Multi-Allergen Detection: Determining Unique Nut and Tree Nut Peptide Markers Using Accurate-Mass Q-TOF LC/MS

Application Note

Food

Authors

Jennifer Sealey-Voyksner
ImmungenX, Durham,
North Carolina, USA

Robert Voyksner
LCMS Limited,
Durham, North Carolina, USA

Jerry Zweigenbaum
Agilent Technologies, Inc.,
Wilmington, Delaware, USA

Abstract

The presence of hidden nut and tree nut allergens in foods can result in serious health issues, necessitating a method capable of detecting and quantifying them at trace levels using a single robust assay. To find specific peptide markers that can be used to determine the presence or absence of specific nuts in food, proteins unique to 11 tree nuts (almond, pecan, cashew, walnut, hazelnut, pine nut, brazil nut, macadamia nut, pine nut, chestnut, and coconut) and peanut were enzymatically digested and analyzed using accurate-mass Q-TOF LC/MS. Each marker peptide was selected by establishing its presence in raw and roasted nuts, processed and unprocessed food, abundance (sensitivity), sequence size, and uniqueness to a specific nut. The National Center for Biotechnology Information (NCBI) nr database searches were performed to confirm peptide identities, and to ensure that the marker peptides chosen were not present in other nuts or common food ingredients such as barley, corn, rice, soy, and wheat. Two marker peptides were selected for each tree nut, and four were selected for peanuts. Analysis of peptide digests from grains such as barley, corn, rice, quinoa, soy, and wheat did not present interferences. The peptide markers were tested to determine if they could be used to screen common foods for the presence of the 11 tree nuts and peanut at sub-ppm levels. Foods containing nuts as listed on the label showed a response for the correct nut. Foods processed using equipment also used to process other tree nuts or peanuts were in certain cases found to contain these other nuts.

Agilent Technologies

Introduction

Though the consumption of nuts is prevalent, an allergic response to tree and other nuts occurs in about 1% of the population in the US and UK, and varies significantly in other parts of the world depending on dietary preferences and nut processing methods [1]. The allergenic component of food is protein. Allergenic proteins, which are typically stable to heat and digestion, are the primary cause of an allergenic response [2]. Nut allergens are usually seed storage proteins including 7S viclin, profilin, glicinin, and 2S albumin [1]. Peanut allergies in particular are responsible for the most fatalities from anaphylactic shock [3]. The primary treatment for food allergy is avoidance. The presence of hidden allergens from manufacturing errors or unintentional contamination can result in serious health issues, and thus necessitates a method capable of detecting and quantifying multi-allergens at trace levels in one robust assay.

Current methodologies for food allergen detection have been reviewed [4], and the challenges well described [5]. These include real-time PCR, radio-allergosorbent tests, enzyme allergosorbent test, rocket immuno-electrophoresis, and enzyme-linked immunosorbant assay (ELISA). ELISA is the most commonly used quantitative method. Immunoassay can be expensive when measuring for multiple allergens, and suffers false positives from cross-reactivity of proteins in nonallergenic foods. Additionally, changes to the nut proteins upon processing can result in false negatives in ELISA. Real-time PCR DNA analysis is indirect and can suffer from a lack of DNA in the sample. Of concern is the effect of food processing on the test used. For example, changes in proteins induced by the Maillard reaction were found to affect hazelnut protein detectability in cookies [6]. There is debate about whether the detection of the protein itself, or marker(s) representing the specific protein generating the allergic response, would be better.

The field of proteomics using LC-MS/MS has enabled the discovery, detection, and quantitation of food allergen markers. Two reviews in 2011 describe allergenomics that includes top down (intact protein analysis) and bottom up (peptide mapping) of proteins to identify (1) the epitopes (protein components responsible for the allergic response) or (2) the proteins that can be used for specific determination of the presence of an allergenic food [7,8]. Faeste *et al.* includes extensive reference to LC/MS and MALDI-MS studies, the protein or food allergen targeted, and provides a number of food allergen targets and marker peptides. However, to date there are no reports of a comprehensive method for the detection of most (11 tree nuts and peanut) in one assay. Other proteomic approaches have failed to verify that the marker peptides will be present in both raw nuts and the ingredients used to make finished foods.

This application note presents an LC-MS/MS proteomic approach to develop a robust one-assay method to detect 11 tree nut and peanut allergens with high specificity. Using this approach, the nut proteins were digested with trypsin, then analyzed by accurate-mass Q-TOF LC/MS. Based on database searching, peptides representing known nut proteins were identified. Unique to this work was the discovery of peptides conserved in processed and uncooked nuts, as well as the confirmation of the uniqueness of each peptide marker for its specific nut. The complementary study of Discovery of Highly Conserved Unique Nut and Tree Nut Peptides by LC/MS/MS for Multi-Allergen Detection [9] provides a detailed description of the method and results.

Experimental

A detailed description of the sources of chemicals and solvents, and experimental procedures can be found in the complementary journal article published in Food Chemistry [9].

Sample preparation

Raw or roasted peanut, almond, pecan, cashew, walnut, hazelnut, pine nut, brazil nut, macadamia, pistachio, chestnut, and coconut were sourced from grocery stores. Raw nuts were roasted in an oven at 176.7 °C for 30 minutes, then ground to a fine power. A 30 mg sample was extracted for 2 hours at 50 °C with 1 mL of 50 mM *tris*-HCl (pH 7.5). Other extraction solvents were evaluated, but the *tris*-HCl was most effective in extracting the larger proteins. The samples were centrifuged at 8,000 rpm for 10 minutes, and the supernatant was analyzed by LC/MS to investigate intact proteins.

As described in the complementary journal article [9], the *tris*-HCl extract was enzymatically digested using trypsin (#T8003-500 mg, Sigma). After 2 hours, the reaction was stopped and the digested extracts were centrifuged at 8,000 rpm for 10 minutes. The supernatant was used for LC-MS/MS peptide analysis.

Q-TOF LC/MS analysis

LC/MS analysis of the nut peptides released from the trypsin digestion was performed using an Agilent 1290 Infinity LC system coupled with an Agilent 6530 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) LC/MS system equipped with an Agilent Jet Stream dual electrospray source. The HPLC and Q-TOF LC/MS parameters for the protein and peptide analyses are shown in Tables 1 and 2.

The Q-TOF LC/MS system was calibrated over the mass range using Agilent tuning solution (p/n G1969-85000). The reference mass standards were delivered to the second nebulizer of the dual Jet Stream electrospray source by an isocratic pump operating at 0.7 mL/min with a 1:100 split resulting in a 7 µL/min flow rate into the electrospray source. The reference masses were *m/z* 322.04812 and 2421.9140 from the Agilent ESI-TOF biopolymer analysis reference mass standards (p/n G1969-85003).

For peptide analysis, Auto MS/MS allowed the acquisition of MS/MS spectra of up to three of the most intense precursor ions (above 15,000 counts in intensity) across a mass range of 300–2,800 *m/z*. Collision energies (CE) were based on the precursor *m/z* using the equations:

Singly charged precursor ions: CE= *m/z* * 0.04

Doubly charged precursor ions: CE= *m/z* * 0.025

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

Table 1. HPLC Parameters

	Protein analysis	**Peptide analysis**
Instrument	Agilent 1290 Infinity LC system	Agilent 1290 Infinity LC system
Column	Agilent Poroshell 300 C18, 2.1 × 75 mm, 2.7 µm column (p/n 660750-902)	Agilent Poroshell 120 C18, 2.1 × 50 mm, 2.7 µm column (p/n 699775-902)
Mobile phases	(A) 95% water with 5% acetonitrile with 0.025% TFA (B) 5% water with 95% acetonitrile with 0.025% TFA	(A) 95% water with 5% acetonitrile with 0.025% TFA (B) 5% water with 95% acetonitrile with 0.025% TFA
Gradient	0 to 60% B in 30 minutes, then to 90% B in 3 minutes	0 to 40% B in 70 minutes, then from 40% B to 60% B at 80 minutes
Flow rate	0.3 mL/min	0.25 mL/min
Post run column equilibration	6 minutes	6 minutes
Column temperature	30 °C	30 °C
Injection volume	5.0 µL	20 µL

Table 2. Q-TOF LC/MS Parameters

	Protein analysis	**Peptide analysis**
Instrument	Agilent 6530 Accurate-Mass Q-TOF LC/MS system	Agilent 6530 Accurate-Mass Q-TOF LC/MS system
Ionization mode	Positive ion electrospray with Agilent Jet Stream technology	Positive ion electrospray with Agilent Jet Stream technology
Instrument mode	2 GHz extended dynamic range	2 GHz extended dynamic range and auto MS/MS
Mass range	300–2,800 *m/z* in 0.5 seconds	MS: 300–2,800 *m/z* in 0.5 seconds MS/MS: 60–2,000 *m/z* at 3 spectra/sec
Collision gas	N/A	Nitrogen
Drying gas	350 °C at 9 L/min	350 °C at 9 L/min
Sheath gas	350 °C at 11 L/min	350 °C at 11 L/min
Nebulizer gas	30 psi	30 psi
Fragmentor	160 V	160 V
Capillary	4,000 V	4,000 V
Nozzle	1,000 V	1,000 V

Data analysis

Agilent MassHunter Workstation software was used to acquire and process data (acquisition version B.05.01, qualitative data analysis version B.06.00).

Agilent Spectrum Mill for MassHunter Workstation (version B.04.00.127) was used to search a green plant database against the LC-MS/MS tryptic digest data to identify peptide sequences in the major peaks of the LC/MS run, and to match these to a specific nut protein. Generated from the full NCBI nr database, the green plant database contained 194,606 entries. Searches were based on a mass tolerance window of 20 ppm for the precursor ion and 50 ppm for the product ions around the measured *m/z*.

The green plant database was also searched using the MS Edman utility tool in Spectrum Mill to confirm the MS/MS amino acid sequences and to verify that the sequences were unique to the nut analyzed and not found in other nuts or green plants.

Peptide standards

The peptide standards and isotopically labeled peptides used to confirm peptide identities and estimate quantities of nuts in a food were from Thermo Scientific Pierce Protein Research (Rockford, IL). The peptides were 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.

Results and Discussion

Roasted versus raw nuts: protein profile differences

Because it is expected that nut proteins, and thus the peptides released by digestion will change with processing, the peptide markers chosen to represent each nut must be present in raw and cooked or processed forms of that nut. To determine the extent of protein variation due to processing, both raw and cooked nut proteins were analyzed to determine changes in their molecular weights.

Figure 1A shows the LC/MS total ion chromatograms (TICs) of roasted and unroasted peanut proteins. The molecular weight of the unroasted peanut extract ranged from 12K to 62K Da, with the majority of the ion current in the 50–62K Da range. The roasting process changed the TIC, reducing the intensity of the 50–62K Da protein. The mass spectrum of the roasted peanuts (Figure 1B) contained numerous peaks, indicating heterogeneity and a 10-fold reduction in signal compared to that of the unroasted peanut (Figure 1C). This could be due to protein glycosylation from the Maillard reaction, which resulted in numerous peaks that could not be deconvoluted to determine discrete protein molecular weights. The unroasted peanut mass spectrum (Figure 1C) at the same retention time was more homogeneous, showing distinctive multiply-charged ions that could be deconvoluted to determine its molecular weight. Table 3 shows that the predominant nut proteins analyzed increased in molecular weight by over 1,000 Da, due to roasting.

Roasting resulted in protein changes, and some of these changes could be through glycosylation. Therefore, the marker peptides chosen must be unaltered by cooking or processing.

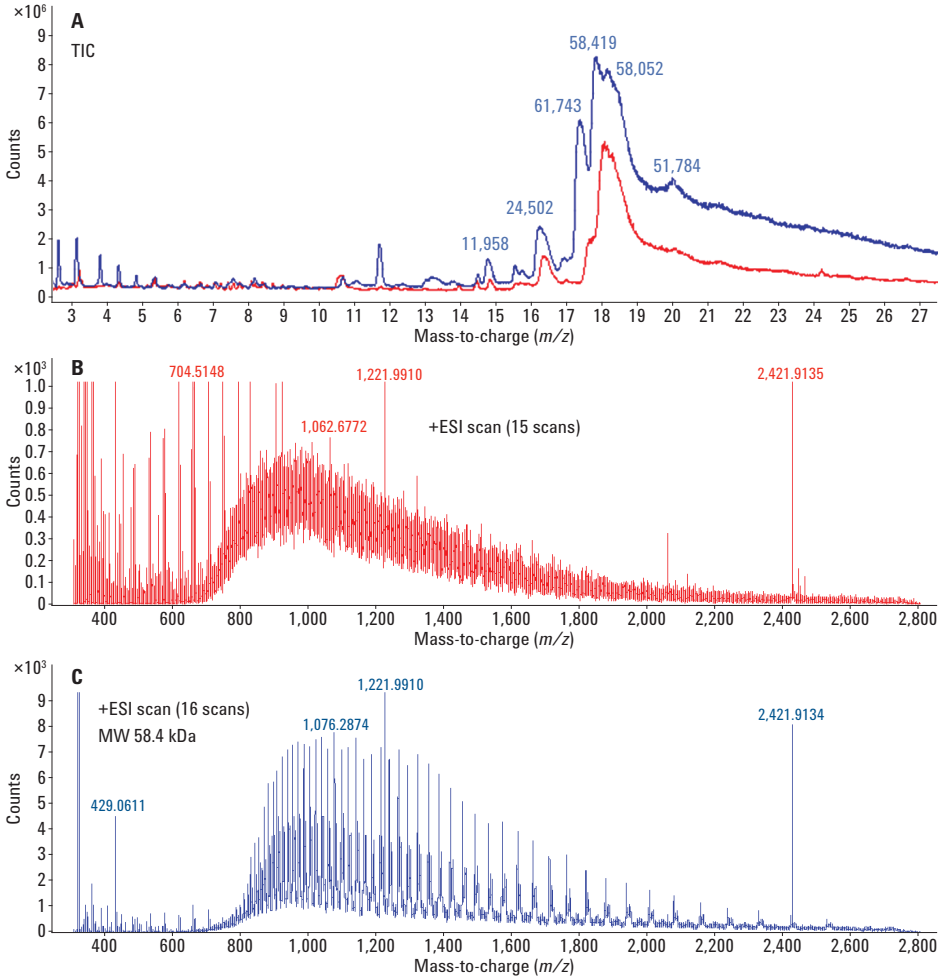


Figure 1. Comparison of protein detected in by Q-TOF LC/MS roasted (red) and unroasted (blue) peanuts. A) LC-MS TIC protein profiles change when peanuts are roasted. B) Mass spectrum of roasted peanut protein at 17.9 minutes. No molecular weight was obtained due to the lack of distinct peaks. (C) Mass spectrum of unroasted peanut protein at 17.9 minutes with a molecular weight of 58,419 Da.

Table 3. Changes in Molecular Weight of the Predominant Protein in Nuts Due to Roasting

Nut	Roasted (Da)	Native (Da)
Almond	60225*	53298
Brazil	51011	49732
Cashew	62293*	50192
Hazelnut	52741*	51601
Macadamia	79383*	57996
Peanut	60536*	58419
Pecan	58069*	51228
Pine	70319*	48012
Walnut	61979*	33093

* < 1,000 Da spread in molecular weight

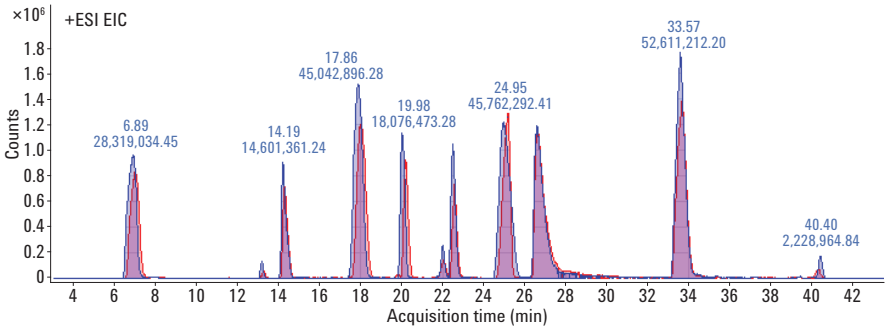
Peptide marker discovery

The selection of peptide markers to uniquely represent each nut protein required a systematic evaluation of peptides formed. The workflow for the selection of marker peptides was:

1. Tryptic digestion, Q-TOF LC/MS detection, and Spectrum Mill search of the NCBI nr database to identify peptide sequences representative of the nut protein. Selected peptides were screened for potential problems with stability including hydrolysis of aspartic acid, oxidation of methionine, and deamidation of glutamine under basic conditions.
2. Selection of (A) the most abundant peptides representative of the nut protein and (B) peptides for which there is less than 20% difference between roasted and unroasted samples.
3. A sequence search of the NCBI nr database verified that the sequences were unique to the nut analyzed and absent in all other nuts and green plants.
4. The identity of the potential marker peptide was confirmed using a synthetic peptide standard.

In step 1, initial peptide selection was limited to those resulting from trypsin digestion as identified by a Spectrum Mill search of the Q-TOF LC/MS data against the NCBI nr plant database. Due to its high mass accuracy, the Q-TOF LC/MS instrument provided a significant advantage during the initial database searching. Reducing mass-tolerance matching errors from 1 Da to less than 20 ppm for parent ions and 50 ppm for product ions greatly reduced the number of erroneous hits obtained when searching a database. For example, the total number of hits obtained from a search of peanut peptides was reduced from 1,190 (22% peanut protein related peptides) to 339 (72% peanut protein related peptides) when the search window was reduced from 1 Da to 20 ppm. Measured mass accuracies for the parent ions were within 2 mmu of the exact masses.

For step 2, Figure 2 shows extracted ion chromatograms for 10 peanut peptides in roasted and unroasted nuts. All of the peptides showed at least 20% area repeatability except for the peptide at 40.4 minutes. The two lowest intensity peptides at 22.2 and 40.4 minutes were dropped from the list due to low or nonmatching intensities for roasted and unroasted peanuts. Peptides greater than 20 amino acids in length were dropped due to cost to synthesize, and their tendency to spread ion current over more multiply-charged ions. Peptides fewer than six amino acids were not chosen due to lack of specificity. To provide broader protein coverage, one or two peptides from different allergen proteins were selected instead of multiple peptides from a single protein, when possible.



RT (min)	Peptide
6.9	AHVQVVDSNGDR
14.2	VLLEENAGGEQEER
17.9	SPDIYNPQAGSLK
20.0	SFNLDEGHALR
22.2	GTGNLELVAVR
22.6	KGSEEEGDITNPINLR
25.0	FNLAGNHEQEFLR
26.0	WLGLSAEYGNLYR
33.6	TANDLNLLILR
40.4	TDSRPSIANLAGENSIIDNLPEEVVANSYR

Figure 2. Comparison of intensities of tryptic peptide markers found in roasted (red) and unroasted (blue) peanuts.

Figure 3 also illustrates the peptide marker selection process. The absolute intensities and the close match of the peptides found in roasted and unroasted almond points to the choice of the four marker peptides indicated by the arrows. Many new peptides discovered in roasted almonds were not found in unroasted almonds, indicating that variation occurs due to roasting. Other nuts analyzed showed similar variations. Typically, there were 40 to 90 new peptides unique to a roasted nut.

Because nuts may be mixed with other nuts and grains during food preparation, a useful assay must not show interferences in food matrices. Therefore, step 3 is to verify the peptide markers that are absent in other nuts and grains after tryptic digestion. In this case, extracted ion chromatograms for the four peanut peptide markers (red trace) were overlaid on extracted ion chromatograms of the same masses of other nut and grain digests listed (black traces) (Figure 4). All traces were normalized to the same vertical intensity. Traces overlaid with the peanut digests included: almond, brazil nut, cashews, hazelnut, macadamia, pecan, pine nut, pistachio, and walnut, and the grains: barley, corn, rice, quinoa, soy, and wheat. Because the peanut peptide markers were not found in the samples, the method is specific to peanuts and has no cross-reactivity, nor interference from other nut or grain products.

To further confirm uniqueness, each target peptide sequence was searched against the NCBI nr plant database. The complementary research [9] lists the peptide markers selected for each nut, including retention time, parent ion, and two confirmatory product ions. Two peptide markers were selected for each tree nut, and four were selected for peanuts.

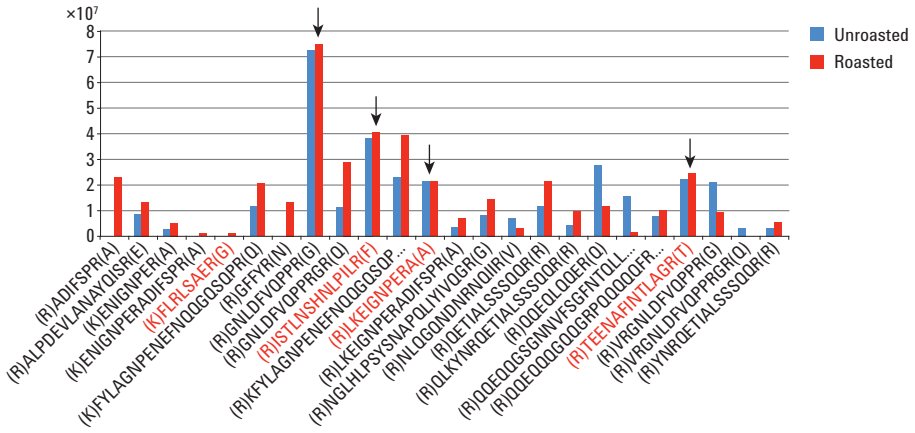


Figure 3. Intensity of response for tryptic peptides from roasted and unroasted almond proteins. The peptides selected (shown by arrows) represented the allergenic protein Pru du 6 (molecular weight 63,016) or prunin (Prunus dulcis) (molecular weight 62,979).

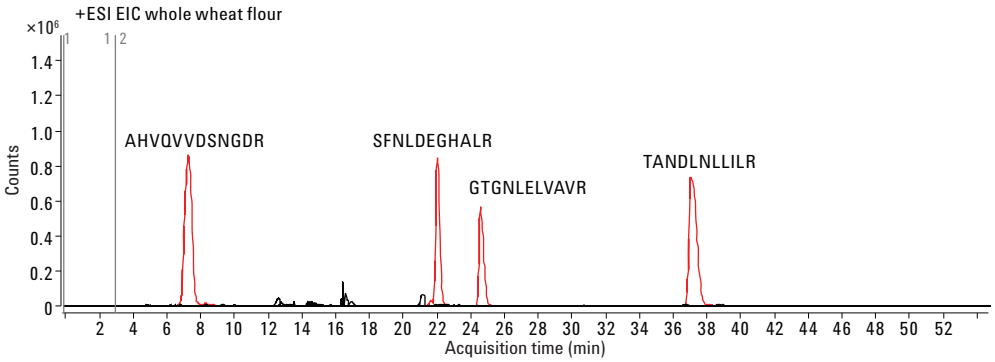


Figure 4. Extracted ion chromatograms (10-ppm window) of the four peanut peptide markers (red trace) overlaid on extracted ion chromatograms of the same masses of the nuts (almond, brazil nut, cashews, hazelnut, macadamia, pecan, pine nut, pistachio, and walnut) and grains (barley, corn, rice, quinoa, soy, and wheat) (black traces). All traces were normalized to the same vertical intensity. The extracted ion chromatograms showed no common ions in that would interfere with the selected peanut peptides.

Detecting nut peptide markers in foods

The peptide markers were tested to determine if they could be used to screen and quantitate food products for the presence of 12 nuts, rapidly in one assay. Detection of the marker peptides was based on retention time, accurate mass within a 10-ppm window, and two confirming MS/MS product ions.

A variety of foods that contained nuts, did not contain nuts, and might contain nuts (because they were manufactured in facilities that also processed nut products) were analyzed. An example of the results is shown in Figure 5 where peanut marker peptides were detected in unroasted peanuts, peanut cookies, and a protein bar, but not in a corn blank.

The complementary study [9] presents a complete list of the foods analyzed and the nuts detected in each. Grains and other foods that should not have contained nuts did not produce false positive responses. Foods that did contain nuts, as stated on the label, showed a positive response for the correct nut. Because many of the food products were processed in facilities that also process tree nut containing foods, other nuts were occasionally detected.

The most significant results were from analyses of foods whose labels did not list specific nuts, but did use a precautionary statement that the product was processed or manufactured with equipment used to process other nuts (Table 4). Macadamia nut cookies showed low intensity peaks for almond, pecan, and walnut marker peptides, in addition to intense macadamia nut marker peptide peaks. Banana cream cake is an example of a food where the label states *processed* or *manufactured on shared equipment with other tree nuts*, but no nut markers were detected. Peanut butter pretzels contained peanuts as stated, but also showed a weak signal for walnuts, which were not listed on the label, nor did the label state that the product was processed on equipment used to process other nuts.

Conclusion

Though hidden peanut and tree nut allergens in foods can result in serious health issues, foods processed using equipment also used to process other tree nuts or peanuts occasionally contain these other nuts. Key to discovering peptide markers that can be used to detect nuts in a variety of foods is verification that, (1) the marker peptide sequences are conserved after cooking and other processing, and (2) the sequences are not found in other nuts or plant grains used in food.

The marker peptides discovered through trypsin digestion of unique nut proteins, analysis by accurate mass Q-TOF LC/MS and NCBI nr database searching enabled the development of a method able to screen raw and roasted nuts, and processed and unprocessed food samples for the presence of 12 nuts in just one assay. The peptide discovery process minimized chances of cross-reactivity and the false positive or negative responses encountered with popular ELISA methods. The Q-TOF LC/MS method presented here was shown to be a tool that could be used to improve the characterization and labeling of food products, and the detection of trace levels of food allergens.

Table 4. Selected Foods and Ingredients Analyzed Highlighting Nut Marker Peptides Found at Greater than 0.1 ppm

Food or ingredient	Peanuts	Almonds	Pecans	Pine nuts	Cashews	Brazil nuts	Hazelnuts	Walnuts	Macadamia nuts	Pistachios	Chestnuts	Coconuts
Banana nut muffin			■				■	■				
Pecan cookie			■									
Banana cream cake												
Peanut butter pretzel	■							■				
Gluten-free crackers			■				■					
Peanut cookie	■											
Macadamia nut cookie		■	■					■	■			
Nut crisp		■	■									

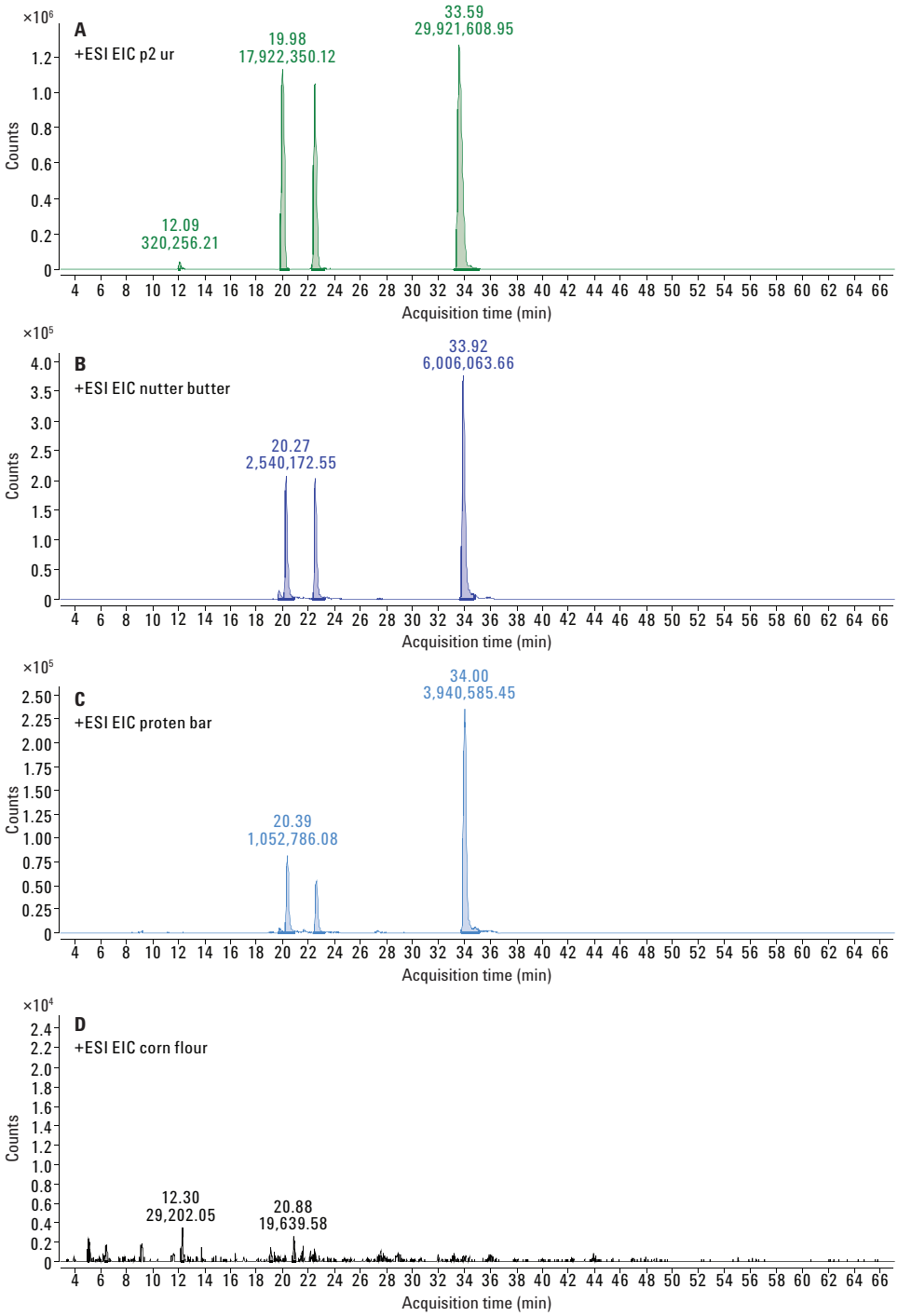


Figure 5. Extracted ion chromatograms of three peanut marker peptides in food. A) unroasted peanuts, B) Nutter butter cookies, C) protein bar, and D) corn blank. The marker peptides are SFNLDEGHALR at a retention time of 20 minutes, GTGNLELVAVR at a retention time of 22.4 minutes, and TANDLNLLILR at a retention time of 33.8 minutes.

References

1. J. F. Crespo, *et al.* Food allergy: nuts and tree nuts. *British Journal of Nutrition*, **96(S2)**, S95 (2007).
2. S. H. Sicherer, H. A. Sampson. Food allergy. *J. Allergy Clin. Immunol.*, **125(2)** Suppl S116-125 (2010).
3. S. A. Bock, A. Muñoz-Furlong, H. A. Sampson. Further fatalities caused by anaphylactic reactions to food, 2001-2006. *J. Allergy Clin. Immunol.*, **119(4)**, 1016-1018 (2007).
4. R. E. Poms, C. L. Klein, E. Anklam. Methods for allergen analysis in food: a review. *Food Addit. Contam.*, **21(1)**, 1-31 (2004).
5. A. J. van Hengel. Food allergen detection methods and the challenge to protect food-allergic consumers. *Anal. Bioanal. Chem.*, **389(1)**, 111-118 (2007).
6. T. Cucu, *et al.* Detection of Hazelnut in Foods Using ELISA: Challenges Related to the Detectability in Processed Foodstuffs. *Journal of AOAC International*, **95(1)**, 149-156 (2012).
7. C. K. Faeste, *et al.* Liquid chromatography and mass spectrometry in food allergen detection. *J. Food Prot.*, **74(2)**, 316-345 (2011).
8. G. Picariello, *et al*. The frontiers of mass spectrometry-based techniques in food allergenomics. *J. Chromatogr. A*, **1218(42)**, 7386-7398 (2011).
9. J. Sealey-Voyksner, J. Zweigenbaum, R. Voyksner. Discovery of Highly Conserved Unique Nut and Tree Nut Peptides by LC/MS/MS for Multi-Allergen Detection. *Food Chemistry* (2015) DOI: 10.1016/j.food-chem.2015.07.043.

For More Information

These data represent typical results. For more information on our products and services, visit our Web site at www.agilent.com/chem.

www.agilent.com/chem

Printed in the USA
August 4, 2015
5991-6095EN

Agilent Technologies