed little variation. For example, there were 10 peanut peptides identified that exhibited a high mass spectrometry response. All of these peptides showed within a 20% area repeatability except for two peptides (TDSRPSIANLAGENSIIDNLPEEVVANSYR and GTGNLELVAVR). These were eliminated from the potential marker peptide list, due to low or nonmatching intensities between roasted and unroasted peanuts. The resulting potential marker peptides for peanuts that did show minimal variation between roasted and unroasted are shown in Table 1. In addition, on a practical note, peptides greater than 20 amino acids in length were eliminated, due to cost of synthesis and spread of the ion current over more multiply charged ions. Also, peptides less than 6 amino acids were not used, due to lack of specificity. When possible, one or two peptides from different allergenic proteins were selected vs multiple peptides from one protein, to offer broader protein coverage. The potential marker peptides for other nuts listed in Table 1 were selected in the same fashion as described for the peanuts, based on the criteria

described. Of interest, almonds are a good example of how much variation can occur between roasted and unroasted nuts, with 58 unique LC/MS peaks detected in roasted almonds that were not detected in the raw almonds. Typically, there were 40 to 90 new peptides unique to each roasted nut.

The next criterion in marker peptide selection was to exclude those present in other nuts, plants and food that are subjected to the sample digestion. This is important, because nuts can be mixed with other grains to make food products typically found in stores or restaurants. A valid assay for the allergenic nuts must not show interferences in food matrices. Figure 2 shows the extracted ion chromatogram for 4 peanut peptides overlaid with 9 tree nuts and 6 grain digests. The chromatogram shows there are no common ions in the other nuts or grains that would interfere with any of the peanut peptides. Also, each sequence for each target peptide was searched against the NCBI nr plant database. If the sequence was not found in any other plant in the database, this would indicate that the target peptide was unique.

Table 1 summarises the suite of peptides discovered for each nut, based on the discussed criteria. There are four peptides in the table that were also identified by other researchers as significant peptide markers. The peptide INTVNSNTLPVLR found in hazelnut and identified as the marker peptide in Table 1 as h2 was also reported by Ansari et al. (2012). Likewise, marker peptides for peanuts have been sought by others and the peptides SFNLDEGHALR (Shefcheck et al., 2004) (p1 in Table 1), AHYQVVDSENGDR (Heick et al., 2011) (p4 in Table 1), and GTGNLELVAVR (Careri, Elviri, Mangia, & Mucchino, 2007; p3 in Table 1) that we found in peanuts were reported by others. From this work flow, 2 peptides were selected for each tree

nut and 4 were selected for peanuts, to serve as markers for each nut. Peptides that were detected in other plants by LC/MS (in immunoassay terms, measured cross-reactivity) or potentially the same peptide found in the NCBI nr database searches of other plants were not selected. The nut peptides did not show common interferences with grains such as barley, corn, rice, quinoa, soy and wheat. However, several nuts, such as walnuts, pecans and hazelnuts had peptide sequences in common, resulting in a lack of specificity for detection of each of these nuts. Those peptides having commonality were not selected as marker peptides to represent that nut. The chosen marker peptides were checked to verify the correct sequence against the acquired MS/MS spectrum, and then were synthesised to confirm the identity of each selected peptide. The selection criteria specified here exceed those enumerated by Johnson et al. (2011).

3.4 Detecting nut peptide markers in foods

The peptides discovered in this work were tested to verify their sequences were conserved after processing and now make it possible to screen or quantitate food products for the specific presence of 12 tree nuts and peanuts in one analysis. ELISA methodology would require multiple assays, one targeted towards each nut or small group of nuts and would suffer from cross-reactivity between some of the nuts. Our multiplexed LC/MS analytical method of detection of these marker peptides in foods was based on retention time, accurate mass (within a 10 ppm window) and 2 product ions from MS/MS. Table 2 lists the marker peptide retention times, accurate masses and MS/MS product ions that were used to detect each

peptide. In addition, the tryptic digest time was evaluated from 0.5, 1, 2, 4 and 8 hours to determine optimal formation of the target peptides, in order to achieve the best possible sensitivity. The 2-hour digestion time was chosen, since it produced a signal similar to the 4 and 8-hour digestion time, and was approximately a factor of 3 greater than the 1-hour digestion.

A variety of commercially available foods were analysed that contained nuts, did not contain nuts, and might contain nuts (because they were manufactured in facilities that also processed nut products). Figure 3 shows a representative extracted ion current chromatogram for 3 peanut marker peptides in several foods that contain peanuts, and in corn, which did not contain peanuts. Clearly it can be seen that the marker peptides were not lost due to food processing and there were no interferences from corn, which did not contain peanuts. Table 3 shows a complete list of the foods analysed and which nuts were detected in each. "Detection", in Table 3, is designated by a peak that is greater than 0.1 ppm of a nut peptide standard response. Spiked food samples and 0.1 and 1 ppm are also listed in Table 3 to demonstrate the food matrix did not interfere with the detection of target peptides. While some 0.1 ppm spikes were not detected (due to matrix effects), all 1 ppm spikes were detected. The 0.1 spike level was chosen because it is close to the detection limit of the method, while the 1 ppm level was chosen because it would be reliably detected in all samples. Current levels of ELISA detection are typically 2.5 -5 ppm. The quantitative measurement of these marker peptides provide the quantitative measure of the protein concentration (concentration of peptide = concentration of protein) of the respective nut in a food sample. Quantitative results that show how the calibrations were performed and relate peptide concentration to protein concentration of the respective nut protein are the subject of another manuscript (in preparation). Upon reviewing

the results shown in Table 3, grains or foods that should not contain nuts did not show false positive responses. Foods that did contain nuts as stated on the label showed a signal for the proper nut. Occasionally, other nuts were detected as well, which seems likely because many of the products were processed in facilities that process a variety of nut foods. The most significant findings are those analyses of food products that did NOT list specific nuts on their product labels, but stated that the product was processed or manufactured on shared equipment with other tree nuts. An example of this is macadamia nut cookies. The spectra showed an intense peak for macadamia nuts marker peptides, but also demonstrated the presence of low intensity marker peptides peaks for almonds, pecans and walnuts. Banana crème cake is an example of a food where the package label stated “processed” or “manufactured on shared equipment with other tree nuts”; however, no nuts were detected in the product. Peanut butter pretzels, manufactured on shared equipment with other nuts, did contain peanuts as stated but also showed a weak signal for walnuts, which were not specifically listed for this food product. Gluten-free crackers are a further example of a product labelled that was processed or manufactured on equipment shared with nuts was and it listed possible tree nuts that could be present. The LC-MS/MS analysis showed that marker peptides for pecans and hazelnuts were detected in the crackers.

4.0 Conclusions

Our enzymatic digestion LC-MS/MS methodology allows for sensitive and specific detection of multiple tree nuts and peanuts in a single analysis. To our knowledge, it is believed that this is

the most comprehensive LC/MS method for detecting and quantifying nuts, covering 11 tree nuts as well as peanuts. Sensitivity magnitudes greater than ELISA are achieved by our assay, thus allowing sub -ppm detection. We stress a key point in this research was in discovering appropriate marker peptides that will successfully detect these nuts in a variety of food products. Methodical selection criteria ensured each peptide sequence is conserved and unique for a particular nut. Verification that marker peptide sequences do not change from cooking or other processing steps is vital, and so is a thorough evaluation that these sequences are not found in other nuts or plant-based food ingredients. This minimised issues of cross-reactivity, as seen in ELISA, leading to false positive responses for the target compound. From the food analysis results presented here, it can be clearly understood that food products can be labelled “processed”, or “manufactured on shared equipment with other tree nuts or peanuts” and often do have other nuts present as unlabelled ingredients. Our LC-MS/MS methodology provides the much-needed means to successfully characterise food products, in order to minimise needless allergenic reactions in people and to ensure accurate food labelling.

5 Literature Cited

- 2003/89/EC. (2003). *Official Journal of the European Union*, L 308/315-318.
 2007/68/EC. (2007). *Official Journal of the European Union*, L 310/311-314.
 Alcocer, M., Rundqvist, L., & Larsson, G. (2012). Ber e 1 protein: the versatile major allergen from Brazil nut seeds. *Biotechnol Lett*, 34(4), 597-610.
 Ansari, P., Stoppacher, N., & Baumgartner, S. (2012). Marker peptide selection for the determination of hazelnut by LC-MS/MS and occurrence in other nuts. *Anal Bioanal Chem*, 402(8), 2607-2615.
 Ben Rejeb, S., Abbott, M., Davies, D., Cleroux, C., & Delahaut, P. (2005). Multi-allergen screening immunoassay for the detection of protein markers of peanut and four tree nuts in chocolate. *Food Addit Contam*, 22(8), 709-715.

- Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., & Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *PROTEOMICS*, 4(6), 1633-1649.
- Bock, S. A., Muñoz-Furlong, A., & Sampson, H. A. (2007). Further fatalities caused by anaphylactic reactions to food, 2001-2006. *J Allergy Clin Immunol*, 119(4), 1016-1018.
- Burks, A. W., Shin, D., Cockrell, G., Stanley, J. S., Helm, R. M., & Bannon, G. A. (1997). Mapping and Mutational Analysis of the IgE-Binding Epitopes on Ara h 1, a Legume Vicilin Protein and a Major Allergen in Peanut Hypersensitivity. *European Journal of Biochemistry*, 245(2), 334-339.
- Careri, M., et al. (2007). Use of specific peptide biomarkers for quantitative confirmation of hidden allergenic peanut proteins Ara h 2 and Ara h 3/4 for food control by liquid chromatography-tandem mass spectrometry. *Anal Bioanal Chem*, 389(6), 1901-1907.
- Careri, M., Elviri, L., Lagos, J. B., Mangia, A., Speroni, F., & Terenghi, M. (2008). Selective and rapid immunomagnetic bead-based sample treatment for the liquid chromatography-electrospray ion-trap mass spectrometry detection of Ara h3/4 peanut protein in foods. *J Chromatogr A*, 1206(2), 89-94.
- Careri, M., Elviri, L., Mangia, A., & Mucchino, C. (2007). ICP-MS as a novel detection system for quantitative element-tagged immunoassay of hidden peanut allergens in foods. *Anal Bioanal Chem*, 387(5), 1851-1854.
- Costa, J., Mafra, I., Carrapatoso, I., & Oliveira, M. B. P. P. (2012). Almond Allergens: Molecular Characterization, Detection, and Clinical Relevance. *Journal of Agricultural and Food Chemistry*, 60(6), 1337-1349.
- Crespo, J. F., James, J. M., Fernandez-Rodriguez, C., & Rodriguez, J. (2007). Food allergy: nuts and tree nuts. *British Journal of Nutrition*, 96(S2), S95.
- Cucu, T., Devreese, B., Trashin, S., Kerkaert, B., Rogge, M., & De Meulenaer, B. (2012). Detection of Hazelnut in Foods Using ELISA: Challenges Related to the Detectability in Processed Foodstuffs. *Journal of AOAC International*, 95(1), 149-156.
- De Leon, M. P., Drew, A. C., Glaspole, I. N., Suphioglu, C., O'Hehir, R. E., & Rolland, J. M. (2007). IgE cross-reactivity between the major peanut allergen Ara h 2 and tree nut allergens. *Molecular Immunology*, 44(4), 463-471.
- Estruch, R., et al. (2013). Primary prevention of cardiovascular disease with a Mediterranean diet. *N Engl J Med*, 368(14), 1279-1290.
- Faeste, C. K., Ronning, H. T., Christians, U., & Granum, P. E. (2011). Liquid chromatography and mass spectrometry in food allergen detection. *J Food Prot*, 74(2), 316-345.
- Food Allergen Labeling and Consumer Protection Act. In. (2004) *Title 2 of PUBLIC LAW 108-282*. United States.
- Heick, J., Fischer, M., & Popping, B. (2011). First screening method for the simultaneous detection of seven allergens by liquid chromatography mass spectrometry. *J Chromatogr A*, 1218(7), 938-943.
- Hourihane, J. O. B., Kilburn, S. A., Nordlee, J. A., Hefle, S. L., Taylor, S. L., & Warner, J. O. (1997). An evaluation of the sensitivity of subjects with peanut allergy to very low doses of peanut protein: A randomized, double-blind, placebo-controlled food challenge study. *Journal of Allergy and Clinical Immunology*, 100(5), 596-600.
- Johnson, P. E., et al. (2011). Current perspectives and recommendations for the development of mass spectrometry methods for the determination of allergens in foods. *J AOAC Int*, 94(4), 1026-1033.
- Kapp, E. A., et al. (2003). Mining a Tandem Mass Spectrometry Database To Determine the Trends and Global Factors Influencing Peptide Fragmentation. *Analytical Chemistry*, 75(22), 6251-6264.
- Kruger, N. J. (2009). The Bradford method for protein quantitation. In *The protein protocols handbook*, (pp. 17-24): Springer.

- Nordlee, J. A., Taylor, S. L., Townsend, J. A., Thomas, L. A., & Bush, R. K. (1996). Identification of a Brazil-Nut Allergen in Transgenic Soybeans. *New England Journal of Medicine*, 334(11), 688-692.
- Picariello, G., Mamone, G., Addeo, F., & Ferranti, P. (2011). The frontiers of mass spectrometry-based techniques in food allergenomics. *J Chromatogr A*, 1218(42), 7386-7398.
- Platteau, C., Cucu, T., Taverniers, I., Devreese, B., De Loose, M., & De Meulenaer, B. (2013). Effect of oxidation in the presence or absence of lipids on hazelnut and soybean protein detectability by commercial ELISA. *Food and Agricultural Immunology*, 24(2), 179-192.
- Poms, R. E., Klein, C. L., & Anklam, E. (2004). Methods for allergen analysis in food: a review. *Food Addit Contam*, 21(1), 1-31.
- Poulsen, M. W., et al. (2013). Advanced glycation endproducts in food and their effects on health. *Food and Chemical Toxicology*, 60(0), 10-37.
- Sealey-Voyksner, J. A., Khosla, C., Voyksner, R. D., & Jorgenson, J. W. (2010). Novel aspects of quantitation of immunogenic sheat gluten peptides by liquid chromatography-mass spectrometry/mass spectrometry. *J Chromatogr A*, 1217, 4167-4183.
- Shelfcheck, K. J., & Musser, S. M. (2004). Confirmation of the Allergenic Peanut Protein, Ara h 1, in a Model Food Matrix Using Liquid Chromatography/Tandem Mass Spectrometry (LC/MS/MS). *Journal of Agricultural and Food Chemistry*, 52(10), 2785-2790.
- Sicherer, S. H., Burks, A. W., & Sampson, H. A. (1998). Clinical features of acute allergic reactions to peanut and tree nuts in children. *Pediatrics*, 102(1), e6.
- Sicherer, S. H., & Sampson, H. A. (2010). Food allergy. *J Allergy Clin Immunol*, 125(2 Suppl 2), S116-125.
- Stevenson, S. E., Houston, N. L., & Thelen, J. J. (2010). Evolution of seed allergen quantification--from antibodies to mass spectrometry. *Regul Toxicol Pharmacol*, 58(3 Suppl), S36-41.
- Teuber, S. S., & Peterson, W. R. (1999). Systemic allergic reaction to coconut (Cocos nucifera) in 2 subjects with hypersensitivity to tree nut and demonstration of cross-reactivity to legumin-like seed storage proteins: New coconut and walnut food allergens. *Journal of Allergy and Clinical Immunology*, 103(6), 1180-1185.
- Van Hengel, A. J. (2007). Food allergen detection methods and the challenge to protect food-allergic consumers. *Anal Bioanal Chem*, 389(1), 111-118.
- Wood, R. A. (2003). The natural history of food allergy. *Pediatrics*, 111(6 Pt 3), 1631-1637.

Figure 1: Workflow for the discovery of marker peptides that represent the selected nuts discussed in this manuscript.

Figure 2: Extracted ion current chromatogram using a 10 ppm mass window for 4 peanut peptide markers (red trace) overlaid with chromatograms for the same masses from other nuts and grains listed below (black traces). All chromatogram traces were normalised to the same vertical intensity. Overlay traces with the peanuts include: almonds, Brazil nuts, cashews, hazelnuts, macadamia, pecans, pine nuts, pistachios and walnuts and grains, including barley, corn, rice, quinoa, soy and wheat.

Figure 3: Extracted ion current chromatogram for 3 peanut marker peptides in: A : unroasted peanuts 10 ppm standard, B: Nutter butter (ingredients state peanuts present), C: protein bar (ingredients state peanuts present), D: nut crisps (peanuts detected and label says manufactured in a facility that processes tree nuts but did not mention peanuts), E: 0.033 ppm peanut standard and F: corn blank.

7.0 Tables

Table 1

Summary of the peptides discovered based on the workflow discussed in the manuscript. Based upon intensity, near equal response between roasted and unroasted and absence of interference, two peptides (4 for peanuts) were selected as markers for each nut.

Table 2

Listing of the marker peptides selected for each peptide, including the retention time, parent ion mass and the 2 confirmatory product ions for peptide.

Table 3

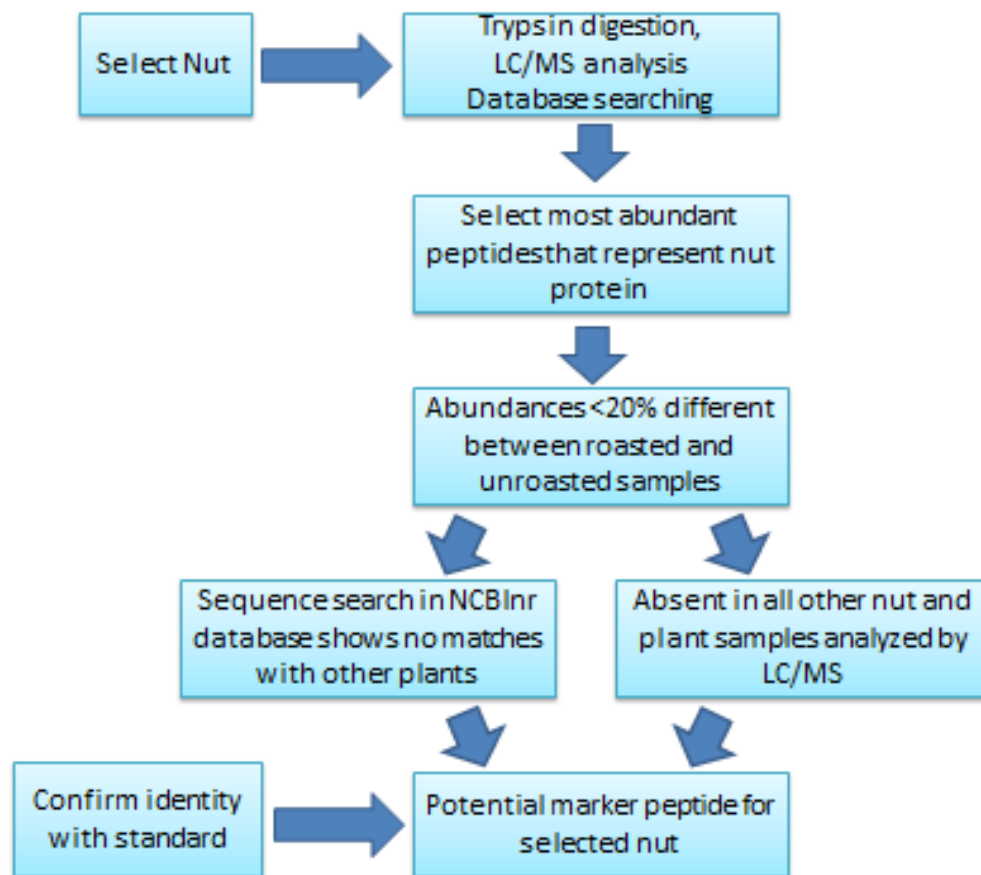
Listing of food products analysed by LC/MS/MS using MRM transitions in Table 2 for nuts above a threshold value of 0.1 ppm. The table lists findings as follows:

1. Nut listed as an ingredient and detected is displayed as a green bar.
2. Nut listed as a possible process contaminant in the facility and detected is displayed as a yellow bar.
3. Nuts are not listed but detected is displayed as a red bar.
4. Nut is not detected; background colour does not change.

Figure 1

543 Workflow for the discovery of marker peptides that represent the selected nuts discussed in
544 this manuscript.

545



546

Figure 2

Extracted ion current chromatogram using a 10 ppm mass window for 4 peanut peptide markers (red trace) overlaid with chromatograms for the same masses from other nuts and grains listed below (black traces). All chromatogram traces were normalized to the same vertical intensity.

Over lay traces with the peanuts include: almonds, brazil nuts, cashews, hazelnuts, macadamia, pecans, pine nuts, pistachios and walnuts and grains including: barley, corn, rice, quinoa, soy and wheat.

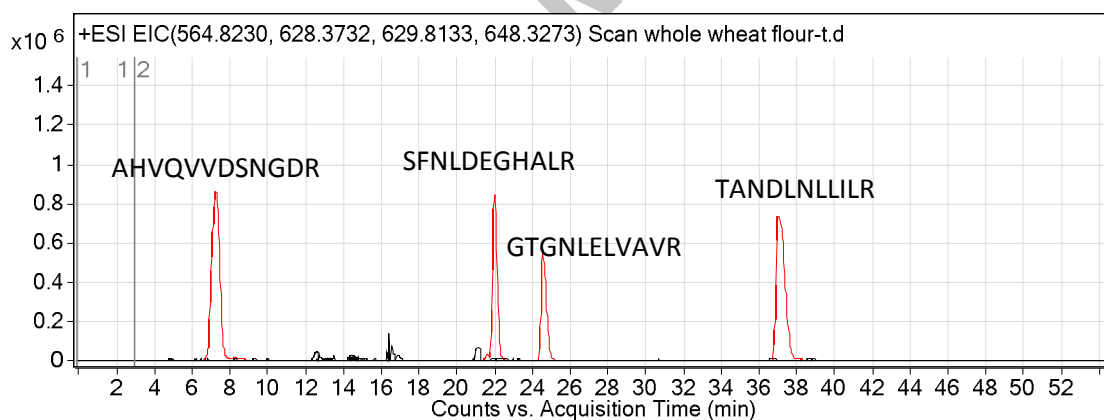
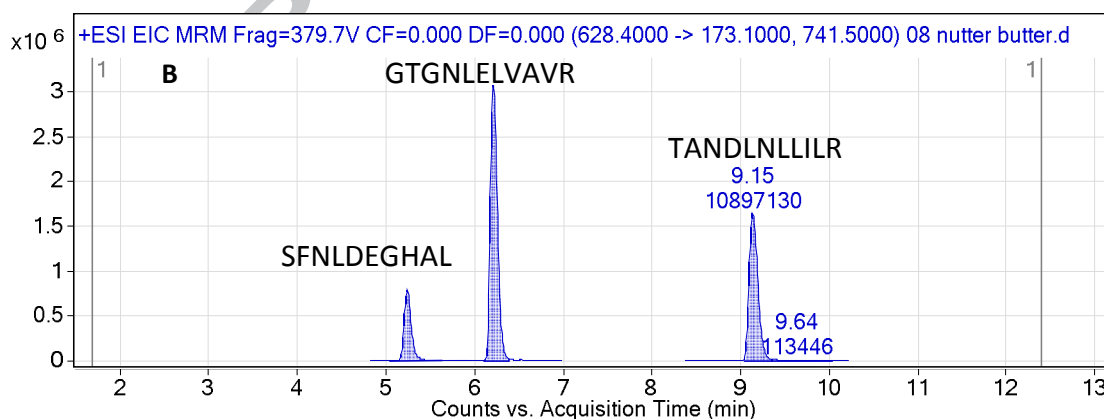
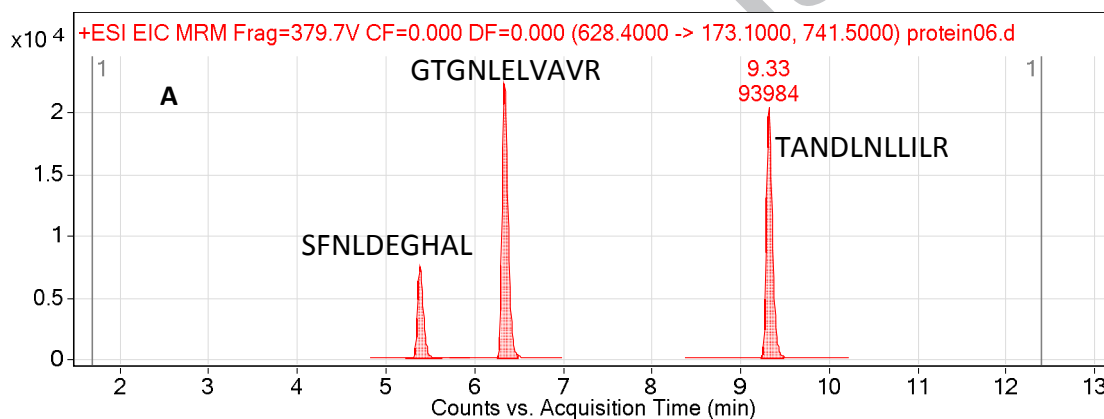
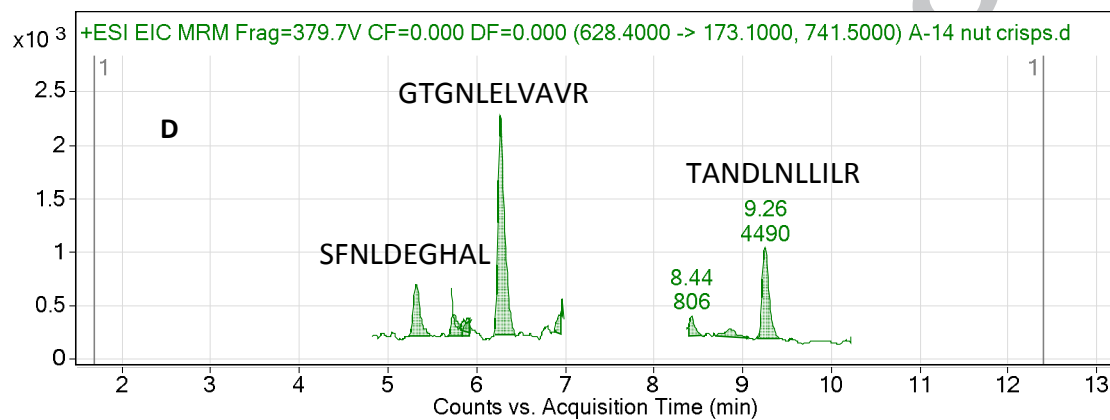
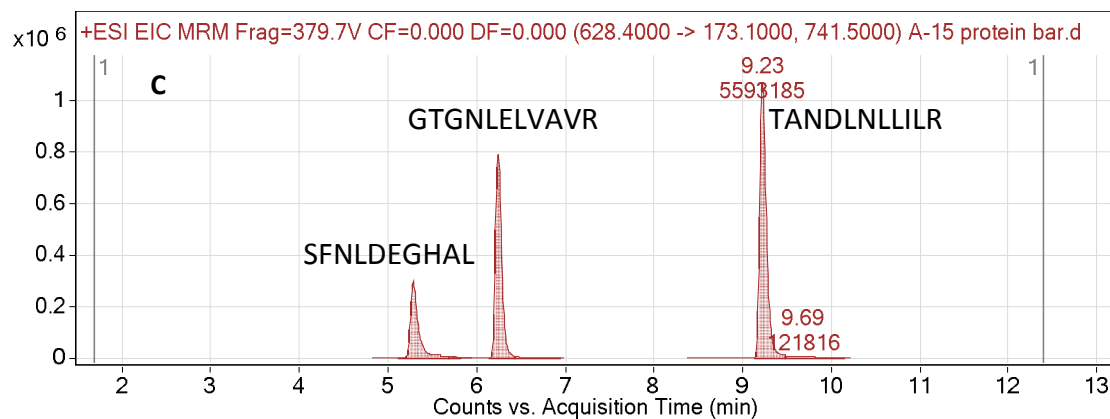


Figure 3

Extracted ion current chromatogram for 3 peanut marker peptides in: A : Unroasted peanuts 10 ppm standard, B: Nutter butter (ingredients states peanuts present), C: Protein Bar (ingredients states peanuts present), D: nut crisps (peanuts detected and label says manufactured in a facility that processes tree nuts but did not mention peanuts), E: 0.033 ppm peanut standard and F: Corn blank. The marker peptides are SFNLDEGHALR at a retention time of 20 min, GTGNLELVAVR at a retention time of 22.4 min and TANDLNLLILR at a retention time of 33.8 min.





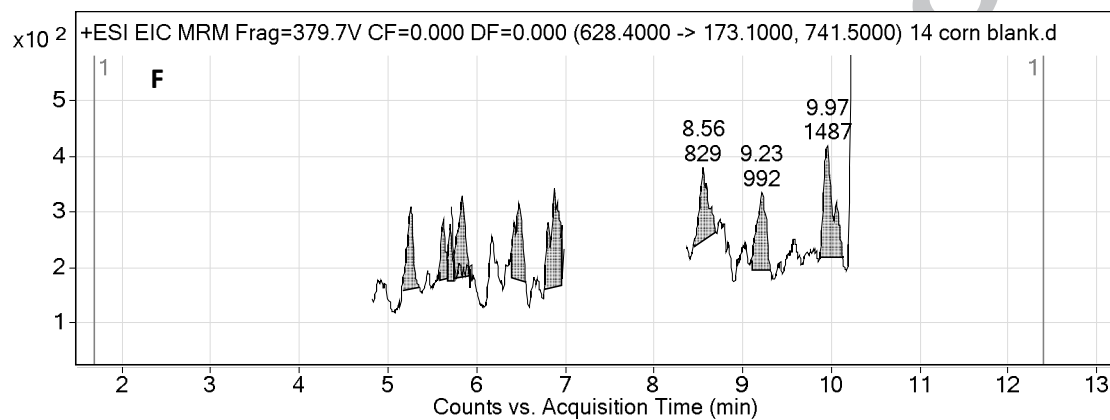
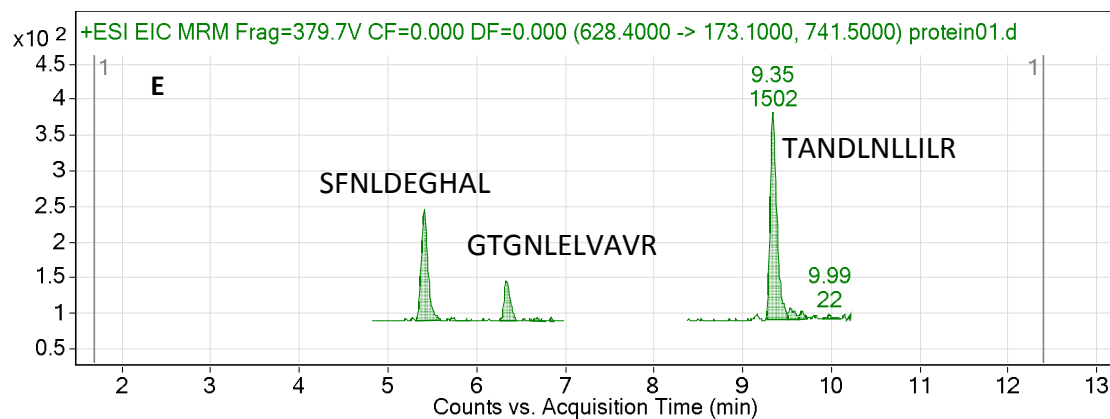


Table 1

<u>Nut (Latin name)</u>	<u>Selecte d peptide</u>	<u>Sequence</u>	<u>MS/MS sequence</u>	<u>Intensit y Ratio unroast ed/roas ted</u>	<u>Peak area (x 10⁶)</u>	<u>Accur ate Mass</u>	<u>Mas s erro r (mm u)</u>	<u>Protein Name</u>	<u>Protein Accessio n #</u>	<u>Prot ein MW (KD a)</u>	<u>Cross activity with other nuts based on LC/MS studies</u>	<u>NCBI nr search results</u>	
Peanut (arachis hypogaea)	p1	SFNLDEGHA				629.8				4656047	66.9		Allergen
		LR	22.8	1.14	135	133	0.2	conarachin	8	73		Ara h 1	
	p2	VLEENAGG				786.8				4656047	66.9		Ara h1
		EQEER	24.6	1.02	123	799	0.6	conarachin	8	73		allergen	
		GTGNLELVA				564.8				4656047	66.9		Ara h1
	p3	VR	17.8	1.11	93	23	0.2	conarachin	8	73		allergen	
		TANDLNLIL				628.3					60.4		Ara h2
		R	21	1.12	330	732	0.2	gly1	9864777	77		and 4	
	p4	SPDIYNPQA				695.3					60.4	Almonds / Pistachio	Ara h1, 2 and 4
		GSLK	21.2	1.19	368	544	1.1	gly1	9864777	77			
		FNLAGNHEQ				787.8					60.4		
		EFLR	23.3	0.98	89	9	0.7	gly1	9864777	77		Pecans	Ara h 3/4
		AHVQVVD				648.3					61.7		
		NGDR	20.2	1.08	309	273	0.1	Arah3/Arah4	5	37			
		WLGLSAEYG				771.3					60.4		
Almond (prunis dulcis)	a1	NLYR	24.1	1.01	186	926	0	gly1	9864777	77			
		TEENAFINTL				718.3					63.0		
	AGR	19.5	0.92	93	637	0.1	prunin	460806	16			prunin	
	a2	ISTLNSHNLPI				739.4					63.0		
		LR	24.9	0.94	161	292	0.3	prunin	460806	16			prunin
		LKENIGNPER	18.9	0.99	42	182	0.4	prunin	460806	16		Cashew	prunin
GNLDFVQPP					571.8					63.0			
Pecan		R	16.8	1.05	72	021	0.1	prunin	460806	16		prunin	
		LLQPVNNPG	18.1	1.19	16	691.8	1.6	vici;jin	6580762	69.9	Walnut	walnut/p	

(carya illinoensis)	QFR				823	precursor				89	ecans	
Brazil nut (bertholletia excelsa)	pe1	ATLTFVSQER	17.8	0.84	30	576.3083	2.1	7s Vicilin	157384600	95.167	pecans	
		LVGFGINGK	18.1	1.09	161	904.5251	0.4	7s Vicilin	157384600	95.167	Pecan	
		pe2	QFQEQQER	17.8	1.03	72	546.7578	0.4	7s Vicilin	157384600	95.167	Pecan
			b1	LAENIPSR	17.7	1.02	317	450.2516	0.3	2S albumin	843953309	16.9
	b2	MAENLPSR	18.7	0.98	393	917.4514	0.5	2S albumin precursor	209738721	16.755	brazil nuts	
		GGQQILADN				645.3			3031386	52.2	brazil nuts	
		VFK	18.9	1.09	203	472	1.2	11s globulin	7	95	brazil nuts	
		GETVFDDNL				583.2			3031386	52.2	brazil nuts	
		R	17.2	0.96	161	789	1	11s globulin	7	95	brazil nuts	
	Cashew (anacardium occidentale)	c1	QLQQQEIQIK	17.7	1.05	123	571.8129	0.1	2s albumin	24473800	16.335	cashew
c2			ADIYTPEVGR	18.9	1.1	169	560.7861	0.2	allergen Ana o 2	25991543	51.995	cashew
		ELYETASELP	19.9	1	216	654.3287	0.1	2s albumin	24473800	16.335	pitachio	
		WLQLSVEK	15.3	1.12	135	1002.5628	0.4	allergen Ana o 2	25991543	51.995	cashew	
pis2		LQELYETASE	19.1	1	131	774.9006	1.8	Pis v 1 allergen 2S albumin	110349081	17.289	pitachio	
pis1	TNGLSQT SQ	19.5	1	76	666.8	0.1	Pis v 2.0101	1103490	56.4	pitachio		

Walnut (juglans)	w1	LAGR				486		allergen11S globulin precursor- Pistacia vera	83	64		
		NIIIEVMEK	15.8	1	72	488.2 64	0.1	vicilin Pis v 5.0101 allergen11S	1337119 74	59.7 72	pitachio	
		AMISPLAGS TSVLR	21.9	1	27	701.8 912	1.8	globulin precursor albumin seed storage protein precursor	1718530 10	53.9 28	pitachio	
		GEEMEEMV QSAR	23.1	0.92	186	698.2 989	0.1	[Juglans regia]	1794252	16.3 73	walnuts	
		GDIFVIPAGH PIAITASQNE NLR	21.1	1.15	72	877.7 701	0.8	vicilin seed storage protein	6580762	69.9 99	pecan/w alnut	
	w2	LLGFGEINGE NNQR	23.6	0.96	37	716.3 722	1.8	vicilin seed storage protein albumin seed storage protein precursor	3132194 4	55.6 74	walnut	
		QQQQQGLR	15.9	1.07	89	493.2 634	1.1	[Juglans regia] 48-kDa glycoprotein precursor	1794252	16.3 73	walnut	
		LLSGIENFR	17.8	1.18	21	524.7 942	1.1	[Corylus avellana]	1933863 0	50.8 55	hazelnut	
		h1	LNALEPTNR	16.1	1.05	59	514.2 81	0.5	11S globulin- like protein	1847908 2	59.1 26	hazelnut

Pine nut (pinus pinea)	h2	INTVNSNTLP				720.9		11S globulin-like protein	1847908	59.1	walnuts, pistachios brazil, cashews	hazelnut walnut, pecan hazelnut, pistchio	
		VLR	21.3	1.04	59	136	0.2	seed storage protein	2	26			
		WLQLSAER	17.5	0.98	21	501.7	0.9		5678803	58.1			
	pin 1	TLAQSLDVEL				773.8		pine globulin-1		55.0			
		ETAR	15.7	1.09	68	464	1.9		20743	01		Pine nut	
	pin2	DYSSFPLAGP				885.4		pine globulin-1		55.0			
		ASSSGGGGR	16.3	1.02	76	102	2.2	cytosolic class 1 small protein	20743	01		Pine nut	
	Chestnut (Castanea)	ch1	VEVEDGNVL				822.4			3819743	17.4		
			QISGER	20.6	0.96	29	156	0.2			81		chestnut
	Coconut (Cocos nucifera)	ch2	AIGADLINNP				1142.		endochitinase		33.9		
DLVATNPTIS			19.7	1.11	22	6134	0.5		1354129	45		chestnut	
Coconut (Cocos nucifera)	co 1	VPGAEQLEQ				599.3		oleosin isoform	1995951	8.23			
		AR	18.9	1.06	38	14	1.3	OLE500c	03	1		coconut	
Macadami a nuts (Macadami a)	co 2	RPPGSEQLE				684.3		oleosin isoform	1995951	8.48			
		QAR	17.8	1.12	7	554	0.7	OLE500a	01	9		coconut	
	m 1	MPHFWSSN	Manually sequenced			567.2		not found in database				not in database	
		K		1.33	39	628	1.2						
	m 2	ELIAENRAAL	Manually sequenced			744.4		not found in database				not in database	
		MTR	Score > 15 indicates excellent match between MS/MS spectrum	1.19	9	033	0.4						
					1 to 24								
				25 to 99									

and peptide sequence

100 to
300
> 300

ACCEPTED MANUSCRIPT

Table 2

<u>Nut (Latin name)</u>	<u>Selected peptide marker (ID)</u>	<u>Sequence</u>	<u>RT (min)</u>	<u>Accurate Mass (m/z)</u>	<u>Charge State</u>	<u>Collision energy (V)</u>	<u>MS/MS mass (m/z)</u>	<u>Fragment type</u>
Peanut (arachis hypogaea)	p1	SFNLDEGHALR	20	629.8133	2	15.7	797.3891 682.3615	y7 y6
	p1 Int standard	SFNLDEGHALR	20	634.8174	2	15.8	807.3973 692.3697	y7 y6
	p2	GTGNLELVAVR	22.2	564.823	2	14.1	686.4175 444.2917	y6 y4
	p3	TANDLNLILR	33.6	628.3732	2	15.7	741.4969 1083.6504	y6 y9
	p3 Int standard	TANDLNLILR	33.6	633.3732	2	15.8	751.5051 1093.6586	y6 y9
	p4	AHVQVVDSENGDR	6.9	648.3273	2	16.2	1087.5488 989.4647	y10 y9
	a1	TEENAFINTLAGR	28.2	718.3637	2	18.0	891.5026 631.3522	y8 y6
	a2	ISTLNSHNLPILR	26.9	739.4292	2	18.5	498.3382 725.4668	y4 y7
	pe1	LVGFINGK	22.1	904.5251	1	36.2	692.3721 488.2855	y7 y5
	pe2	QFQEQQER	14.4	546.7578	2	13.7	689.3171 817.3772	y5 y6
Brazil nuts (bertholletia excelsa)	b1	LAENIPSR	11.3	450.2516	2	11.3	359.2031	y3
	b2	MAENLPSR	10.5	917.4514	1	36.7	715.3746 586.3291	y6 y5

[illegible]

R= Isotopically labeled
Arginine $^{13}\text{C}_6$ $^{15}\text{N}_4$

Table 3

[illegible]

Candy corn	none listed	Peanuts, tree nuts												
Instant Oat meal	walnuts	tree nuts												
Misco soup cup	none listed	none listed												
Falafil mix	none listed	peanuts, tree nuts												
cracker jacks	peanuts	none listed												
hot souce	none listed	none listed												
cereal	walnuts	none listed												
cracker crisps	chestnut	none listed												
chestnut puree	chestnut	none listed												
coconut nut cake	almonds, pistachio	tree nuts												
Gronola bar	none listed	peanuts, almomds, pecans												
coffee coated almonds	almonds	peanuts, tree nuts												

	Listed ingredient
	Nut listed as possible contamination
	Nut not listed as a possible contamination

Highlights:

- Discovery of peptide markers to detect nuts in foods.
- The discovered marker peptides were present in both raw and processed nuts.
- Enzymatic digestion and LC/MS/MS analysis to identify nut peptides from proteins.
- 12 tree nuts and peanut uniquely identified based on peptide markers.
- Can detect all tree nut and peanut allergens in one analysis.