The Analysis of Unintended Open Reading Frame ORF-130 Expression in Maize Event MZIR098 by LC-MS/MS as Part of the Allergenicity Risk Assessment of Genetically Modified Crops

Ivan Birukou[a1], Kathryn Bailey[b], James Sugar[a], Scott McClain[c], and Scott Young[a]

[a] Syngenta Seeds, LLC, P.O. Box 12257, 9 Davis Drive, Durham, NC 27709, USA
[b] Syngenta Crop Protection, LLC, P.O. Box 12257, 9 Davis Drive, Durham, NC 27709, USA
[c] SAS Institute, Inc, 100 SAS Drive, Cary, NC 27513, USA.

ABSTRACT

Rigorous safety assessments are required for genetically modified (GM) crops to support their use as food or feed. These assessments include bioinformatics analysis of the amino acid sequences of all potential open reading frames (ORFs) encoded in the exogenous DNA introduced into the plant through genetic modification. Sequence similarities of potential DNA products to known or potential toxins and allergens trigger subsequent assessments of potential hazards for the consumer or the environment. The bioinformatics analysis of Syngenta maize event MZIR098 insert DNA revealed a total of 415 ORFs. One putative ORF (*orf-130*[2]) triggered further risk assessment based on its sequence similarities with known allergens, including peanut allergen Ara h 1, wheat glutenin, and others. We examined the potential expression of *orf-130* at a protein level to exclude the possibility of exposure during consumption of MZIR098 maize. Specifically, we developed and validated a liquid chromatography coupled to a tandem mass spectrometry (LC-MS/MS) method, called parallel reaction monitoring (PRM). This measure of expression was quantitative and targeted against the peptides unique to ORF-130 amino acid sequence. The linear range of the method was 25-5000 amol ORF-130/µl. Both, limit of detection (LOD) and lower limit of quantitation (LLOQ) of the method, were 12 µg ORF-130 per gram of dry MZIR098 maize leaf. The method was accurate and specific to ORF-130. PRM analysis of MZIR098 extracts did not detect putative ORF-130 above the limits of detection/quantitation. Furthermore, the broad proteomic data dependent acquisition (DDA) analysis of the MZIR098 extracts demonstrated no evidence of ORF-130 corroborating the PRM analysis. Our results confirm that *orf-130* is not characterized as a functional gene in Syngenta maize event MZIR098 insert DNA and does not express gene products at detectable levels.

Key Words: parallel reaction monitoring, open reading frame, risk assessment, GM crops, LC-MS/MS, spectroscopy

Introduction

Traditionally, genetically modified (GM) crops are developed using recombinant DNA technology, in which one or more exogenous genes are introduced into the plant genome. The expression of these genes confers

[1] Corresponding Author: Ivan Birukou. Email: ivan.birukou@syngenta.com
[2] In this article, *orf-130* refers to a DNA sequence of the open reading frame ORF-130, while ORF-130 refers to the putative protein ORF-130, a product of *orf-130* expression.

desirable traits to the plant, such as insect resistance or herbicide tolerance. The food and feed safety goals focus on conferring these desirable phenotypes without introducing potential hazards for consumers or the environment.

To be approved for cultivation and consumption, the GM plant must undergo consumer and environmental safety assessments. The consumer assessment relies on comparisons of the GM plant to its parental, non-GM counterpart, which has a long history of safe use for consumers. It aims to establish that the GM plant is as safe as its conventional counterpart, with the only difference being the introduction of the intended traits; i.e., protein(s) (Codex, 2009; Privalle et al., 2012). The newly expressed transgenic proteins also undergo a safety assessment, which includes an assessment of the mode of action of proteins, history of safe use in food and feed, biochemical characteristics (*e.g.*, stability under processing/cooking conditions and conditions found in digestive tract), and bioinformatic approaches that evaluate its amino acid sequence similarity to known/potential toxins and allergens (Codex, 2009; Delaney et al., 2018). In addition to the intended trait protein genes, the presence of any unintended, putative open reading frames (ORFs) introduced along with the insert DNA must also be evaluated, specifically, the possibility that these ORFs would encode proteins of potential harm if they were to be expressed. Therefore, all putative ORFs derived from the insert DNA and the junction sequences between the insert and the flanking genomic DNA are evaluated using bioinformatic approaches for amino acid sequence similarity to known toxins and allergens (Codex, 2009). This holistic approach to the molecular characterization of inserted trait DNA is the foundation of the safety assessment. Any ORFs that demonstrate sufficient sequence similarities are further investigated to determine if translation and subsequent exposure upon consumption of the GM plant are possible (EFSA, 2010; Ladics et al., 2011).

Event MZIR098 maize developed by Syngenta contains a single, contiguous insert which includes three transgenes, *ecry3.1Ab*, *mcry3A*, and *pat*, and confers control of three of the major corn rootworm pests in North America (Walters et al., 2020). Sequence analysis of the MZIR098 insert DNA and the junction regions identified a total of 415 ORFs (in these analyses ORF is defined as DNA sequence contained between two stop codons (Kobayashi et al., 2002; Ladics et al., 2011; Riechmann et al., 1999)), ranging in length from 8 to 659 amino acids. When searched against the COMprehensive Protein Allergen REsource (COMPARE) allergen database (COMPARE, 2017), a curated, peer-reviewed database of clinically relevant known/putative protein allergens (https://comparedatabase.org/), one of the identified MZIR098 open reading frames, ORF-130, generated sequence alignments with widespread allergens, such as seed storage protein Ara h 1 from *Arachis hypogaea*, Pru Du Amandin and prunin 1 from plum, bovine collagen alpha-2(I) chain and wheat glutenin. The identified bioinformatics alignments triggered downstream analyses of potential exposure of a consumer to the putative ORF-130 protein.

While molecular analyses, such as RT-PCR or Northern blot, can be used to demonstrate presence or absence of a transcript of an unintended ORF in question, they can carry a risk for false positives or ambiguous results and may trigger a follow-up assay to confirm the presence/absence of the corresponding unintended protein. Analyzing potential expression of a putative ORF on the protein level is the preferred approach.

There are a few methods that could be applied to analyze proteins of interest in GM crops,

among which the enzyme-linked immunosorbent assay (ELISA) (Grothaus et al., 2006; Zhang et al., 2014) and Western blot (Taylor et al., 2013; Taylor & Posch, 2014) are the most frequently used. However, both ELISA and Western blot are highly dependent on the availability of a purified target protein, required for (i) generation of antibody with desired specificity, (ii) to be used as standard, and (iii) positive control within the assay. Liquid chromatography/tandem mass spectrometry (LC-MS/MS) techniques are a great alternative to antibody-based approaches described above (Aebersold & Mann, 2003; Lange et al., 2008; Pitt, 2009; Vidova & Spacil, 2017). Targeted LC-MS/MS quantitative techniques differ from the antibody-based methods in that the intact form of the target protein is not analyzed. Instead, the sample is digested with the appropriate protease (typically trypsin), and selected proteotypic peptides unique to the target protein are monitored (Liebler & Zimmerman, 2013). Thus, the full-length target protein is not necessary for the method development because synthetic proteotypic peptides can be used as standards and controls, which are much easier to generate in necessary quantities, especially when the target protein is unlikely to be expressed. Parallel reaction monitoring (PRM) is one of the more recent targeted LC-MS/MS quantitative methods (Peterson et al., 2012; Rauniyar, 2015). It is typically performed on a hybrid mass spectrometer, such as quadrupole-orbitrap (Q-OT) (Bourmaud et al., 2016; Michalski et al., 2011). During the first step of PRM analysis, a proteotypic peptide is selected in the quadrupole section of mass spectrometer which filters ions based on their mass-to-charge ratio. After that, the selected peptide is fragmented in the collision cell *via* higher energy collision-induced dissociation (HCD) (Olsen et al., 2007). Finally, all resultant fragment ions of the target peptide are analyzed in the orbitrap mass analyzer simultaneously during the last step of the PRM duty cycle, and the combined intensities of the selected fragment ions are used to quantitate the amount of the target peptide, and the corresponding protein in question, by extension. High mass accuracy and resolution (e.g. (Bekker-Jensen et al., 2020)) makes PRM a highly specific method for analyses of complex mixtures such as plant extracts.

Targeted LC-MS/MS methods suffer from instrumental limitations due to which only a small set of peptides from the protein of interest can typically be monitored. In contrast, a discovery proteomics approach called data dependent acquisition (DDA, (Aebersold & Mann, 2016)) can, theoretically, detect a significantly larger number of peptides from the protein in question, if not all of them. The downside of this method is that it typically does not quantitate the identified proteins, but rather provides the evidence of the presence/absence of a protein. However, the combination of DDA and PRM methods should provide sufficient evidence in the analysis of the expression of the unintended ORFs.

In this work, we describe an LC-MS/MS approach to evaluate the risk of consumer exposure to the translation product of an unintended open reading frame *orf-130* identified within the MZIR098 insert DNA. We developed and validated a targeted quantitative PRM method to demonstrate the lack of expression of putative ORF-130 in MZIR098. In addition, we applied a qualitative proteomic DDA approach to evaluate protein expression in MZIR098 maize broadly and look for the evidence of ORF-130 peptides not targeted by the PRM method.

Materials and Methods

Materials

The test material used in the LC-MS/MS analyses was lyophilized transgenic maize leaf tissue (growth stage V6-V8, leaf collar method, (Abendroth et al., 2011)) from the event MZIR098 that contains crystal insecticidal proteins mCry3A and eCry3.1Ab, as well as phosphinothricin acetyltransferase (PAT) protein. The control plant material was lyophilized non-transgenic maize V6-V8 leaf tissue, which is non-transgenic, near-isogenic to MZIR098 maize and was used to generate standard and blank samples.

Two surrogate non-labeled (light) peptides that are unique to the putative ORF-130 protein, APVPQGGEDR (MW = 1025 g/mol) and GPLEGLHGAGGGPDGPEDR (MW = 1788 g/mol), were used in this study as standards, as well as their corresponding stable isotope-labeled (SIL, heavy) counterparts, APVPQGGED[C13N15-HeavyR] (MW = 1035 g/mol) and GPLEGLHGAGGGPDGPED[C13N15-HeavyR] (MW = 1798 g/mol). The synthetic peptides were supplied by New England Peptide (Gardner, MA) and were stored at -20°C until use.

Analysis of ORF-130 overexpressed in *E. coli* by DDA

The DNA sequence for the putative ORF-130 (the sequence did not include the first 16 amino acid residues and starts with the first Met residue at position 17, as this would be the most likely translation product of putative *orf-130*) was synthesized and cloned into plasmid pET29a (MilliporeSigma, Burlington, MA), transformed into *Escherichia coli* BL21* (DE3) cells (Thermo Fisher Scientific, Rockford, IL) and overexpressed following isopropyl β-D-1-thiogalactopyranoside (IPTG, Teknova, Hollister, CA) induction. The *E. coli* lysate containing transgenically expressed ORF-130 polypeptide was spiked into the leaf extract of non-transgenic near-isogenic to MZIR098 maize at 1:200 ratio and digested using trypsin following the procedure described below in "Extraction and processing of plant material for PRM and DDA" section. The resultant spiked maize extract digest was desalted by solid-phase extraction utilizing C18 reverse phase resin, as described below in "Sample desalting" section, to minimize matrix effects or interferences and reduce ion suppression. The sample was subjected to data dependent acquisition analysis using Thermo Scientific™ EASY-nLC 1200 coupled to a Thermo Scientific™ Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Rockford, IL) using DDA parameters described in "LC-MS/MS analysis" section. The obtained MS/MS spectra were analyzed using MaxQuant™ 1.6.2.10 software (Cox & Mann, 2008) and searched against B73 maize proteome as well as *E. coli* proteome which included the sequence of ORF-130 polypeptide. The search parameters are described in "LC-MS/MS analysis" section.

Extraction and processing of plant material for PRM and DDA

Crude extracts of MZIR098 and non-transgenic, near isogenic maize leaf tissue were prepared by combining 20 mg of lyophilized tissue with 2 ml of phosphate buffered saline (PBS, Sigma-Aldrich, Inc., St. Louis, MO) with 0.1% RapiGestSF surfactant (Waters Corporation, Milford, MA) and homogenizing the mixture using the Omniprep™ homogenizer (Omni International, Kennesaw, GA). Extracts were then centrifuged for 15 minutes at 25,000 x g and 4°C to remove insoluble material. The supernatants were transferred to a fresh LoBind Eppendorf vial (Eppendorf North America, Enfield, CT) and diluted 9.8-fold

with PBS with 0.1% RapiGestSF. The diluted extracts were mixed with an equal volume of 2,2,2-trifluoroethanol (TFE, Sigma-Aldrich, Inc., St. Louis, MO) and incubated with shaking at room temperature for 30 minutes in Eppendorf Thermomixer (Eppendorf North America, Enfield, CT). After that, 300 µl of extract was mixed with 1200 µl of 100 mM ammonium bicarbonate and 30 µl of 0.1 µg/µl trypsin (Promega, Madison, WI) at 37°C for 18 hours while shaking in Eppendorf Thermomixer (Eppendorf North America, Enfield, CT). After digestion, the extracts were acidified by adding 170 µl formic acid (FA, Thermo Fisher Scientific, Rockford, IL). Finally, for PRM samples 10 µl of fortification solution containing 10X concentrations of both non-labeled and SIL peptides (non-zero standards) or just SIL peptides (test samples and zero standards) were added to 190 µl of the acidified extracts. The PRM analysis samples were prepared by spiking the MZIR098 extract digests with SIL peptides at 200 amol/µl. DDA samples were the MZIR098 extract digests without the addition of non-labeled or SIL peptides.

PRM standards and double blanks

The standards which were used in PRM analysis to generate the calibration curve were prepared by fortifying the non-transgenic extract digests with seven concentrations (25, 50, 100, 200, 500, 1000, 5000 amol/µl) of each of the two non-labeled peptides and one fixed concentration of the corresponding SIL peptides (200 amol/µl). A zero standard was also included and contained the same fixed concentrations of the two SIL peptides (200 amol/µl each) and without any non-labeled peptides. Double blank PRM samples were the non-transgenic extract digests without the addition of the non-labeled or SIL peptides.

Sample desalting

PRM and DDA samples, as well as PRM standards and double blanks were desalted using Pierce™ C18 Spin Columns (Thermo Fisher Scientific, Rockford, IL) to minimize matrix effects or interferences and reduce ion suppression. Specifically, C18 columns were conditioned in methanol and washed twice in 1% FA, the binding of the sample to C18 column was achieved by applying the sample twice. The bound sample was washed twice in 1% FA and eluted twice with 80% acetonitrile/1% FA. The combined eluate was completely evaporated using CentriVap refrigerated concentrator (Labconco, Fort Scott, KS) and resuspended in 3% acetonitrile/0.1% FA before LC-MS/MS analysis.

Table 1 - *Gradient parameters for ORF-130 PRM assay*

Time, min	Duration, min	Flow, nl/min	Solvent B, %
0	0	500	0
30	30	500	35
40	10	500	35
41	1	500	90
56	15	500	90
57	1	500	0
60	3	500	0

LC-MS/MS analysis

For PRM analysis, 4 µl of sample were injected onto a Thermo Scientific™ EASY-nLC 1200 coupled to a Thermo Scientific™ Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer. Peptide separation was achieved using a ReproSil-Pur column with particle size 3 µm and column dimensions 75 µm x 25 mm (New Objective, Littleton, MA). The column was initially equilibrated with the solvent A (3% acetonitrile in 0.1% FA/water). The LC gradient is provided in Table 1. Analytes were measured in positive ion mode using the Nanospray Flex ion source. Data acquisition was performed using Thermo Scientific™ Xcalibur™ software. The MS/MS resolution was 35,000, automated gain control (AGC) target was $1x10^5$ with maximum ion time of 110 ms. The isolation window was m/z 1.6 with an offset of m/z 0.4. Normalized collision energy was 27 eV. The PRM results were analyzed using Skyline 4.2.0 software (Henderson et al., 2018).

Table 2 - *Gradient parameters for ORF-130 DDA assay*

Time, min	Duration, min	Flow, nl/min	Solvent B, %
0	0	300	0
5	5	300	0
245	240	300	35
265	20	300	35
266	1	300	100
281	15	500	100
282	1	500	0
292	10	500	0

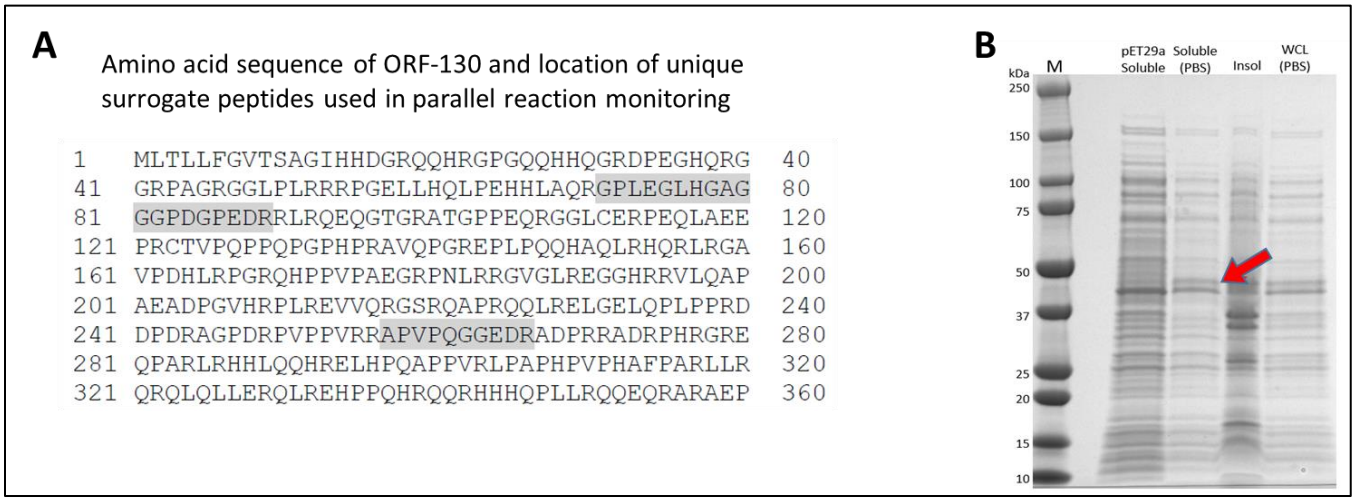
In the DDA analysis, 6 µl of sample were injected onto the same ReproSil-Pur column, which was initially equilibrated with the solvent A (3% acetonitrile in 0.1% FA/water). The LC gradient is provided in Table 2. Analytes were measured in positive ion mode using the Nanospray Flex ion source. Data acquisition was performed using Thermo Scientific™ Xcalibur™ software. The *Full MS/ dd-MS^2 (TopN)* method was used. Full MS resolution was set at 70,000, AGC target was $3x10^6$ with maximum ion time of 40 ms. The scan range was 300-1600 m/z. The data dependent MS/MS resolution was set at 17,500, AGC target was $1x10^5$ with maximum ion time of 120 ms. The scan range was m/z 200-2000. The top 10 ions were selected for fragmentation. The isolation window was m/z 1.5 with an offset of m/z 0.2. Normalized collision energy was 27 eV. The DDA results of duplicate MZIR098 samples were analyzed using MaxQuant™ 1.6.2.10 software (Cox & Mann, 2008). Specifically, the obtained LC-MS/MS data were searched against the proteome of the non-transgenic B73 maize near-isogenic to the MZIR098, which also included the sequences of transgenic PAT, mCry3A, and eCry3.1Ab proteins expressed in MZIR098, as well as putative protein ORF-130. The

search included variable protein modifications such as N-terminal acetylation and Met oxidation and up to 2 miscleavages were allowed. For the decoy database approximately 10% of the searched proteome was reverted and the false detection rate was set to 0.01.

RESULTS AND DISCUSSION

PRM method development

Quantitative LC-MS/MS methods rely heavily on the ability to identify the peptide unique to the target protein in a complex mixture and measure the area of its chromatographic peak accurately. The area of the peptide's chromatographic peak strongly depends on the MS signal response of the peptide (Picotti & Aebersold, 2012). Even though there are ways to predict, to a certain degree, the LC-MS/MS properties of peptides from their sequences (Gessulat et al., 2019; Guan et al., 2019; Searle et al., 2020; Tiwary et al., 2019), the most accurate approach to identify the peptides with the optimal MS signal response is to directly measure these properties by digesting the target protein and analyzing the resultant digest by LC-MS/MS.



Note: A. Deduced amino acid sequence of putative ORF-130 polypeptide and the location of proteotypic peptides (highlighted in grey) selected for the PRM assay. B. SDS-PAGE gel demonstrating overexpression of ORF-130 in *E. coli*. Lane "pET29a Soluble" corresponds to the soluble fraction of *E. coli* transformed with an empty vector. Lane "Soluble (PBS)" - soluble fraction of a lysate from E. coli overexpressing putative ORF-130 protein. The red arrow indicates a Coomassie-stained band corresponding to the molecular weight of ORF-130 (approx. 41 kDa). Lane "Insol" – insoluble fraction of a lysate from *E. coli* overexpressing putative ORF-130 protein. Lane "WCL (PBS)" – a whole cell lysate of *E. coli* overexpressing putative ORF-130 protein.

Figure 1 - *Amino acid sequence of ORF-130 and its overexpression in Escherichia coli*

Sequence analysis of the putative ORF-130 polypeptide suggests that it is a low sequence complexity protein (Wootton, 1994; Wootton & Federhen, 1996) with a high frequency of proline residues (15.8%, Figure 1A), which makes it likely to be unfolded and, hence, challenging to purify. To avoid the purification challenges, we overexpressed putative ORF-130 protein in *E. coli* and used the clarified lysate in LC-MS/MS method development, instead of a purified protein. Figure 1B illustrates a Coomassie stained SDS-PAGE gel of cell lysates from *E. coli* transformed either with a vector encoding ORF-130 DNA sequence or an empty vector. Compared with soluble lysate fractions of *E.*

coli transformed with an empty pET29a vector, a discernable Coomassie-stained band corresponding to the molecular weight of ORF-130 (approx. 41 kDa) was identified in the lysate fractions from ORF-130 overexpressing *E. coli* culture on the gel (Figure 1B).

Table 3 - *Summary of DDA analysis of putative ORF-130 protein overexpressed in E. coli*

Sequence	Miscleavages	Unique	Intensity Replicate 1	Intensity Replicate 2	Intensity Replicate 3
AGPDRPVPPVR	1	yes	1203700	2075500	1095200
APVPQGGEDR*	0	yes	0	675800	440630
ATGPPEQR	0	yes	0	0	1072400
GPLEGLHGAGGGPDGPEDR*	0	yes	603640	2115300	1156300
RAPVPQGGEDR	1	yes	1323400	2256800	1537300

Note: Only the five most intense peptides are provided. Miscleavages indicate the number of trypsin cut sites present within the peptide which have not been processed by the enzyme. Intensity corresponds to the chromatographic peak intensity of a peptide in the LC-MS/MS run.

* Indicates the peptides selected for further PRM method development.

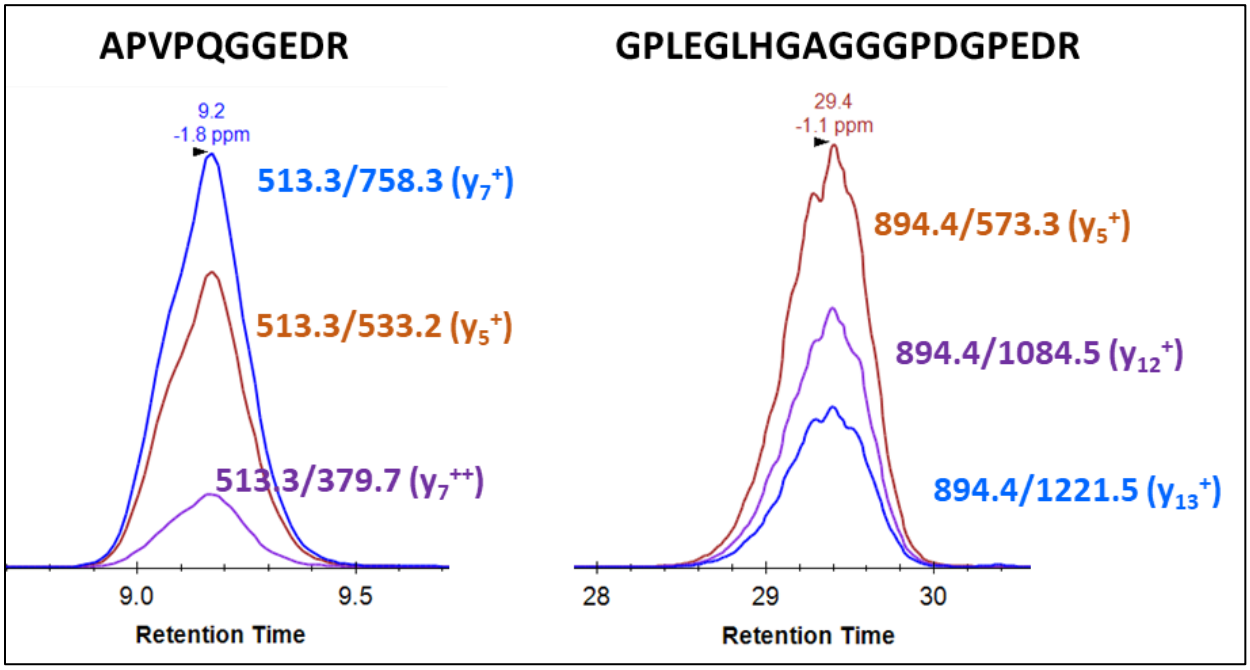
Table 3 lists the five most intense ORF-130 peptides identified by a DDA approach in non-transgenic maize extracts fortified with the lysate of *E. coli* which overexpress putative ORF-130. Among these five peptides, we identified ORF-130 peptides which i) are efficiently cleaved by trypsin, ii) are specific to ORF-130 and are not present in any of the maize proteins, iii) ionize well and can be reliably detected by Q Exactive Plus mass spectrometer in the maize proteome background even in DDA mode, and iv) conform to peptide selection criteria for PRM (peptide length of 8-25 amino acids; peptides containing Cys and Met residues and potential rugged ends, as well as peptides within 5 amino acids from N terminus were excluded). In addition, the MS/MS spectra of the identified peptides provided information on the charge states and transitions to be used in the PRM analysis.

Of the five identified peptides, APVPQGGEDR and GPLEGLHGAGGGPDGPEDR were selected as the proteotypic peptides to be used in the PRM method because they i) had no miscleavages, ii) were identified in at least 2 of 3 replicate LC-MS/MS analyses of the *E. coli* lysate containing overexpressed putative ORF-130 protein spiked into the maize leaf extract, and, more importantly, iii) come from different regions of the putative ORF-130 protein: APVPQGGEDR peptide is located near the C-terminus, while GPLEGLHGAGGGPDGPEDR peptide is located closer to the N-terminus of putative ORF-130 polypeptide (see Figure 1A, the selected peptides are highlighted in grey).

Table 4 - *ORF-130 surrogate peptides and transitions monitored in PRM assay*

Mass, Da	Peptide type	Charge state	Transitions, m/z
APVPQGGEDR			
1025	light	+2	513.3/758.3; 513.3/533.2; 513.3/379.7
1035	heavy	+2	518.3/768.4; 518.3/543.2; 518.3/384.7
GPLEGLHGAGGGPDGPEDR			
1788	light	+2	894.4/1221.5; 894.4/1084.5; 894.4/573.3
		+3	596.4/1084.5; 596.9/573.3; 596.9/817.4
1798	heavy	+2	899.4/1231.5; 899.4/1094.5; 899.4/583.3
		+3	600.0/1094.5; 600.0/583.3; 600.0/822.4

Note: In the transition notation, the number before "slash" corresponds to the mass-to-charge ratio of the target peptide and the number after "slash" corresponds to the monitored fragment ion, produced by the target peptide. Light peptides do not contain amino acids with heavy isotopes. Heavy peptides contain Lys or Arg with carbon-13 and nitrogen-15.



Note: Only 2+ charge state of GPLEGLHGAGGGPDGPEDR peptide is included for clarity. Retention times and mass accuracies are indicated at the apexes of peaks. Fragment ion types are provided in parentheses.

Figure 2 – Co-eluting transition triplets of light APVOGGEDR and GPLEGLHGAGGGPDGPEDR proteotypic ORF-130 peptides

Table 4 provides the details of the synthetic ORF-130 surrogate peptides and their corresponding SIL analogs used in PRM assay, as well as transitions used in

quantitative analysis. In the current study, SIL analogs of the target peptides were used as internal standards to compensate for any analyte signal losses (Kettenbach et al., 2011). For quantitation purposes, the three most intense transitions for each target peptide and charge state were selected (see Table 4). Figure 2 shows the examples of selected co-eluting transition triplets for light parent ions m/z 513.3 and m/z 894.4, corresponding to peptides APVPQGGEDR and GPLEGLHGAGGGPDGPEDR, respectively.

Previously it has been demonstrated that even without chromatographic separation, three transitions are sufficient to correctly identify a target peptide utilizing high resolution and accurate mass (HRAM) LC-MS methods (Peterson et al., 2012). In addition, +2 and +3 charge states of the GPLEGLHGAGGGPDGPEDR peptide were selected as both charge states demonstrated comparable signal intensity.

PRM method validation

Before application, the PRM method was validated with respect to its linearity, limit of detection (LOD), lower limit of quantitation (LOQ), and specificity. The ***linearity*** was assessed *via* linear regression analysis of a standard curve. Duplicate standards were prepared by fortifying the trypsin treated non-transgenic maize leaf extract with the addition of increasing concentrations of the non-labeled (light) peptides ranging from 25 to 5000 amol/µl and a fixed concentration of the SIL peptide (200 amol/µl). The method was considered linear if the R^2 of the calibration curve was ≥ 0.99 and at least 5 of 7 consecutive standards were found within 25% of their nominal concentrations.

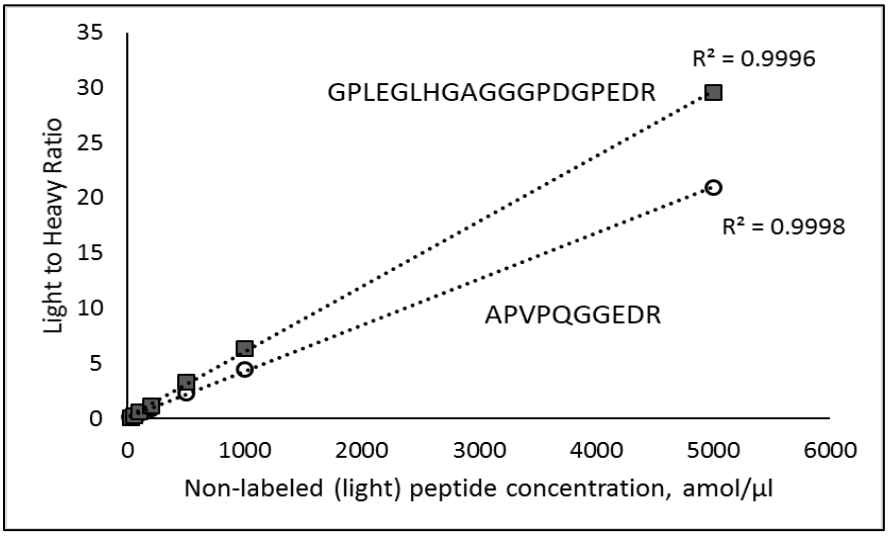


Figure 3 - *Standard curves of APVPQGGEDR (circles) and GPLEGLHGAGGGPDGPEDR (squares) proteotypic ORF-130 peptides*

Table 5 -*Linear range and LOD determination for the peptides APVPQGGEDR and GPLEGLHGAGGGPDGPEDR*

	APVPQGGEDR			GPLEGLHGAGGGPDGPEDR		
Standard, amol/µl	**MBCC, amol/µl**	**% Difference**	**Area ratio CV, %**	**MBCC, amol/µl**	**% Difference**	**Area ratio CV, %**
0	5.4	ND	138%	0.2	ND	11%
25	32.1	28%	47%	23.1	-8%	5%
50	71.9	44%	44%	49.6	-1%	0%
100	110.0	10%	3%	108.6	9%	5%
200	205.0	2%	6%	199.6	0%	11%
500	557.7	12%	0%	560.6	12%	4%
1000	1066.6	7%	12%	1082.9	8%	11%
5000	5016.6	0%	2%	5008.2	0%	3%

Note: The average MBCC values of two replicates for each peptide are shown. The % CV of the area ratios of two replicates are provided. MBCC – mean back calculated concentration of the standard; % Difference is calculated as (MBCC-Standard)/Standard. ND – not determined. CV – coefficient of variation.

Figure 3 demonstrates the calibration curves for the peptides APVPQGGEDR and GPLEGLHGAGGGPDGPEDR in maize leaf extracts. Table 5 provides the mean back calculated concentrations (MBCC) and % difference from the nominal concentrations for each standard. All seven standards of the GPLEGLHGAGGGPDGPEDR peptide satisfied the acceptance criteria for linearity and, as a result, the linear range for this peptide was determined to be 25 - 5000 amol/µl. The peptide APVPQGGEDR exhibited linear behavior in the range 100-5000 amol/µl. The two lowest concentrations of the standards (25 and 50 amol/µl) did not meet linearity acceptance criteria as their % difference from the nominal concentrations were 28% and 44%, respectively (Table 5).

The ***LOD*** describes the lowest concentration of each unique surrogate peptide that can be reliably detected in the leaf extract by LC-MS/MS. The empirical LOD was determined as the concentration of the lowest non-zero standard that yielded a MBCC greater than that of the zero standard, and for which the % CV of the mean peak area ratio of duplicate measurements was ≤25.0%. The data generated to establish the LOD for the two surrogate peptides are summarized in Table 5. For the GPLEGLHGAGGGPDGPEDR peptide, 25 amol/µl was determined as the LOD because it is the concentration of the lowest non-zero standard that yielded a MBCC greater than that of the zero standard and demonstrated mean %CV of the peak area ratio of 5% among duplicate measurements (Table 5). The LOD of APVPQGGEDR peptide was determined to be 100 amol/µl, as the % CV of the mean peak area ratio of this standard met the acceptance criteria described above, while the % CV for the 25 and 50 amol/µl standards were 47% and 44%, respectively, and thus failed to meet the LOD acceptance criteria (Table 5).

The lower limit of quantitation (***LLOQ***) is the lowest concentration at which the analyte can be reliably detected but also at which certain predefined criteria for precision and recovery (i.e. bias) are met (Armbruster & Pry, 2008). In this method we applied the following LLOQ criteria commonly used in bioanalytical method validations of ligand binding assays (DeSilva et al., 2003; FDA, 2018): i) precision is <25% CV among duplicates; ii) MBCC is within ±25% of the nominal value. In the current study, the lowest standard concentration of GPLEGLHGAGGGPDGPEDR peptide (25 amol/µl) demonstrated 5% CV in duplicate measurements and average MBCC of 23.1 amol/µl, which is only 7.6% lower than the nominal concentration (25 amol/µl) and, thus, was determined a LLOQ for this peptide (Table 5). For the peptide APVPQGGEDR, 25 amol/µl and 50 amol/µl standards produced %CV and MBCC deviation more than 25% (Table 5). The LLOQ for this peptide was determined to be 100 amol/µl, because this standard demonstrated 3% CV in duplicate measurements and the average MBCC was 110% of the nominal concentration (Table 5).

It is not uncommon that the specific MS signal response for different peptides even from the same protein can vary significantly and consequently result in different validation parameters (Picotti & Aebersold, 2012). Thus, it is crucial that the best performing peptide (with the highest signal response) is selected for quantitation purposes to obtain the highest possible sensitivity of the method. In the current study, the GPLEGLHGAGGGPDGPEDR demonstrated the best linearity within the 0 – 5000 amol/µl range, LOD and LLOQ among the two ORF-130 peptides, and, thus, was considered a quantitative peptide. As a result, both LOD and LLOQ of the method were established as 25 amol/µl, which is equivalent to 12 µg of ORF-130 per 1 g dry weight (DW) of MZIR098 maize leaf tissue (the conversion can be found in the Supplemental Figure 1).

The LOD and LLOQ for APVPQGGEDR peptide were found to be 100 amol/µl (equivalent to 47 µg/DW). Therefore, APVPQGGEDR peptide was monitored solely for qualitative purposes to confirm that the quantitative peptide is representative of the whole protein (i.e., the digestion was complete, there is no partial truncation of the protein, or other types of modification).

The ***specificity*** of the method was evaluated by assessing interference and carryover for ORF-130 surrogate and SIL peptides in double blanks and zero standards. The interference for each ORF-130 surrogate peptide was determined as the peak area ratio of zero standard divided by mean peak area ratio of the LOD sample. The interference for SIL peptides was determined as the peak area of SIL peptide in double blank sample divided by mean peak area of SIL peptide in LOD sample. To assess carryover of each ORF-130 surrogate and SIL peptide we evaluated the double blank samples injected after the highest concentration standard (5000 amol/µl). The % carryover for ORF-130 surrogate peptides was calculated as the ratio of the ORF-130 surrogate peptide peak area in a double blank to the mean peak area of the corresponding peptide in LOD sample. The % carryover for SIL peptides was calculated as the ratio of the SIL peptide peak area in a double blank to the mean peak area of the corresponding SIL peptide in 5000 amol/µl standard. It is important to note that since the double blank samples lack either ORF-130 surrogate or SIL peptides, the corresponding peak areas were measured by integrating the background noise of the double blank chromatograms using the peak boundaries of ORF-130 surrogate peptide or SIL peptide in LOD samples or highest concentration standard (5000 amol/µl). The

summary of the assessment of interferences and carryover for non-labeled and SIL peptides are provided in Supplemental Table 1. The method was considered specific as % interference and % carryover were ≤20.0% for ORF-130 surrogate peptides and ≤5.0% for SIL-peptides. In addition, the specificity of the selected peptides was confirmed by demonstrating the uniqueness of the peptides to ORF-130 protein, as the corresponding peptide sequences did not match any other protein from the complete B73 maize proteome when searched using Skyline software.

Table 6 - *Detection of putative ORF-130 protein in MZIR098 leaf extracts by PRM*

	Peak area ratio	**ORF130, µg/g DW**	**Peak area ratio**	**ORF-130, µg/g DW**
	MZIR098 Extract 1		MZIR098 Extract 2	
APVPQGGEDR	0	<LOD	0.0078*	<LOD
GPLEGLHGAGGGPDGPEDR	0.028*	<LOD	0.0026*	<LOD

Note: *Values greater than 0 are the result of integrating the matrix background, which includes minor interfering peaks and noise. LOD – limit of detection. DW – dry weight.

Detection of a putative protein from ORF-130 by PRM

Table 6 summarizes the results of the PRM analysis of duplicate MZIR098 leaf extracts. In the two MZIR098 extracts analyzed, the chromatographic peaks corresponding to the validated transitions of the quantitative ORF-130 peptide GPLEGLHGAGGGPDGPEDR and qualitative ORF-130 peptide APVPQGGEDR were not detected. Although the values of peak area ratios for peptide GPLEGLHGAGGGPDGPEDR in MZIR098 extract 1 and both peptides in MZIR098 extract 2 had non-zero values, these chromatograms did not demonstrate co-eluting transitions and the values were the result of integrating the chromatographic background noise and minor interfering peaks using the boundaries of the corresponding SIL peptide peaks (Supplemental Figure 2). In addition, all three non-zero peak area ratios were below the determined LODs for both peptides, which supports the conclusion that *orf-130* does not express a protein in MZIR098 maize leaf at detectable levels.

The mass accuracy for both light and heavy forms of both peptides was appropriate. It varied across all the analyzed standards and blank samples from -2.7 to +1.0 ppm with an average of -0.9 ppm and standard deviation of 0.8 ppm. The retention time for peptide APVPQGGEDR was very consistent, varying from 9.0 to 9.8 min across all analyzed samples. The retention time variation for the longer peptide, GPLEGLHGAGGGPDGPEDR, was much more significant: in both replicate experiments the difference between the highest and the lowest retention time was approximately 10 min.

Most likely explanation for this retention time variation is that peptide GPLEGLHGAGGGPDGPEDR eluted closer to the end of the gradient as opposed to the peptide APVPQGGEDR which elutes in the beginning of the gradient. In addition, the chromatographic separation was performed using nanoLC column at 500 nl/min flow rate reaching the upper pressure limit of 300 Bar which also contributed to the retention time instability. Nevertheless, the combination of good mass accuracy and co-elution of all

measured transitions of each peptide, co-elution of both heavy and light forms of each peptide, as well as co-elution of both charge states of GPLEGLHGAGGGPDGPEDR peptide confirm the specificity of the utilized method and support the validity of our results.

DDA analysis of complete MZIR098 proteome

In addition to quantitative analysis of the expression of putative ORF-130 protein in MZIR098 maize using the PRM method, a qualitative analysis of MZIR098 maize proteome was performed by a DDA method. In contrast to our PRM method in which only two peptides of the ORF-130 protein are analyzed, DDA analysis, albeit less sensitive than PRM, has the ability to detect ORF-130 peptides not selected for targeted analysis. In total, we were able to identify 2,497 proteins expressed in MZIR098 maize (Supplemental Data). These included transgenically expressed PAT, mCry3A, and eCry3.1Ab, for which we identified 5, 19, and 27 peptides, correspondingly. There were no ORF-130 peptides identified, further confirming the PRM results (Supplemental Data). Previously, the mean expression rates of PAT, mCry3A, and eCry3.1Ab in MZIR098 maize V6 leaves (leaf collar method, (Abendroth et al., 2011)) were determined during field trials using validated ELISA methods and were found to be 7, 76, and 200 µg/g DW, respectively (data not published), which correlate well with the total number of peptides for PAT (5 peptides), mCry3A (19 peptides), and eCry3.1Ab (27 peptides) identified during the current DDA analysis of MZIR098. It is important to note that DDA method, although being purely qualitative, was able to identify five peptides of even a low expressor such as PAT. Based on our results, the combination of PRM and DDA provide high confidence in our data and support the conclusion of ORF-130 not being expressed at detectable levels in MZIR098 maize.

CONCLUSIONS

It is not uncommon for transgenic insert DNA to contain putative ORF sequences when all six reading frames are considered. These ORF sequences are considered "unintended" because they are not part of the codon optimization necessary to express the intended trait protein(s). Recent requirements of regulatory agencies from European Union (EFSA) and Japan (MAFF, MHLW) define an ORF as not only a DNA sequence between a start and stop codon, but also between two stop codons which significantly increases the number of these putative ORFs. Although it is not uncommon to see putative DNA ORFs yield alignments with known toxins or allergens using very conservative bioinformatics methods, their theoretical amino acid sequences may not be translated and expressed in plant tissues. An assessment of expression at the protein level, as in this work, is a conservative approach that matches with the conservative consideration of theoretical protein expression present in some regulatory guidelines.

Development of a protein analysis method for unintended ORFs is complicated if these ORFs are unlikely to be expressed. As a result, it may be very resource and time consuming, if possible at all, to generate a standard and a control to use in the analysis. And, in general, proving the absence of expression is highly challenging and requires very specific and sensitive technology.

LC-MS/MS method, such as PRM, is the most suitable for the analysis of expression of unintended ORFs, compared to other protein analysis methods, such as ELISA and Western blot. The availability of purified target protein is not essential, and the method can be developed only by having the sequence of the ORF in question.

Furthermore, mass spectrometry approach can be multiplexed to allow testing of large numbers of unintended ORFs simultaneously. And finally, PRM provides the highest specificity, because the detection of the target is a direct analysis of its chemical structure (e.g., amino acid sequence). The addition of untargeted proteomic method such as DDA boosts the confidence in the results even further as it lets us to look more broadly at the proteome of the analyzed plant for the evidence of presence/absence of the unintended ORF.

In the current work, we successfully applied targeted PRM analysis of MZIR098 maize to demonstrate the absence of ORF-130 expression. The limit of detection and quantitation of the PRM method was 12 µg ORF-130 per 1 g of lyophilized MZIR098 maize leaf tissue. The method was linear over two orders of magnitude of putative ORF-130 protein concentration, with low ppm accuracy and high specificity to ORF-130, which was confirmed by negligible interferences. In addition, we applied a broad proteomic DDA approach which demonstrated thorough coverage of possible expressed gene products (approximately 2500 proteins) including intended transgenic proteins PAT, mCry3A, and eCry3.1Ab. Our results suggest that the putative *orf-130* is not a functional gene sequence, at the limits of the methods herein, and support the design and stability of the intended expressed proteins in MZIR098 maize. Thus, the bioinformatic screening results of the *orf-130* and associated alignments with allergens do not constitute a risk for a consumer. The approach herein confirms the intended molecular design of the inserted trait DNA in maize MZIR098 and demonstrates a lack of expression and exposure to theoretical peptide/protein products of putative *orf-130*.

ACKNOWLEDGEMENTS

The authors want to acknowledge Chris Fleming and Rich Sessler for their help with DDA analysis of the *E. coli* lysate containing overexpressed ORF-130.

Conflict of Interest

The authors have no conflict of interest to declare.

REFERENCES

Abendroth, L. J., Elmore, R. W., Boyer, M. J., & Marlay, S. K. (2011). Corn growth and development. *Iowa State University Extension, PMR 1009*, 49.

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature*, *422*(6928), 198-207. https://doi.org/10.1038/nature01511

Aebersold, R., & Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. *Nature*, *537*(7620), 347-355. https://doi.org/10.1038/nature19949

Armbruster, D. A., & Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *Clin Biochem Rev*, *29 Suppl 1*, S49-52. https://www.ncbi.nlm.nih.gov/pubmed/18852857

Bekker-Jensen, D. B., Martinez-Val, A., Steigerwald, S., Ruther, P., Fort, K. L., Arrey, T. N., Harder, A., Makarov, A., & Olsen, J. V. (2020). A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients. *Mol Cell Proteomics*, *19*(4), 716-729. https://doi.org/10.1074/mcp.TIR119.001906

Bourmaud, A., Gallien, S., & Domon, B. (2016). Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications. *Proteomics*, *16*(15-16), 2146-2159. https://doi.org/10.1002/pmic.201500543

Codex. (2009). *Foods derived from modern biotechnology*.

COMPARE. (2017). *Comprehensive Protein Database International Life Sciences Institute; Health and Environmental Sciences Institute Protein Allergenicity Technical Committee* http://comparedatabase.org

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*, *26*(12), 1367-1372. https://doi.org/10.1038/nbt.1511

Delaney, B., Goodman, R. E., & Ladics, G. S. (2018). Food and Feed Safety of Genetically Engineered Food Crops. *Toxicol Sci*, *162*(2), 361-371. https://doi.org/10.1093/toxsci/kfx249

DeSilva, B., Smith, W., Weiner, R., Kelley, M., Smolec, J., Lee, B., Khan, M., Tacey, R., Hill, H., & Celniker, A. (2003). Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm Res*, *20*(11), 1885-1900. https://doi.org/10.1023/b:pham.0000003390.51761.3d

EFSA. (2010). Scientific Opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. *EFSA Journal*, *8*(7), 1700.

FDA, U. (2018). Bioanalytical method validation. Guidance for industry. In. Washington, DC: US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM).

Gessulat, S., Schmidt, T., Zolg, D. P., Samaras, P., Schnatbaum, K., Zerweck, J., Knaute, T., Rechenberger, J., Delanghe, B., Huhmer, A., Reimer, U., Ehrlich, H. C., Aiche, S., Kuster, B., & Wilhelm, M. (2019). Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat Methods*, *16*(6), 509-518. https://doi.org/10.1038/s41592-019-0426-7

Grothaus, G. D., Bandla, M., Currier, T., Giroux, R., Jenkins, G. R., Lipp, M., Shan, G., Stave, J. W., & Pantella, V. (2006). Immunoassay as an analytical tool in agricultural biotechnology. *J AOAC Int*, *89*(4), 913-928. https://www.ncbi.nlm.nih.gov/pubmed/16915826

Guan, S., Moran, M. F., & Ma, B. (2019). Prediction of LC-MS/MS Properties of Peptides from Sequence by Deep Learning. *Mol Cell Proteomics*, *18*(10), 2099-2107. https://doi.org/10.1074/mcp.TIR119.001412

Henderson, C. M., Shulman, N. J., MacLean, B., MacCoss, M. J., & Hoofnagle, A. N. (2018). Skyline Performs as Well as Vendor Software in the Quantitative Analysis of Serum 25-Hydroxy Vitamin D and Vitamin D Binding Globulin. *Clin Chem*, *64*(2),

408-410. https://doi.org/10.1373/clinchem.2017.282293

Kettenbach, A. N., Rush, J., & Gerber, S. A. (2011). Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. *Nat Protoc*, *6*(2), 175-186. https://doi.org/10.1038/nprot.2010.196

Kobayashi, Y., Dokiya, Y., Kumazawa, Y., & Sugita, M. (2002). Non-AUG translation initiation of mRNA encoding plastid-targeted phage-type RNA polymerase in Nicotiana sylvestris. *Biochem Biophys Res Commun*, *299*(1), 57-61. https://doi.org/10.1016/s0006-291x(02)02579-2

Ladics, G. S., Cressman, R. F., Herouet-Guicheney, C., Herman, R. A., Privalle, L., Song, P., Ward, J. M., & McClain, S. (2011). Bioinformatics and the allergy assessment of agricultural biotechnology products: industry practices and recommendations. *Regul Toxicol Pharmacol*, *60*(1), 46-53. https://doi.org/10.1016/j.yrtph.2011.02.004

Lange, V., Picotti, P., Domon, B., & Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol*, *4*, 222. https://doi.org/10.1038/msb.2008.61

Liebler, D. C., & Zimmerman, L. J. (2013). Targeted quantitation of proteins by mass spectrometry. *Biochemistry*, *52*(22), 3797-3806. https://doi.org/10.1021/bi400110b

Michalski, A., Damoc, E., Hauschild, J. P., Lange, O., Wieghaus, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., & Horning, S. (2011). Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics*, *10*(9), M111 011015. https://doi.org/10.1074/mcp.M111.011015

Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., & Mann, M. (2007). Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods*, *4*(9), 709-712. https://doi.org/10.1038/nmeth1060

Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., & Coon, J. J. (2012). Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics*, *11*(11), 1475-1488. https://doi.org/10.1074/mcp.O112.020131

Picotti, P., & Aebersold, R. (2012). Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. *Nature Methods*, *9*(6), 555-566. https://doi.org/10.1038/nmeth.2015

Pitt, J. J. (2009). Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev*, *30*(1), 19-34. https://www.ncbi.nlm.nih.gov/pubmed/19224008

Privalle, L. S., Chen, J., Clapper, G., Hunst, P., Spiegelhalter, F., & Zhong, C. X. (2012). Development of an agricultural biotechnology crop product: testing from discovery to commercialization. *J Agric Food*

Chem, *60*(41), 10179-10187. https://doi.org/10.1021/jf302706e

Rauniyar, N. (2015). Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. *Int J Mol Sci*, *16*(12), 28566-28581. https://doi.org/10.3390/ijms161226120

Riechmann, J. L., Ito, T., & Meyerowitz, E. M. (1999). Non-AUG initiation of AGAMOUS mRNA translation in Arabidopsis thaliana. *Mol Cell Biol*, *19*(12), 8505-8512. https://doi.org/10.1128/MCB.19.12.8505

Searle, B. C., Swearingen, K. E., Barnes, C. A., Schmidt, T., Gessulat, S., Küster, B., & Wilhelm, M. (2020). Generating high quality libraries for DIA MS with empirically corrected peptide predictions. *Nature Communications*, *11*(1), 1548. https://doi.org/10.1038/s41467-020-15346-1

Taylor, S. C., Berkelman, T., Yadav, G., & Hammond, M. (2013). A defined methodology for reliable quantification of Western blot data. *Mol Biotechnol*, *55*(3), 217-226. https://doi.org/10.1007/s12033-013-9672-6

Taylor, S. C., & Posch, A. (2014). The design of a quantitative western blot experiment. *Biomed Res Int*, *2014*, 361590. https://doi.org/10.1155/2014/361590

Tiwary, S., Levy, R., Gutenbrunner, P., Salinas Soto, F., Palaniappan, K. K., Deming, L., Berndl, M., Brant, A., Cimermancic, P., & Cox, J. (2019). High-quality MS/MS spectrum prediction for data-dependent and data-independent acquisition data analysis. *Nat Methods*, *16*(6), 519-525. https://doi.org/10.1038/s41592-019-0427-6

Vidova, V., & Spacil, Z. (2017). A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition. *Anal Chim Acta*, *964*, 7-23. https://doi.org/10.1016/j.aca.2017.01.059

Walters, F. S., Young, S., & Graser, G. (2020). Meeting technical challenges for protein characterization and surrogate equivalence studies that resulted from insecticidal protein co-expression in maize event MZIR098. *Transgenic Res*, *29*(1), 109-124. https://doi.org/10.1007/s11248-019-00183-w

Wootton, J. C. (1994). Non-globular domains in protein sequences: automated segmentation using complexity measures. *Comput Chem*, *18*(3), 269-285. https://doi.org/10.1016/0097-8485(94)85023-2

Wootton, J. C., & Federhen, S. (1996). Analysis of compositionally biased regions in sequence databases. *Methods Enzymol*, *266*, 554-571. https://doi.org/10.1016/s0076-6879(96)66035-2

Zhang, S., Garcia-D'Angeli, A., Brennan, J. P., & Huo, Q. (2014). Predicting detection limits of enzyme-linked immunosorbent assay (ELISA) and bioanalytical techniques in general. *Analyst*, *139*(2), 439-445. https://doi.org/10.1039/c3an01835k

Supplemental Materials

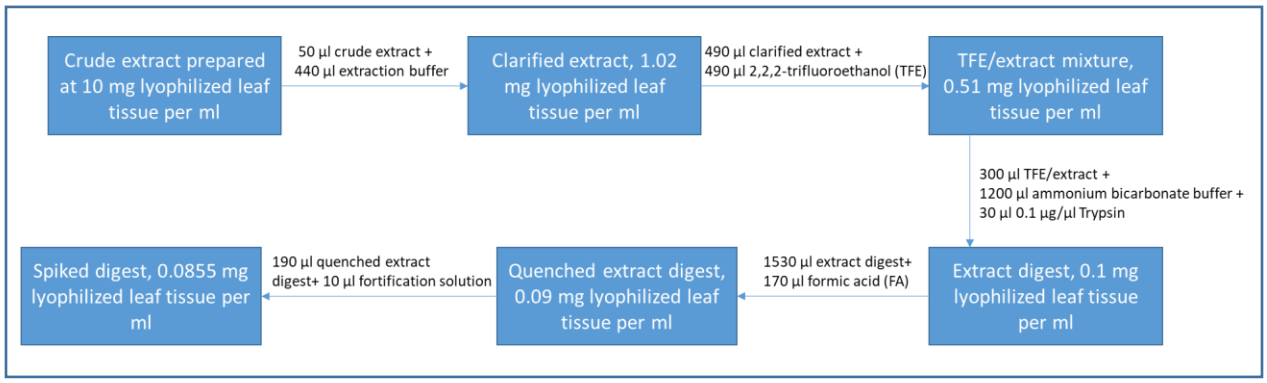


Figure S1A. Maize leaf sample extraction and digestion

$$LOD\&LLOQ\ (GPLEGLHGAGGGPDGPEDR\ peptde) = \frac{25\ \frac{amol}{\mu l}\ ORF130 \times 40606 \frac{g}{mol}\ (MW\ of\ ORF130)}{0.0855 \frac{mg}{ml}\ lyophilized\ leaf\ tissue\ in\ the\ sample} =$$

$$\frac{25 \times 10^{-12}\ \frac{mol}{l}\ ORF130 \times 40606 \frac{g}{mol}}{0.0855 \frac{g}{l}\ lyophilized\ leaf\ tissue\ in\ the\ sample} = 11.87 \times 10^{-6}\ \frac{g(ORF130)}{g\ (lyophilized\ leaf\ tissue)} = 12 \frac{\mu g(ORF130)}{g\ (lyophilized\ leaf\ tissue)}$$

$$LOD\&LLOQ(APVPQGGEDR\ peptde) = \frac{100\ \frac{amol}{\mu l}\ ORF130 \times 40606 \frac{g}{mol}\ (MW\ of\ ORF130)}{0.0855 \frac{mg}{ml}\ lyophilized\ leaf\ tissue\ in\ the\ sample} =$$

$$\frac{100 \times 10^{-12}\ \frac{mol}{l}\ ORF130 \times 40606 \frac{g}{mol}}{0.0855 \frac{g}{l}\ lyophilized\ leaf\ tissue\ in\ the\ sample} = 47.49 \times 10^{-6}\ \frac{g(ORF130)}{g\ (lyophilized\ leaf\ tissue)} = 47 \frac{\mu g(ORF130)}{g\ (lyophilized\ leaf\ tissue)}$$

Figure S1B. LOD and LLOQ conversions for ORF-130 peptides *GPLEGLHGAGGGPDGPEDR* and *APVPQGGEDR*.

Supplemental Tables.

Table S1 - Assessment of interferences and carryover for non-labeled and SIL peptides

	Interference				Carryover			
	Non-labeled		SIL		Non-labeled		SIL	
	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2
APVPQGGEDR	0%	10%	0%	0%	0%	2%	0%	0%
GPLEGLHGAGGGPDGPEDR	1%	1%	1%	1%	3%	8%	1%	1%