

Article

Development of an isoform-specific tandem mass spectrometry assay for absolute quantitation of maize lipid transfer proteins

Jay J Thelen

J. Agric. Food Chem., **Just Accepted Manuscript** • DOI: 10.1021/jf504708u • Publication Date (Web): 24 Dec 2014

Downloaded from <http://pubs.acs.org> on January 14, 2015

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



ACS Publications
High quality. High impact.

Journal of Agricultural and Food Chemistry is published by the American Chemical Society, 1155 Sixteenth Street N.W., Washington, DC 20036
Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

Development of an isoform-specific tandem mass spectrometry assay for absolute quantitation of maize lipid transfer proteins

Journal:	<i>Journal of Agricultural and Food Chemistry</i>
Manuscript ID:	jf-2014-04708u.R1
Manuscript Type:	Article
Date Submitted by the Author:	19-Dec-2014
Complete List of Authors:	Stevenson, Severin; University of Missouri, Biochemistry McClain, Scott; Syngenta Crop Protection, LLC, Thelen, Jay; University of Missouri-Columbia, Biochemistry

SCHOLARONE™
Manuscripts

Title: Development of an isoform-specific tandem mass spectrometry assay for absolute quantitation of maize lipid transfer proteins

Running title: Isovariant-specific quantitation of maize LTP allergens

Authors: Severin E. Stevenson¹, Scott McClain², and Jay J. Thelen¹

Affiliations:

¹University of Missouri, Department of Biochemistry, Christopher Bond Life Sciences Center, Columbia, Missouri 65211

²Syngenta Crop Protection, LLC, 3054 East Cornwallis Road, Research Triangle Park, North Carolina 27709

Corresponding author:

Jay J. Thelen

University of Missouri

Department of Biochemistry

109 Life Sciences Center

1201 Rollins Ave

Columbia, MO 65211

Ph: 573.884.1325

Email: thelenj@missouri.edu

Abstract

Precise and accurate quantitation of maize grain allergens is important for seed and food industries. The major allergen in maize grain is Zea m 14, a lipid transfer protein (LTP). The B73 maize genome encodes for at least six LTPs sharing between 15%-87% sequence identity to Zea m 14. Phylogenetic analysis of the maize LTP family revealed one gene that corresponds to Zea m 14 (denoted as LTPa) and two other genes sharing 43% (LTPc) and 74% (LTPb) identity with Zea m 14 that are putative homologs. Using stable isotope peptide mimics as internal standards for LTPs, we present a multiple reaction monitoring mass spectrometry approach for multiplexed, absolute quantitation of LTP proteins and alternative transcript models therein. To validate quantitative accuracy, a redundant peptide, simultaneously representing the two most abundant LTPs, was included. Analysis of 21 maize varieties revealed LTPa was most prominently expressed in maize grain, ranging from 9 to 32 μg LTP/mg protein. Proteins belonging to the LTPb and LTPc gene models were also expressed but at approximately 10- and 100-fold lower levels than LTPa, respectively. The quantitative results provided by the redundant peptide show around 95% agreement with the sum of the two unique peptides, thus providing support for the LTP gene models and validating the accuracy of this method. Though not all Zea m 14-related LTPs are abundant in grain, their high sequence homology and detectable expression in maize grain signify that LTPb and LTPc are putative allergens and should be accounted for in any quantitation strategy for maize LTP allergens.

Keywords:

LC-MS/MS, MRM, LTP, lipid transfer protein, allergen, maize, Zea mays, AQUA-MRM, quantitation

Introduction

Reports on the food allergenicity of maize grain are few despite significant maize consumption on a global basis ^{1, 2}. Some cases of allergenicity to maize have been reported ³⁻⁸; though not all cases can be confirmed as representing true food allergenicity. For example, research by Jones et al. ⁷ and separately, Pasini et al. ⁹ have found that less than half of subjects with positive skin prick tests (SPT) or CAP-FEIA (fluorescence enzyme immunoassay) to maize reacted clinically to an oral challenge. In 2000, Pastorello et al. discovered a salt-soluble IgE-binding component of maize flour, a 9 kDa protein, Zea m 14, and suggested it represents the major allergen of maize responsible for food induced allergic reactions. Zea m 14 was identified as lipid transfer protein (LTP), a ubiquitous protein belonging to the pathogen resistant group of plant proteins. Sera from 86% of patients with systemic symptoms after maize grain ingestion and positive SPT exhibited immunoreactivity to this 9 kDa protein ^{3, 10}. LTP is now acknowledged to be the major allergen in maize grain.

At the cellular level, LTPs are generally located in the apoplast and most of these proteins contain eight conserved cysteines forming four disulfide bridges, which make them highly resistant to loss of transfer function at high temperature or pH changes ¹¹. As a group, plant LTPs display a high degree of structural homology as revealed by sequence alignment and hydrophobic cluster analysis ^{12, 13}. In 1994, a unifying nomenclature for pathogenesis-related proteins (PRs) was proposed on the basis of their sequence, enzymatic, or biological activity ¹⁴. Recently this list has been extended to contain 14 families, including the PR-14-type proteins referred to as LTPs.

Expression analysis by immunoblotting of Zea m 14, based on IgG polyclonal antibody detection specific for an LTP peptide, indicated that its abundance varies in mature grain from various maize hybrids grown under different conditions. Significant differences in relative LTP abundance (up to 10-fold differences) between maize hybrids were demonstrated and occurred across a greater range (up to 20-fold) in non-irrigated plots, indicating differences in allergenic

potential^{15, 16}. The natural variation in protein allergen expression in food crops is in part due to differences in the genetic backgrounds of commercial non-genetically modified (GM) varieties, and interactions with the environment such as: nutrient and water availability, differences in moisture, temperature, plant pathogens, harvest timing, and storage conditions¹⁷. Recent work in soybean shows that both genetics and environmental differences contribute to allergen expression levels with environment playing the larger role among related breeding lines¹⁸. Although LTP from botanically related plants share IgE-binding epitopes¹⁹, it has been shown that patients allergic to LTP show a high variability in the IgE reactivity to vegetable foods, including maize¹⁵. However, at present it remains unclear how patient allergy to maize LTP may vary in response to the actual exposure levels in various maize-derived foods.

The technical objectives for developing a LTP detection method were focused on verifying the number of related LTP isoforms and whether these can be quantified individually with precision, reliability, and accuracy. Currently there is no rapid quantitative method for analyzing multiple isoforms of LTP in a multiplexed manner. Western blot analysis represents a well-established, robust, and reliable method to accurately detect specific proteins within a multitude of matrix compounds, such as those found in seeds, including endogenous allergens in their host plant matrices^{20, 21}. LTP was first characterized by IgE-binding studies using the western blot format²². Typically, LTP has been studied with human serum and detecting IgE antibody binding, as is the case with most allergens for the purpose of characterizing patient responses, as well as identifying newly discovered protein allergens. However, for the purpose of characterizing LTP content (and isoforms therein) in maize a more selective and accurate analytical approach is desirable and may be preferred for supporting further study of all related LTP isoforms and their content in foods, allergen preparations (used for clinical allergy studies), and GMO plants (as a means to identify allergen risk).

Materials and Methods

79 *Reagents*

80 All reagents were purchased from Research Products International Corp. (Mount Prospect, IL,
81 USA) unless stated otherwise. Organic solvents were purchased from Sigma-Aldrich (St. Louis,
82 MO, USA) unless stated otherwise. AQUA™ peptide standards, containing either a C-terminal
83 heavy lysine or arginine, were custom synthesized by Sigma Life Science Custom Products
84 (The Woodlands, TX, USA).

85

86 *LTP homology assessment*

87 Phylogenetic analysis was performed using the “One-click” Phylogeny.fr
88 (http://www.phylogeny.fr/version2_cgi/simple_phylogeny.cgi)²³. Pair-wise alignments were
89 performed using ClustalO on UniProtKB (<http://www.uniprot.org/uniprot/>), and percent identity
90 values were recorded.

91

92 *Maize Grain Samples*

93 The goal was to develop a LTP analytical capability using a broad sampling of maize across
94 various breeding backgrounds and production locations under varying environmental conditions.
95 Samples of grain (approximately 50 grams of whole grain) were taken as sub-samples from
96 bags of mature and harvested maize. The maize varietal characteristics and breeding histories
97 were not taken into account in the LTP analysis in this work in terms of data interpretation.
98 Supplemental Table 1 is provided to show the variety of production years, hybrid varieties and
99 production locations.

100

101 *Protein extraction*

102 Total protein was extracted from a sample subset of five maize kernels from each maize variety.
103 Five replicates were prepared for each variety. Each sample was homogenized and total
104 protein was extracted using a phenol partition protocol²⁴ modified for smaller volumes. Briefly,

five maize kernels from each replicate in each variety were weighed and pulverized for 30 sec using a 50 mL Waring blender. One hundred milligrams (+/- 5mg) were portioned for each sample, which represents a single biological replicate. Extraction buffer (1:1 mix of sucrose buffer (0.9 M sucrose, 100 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.4% β -mercaptoethanol) and Tris buffered phenol (pH 8.0) was added to each sample in a 1:60 ratio (mL buffer : mg sample). Samples were shaken for 1 hr at 4 °C using a Bead Ruptor 24 (Omni International, Kennesaw, GA, USA) on speed 1. Sample tubes were centrifuged at 17000 x g at 4 °C for 5 minutes in a benchtop centrifuge. The upper phenol phase was transferred to a new tube and was back extracted with 1 volume of fresh sucrose buffer by vortexing lightly at room temperature for 5 minutes and repeating the centrifugation step above. The phenol phase was removed and protein precipitated with five volumes of ice-cold precipitation solution (0.1 M ammonium acetate in 100% methanol).

The samples were vortexed and incubated for 16 hours at -20 °C to precipitate the protein. Protein was pelleted by centrifugation at 4000 × g for 10 min at 4 °C, and then washed twice with 1 mL of 0.1 M ammonium acetate in methanol, twice with 1 mL of ice-cold 80% (v/v) acetone in water, and once in 1 mL of 70% (v/v) ethanol in water. Protein pellets were stored in 0.5 mL of 70% (v/v) ethanol at -20 °C. Each of the five replicates was analyzed separately using mass spectrometry.

Preparation of samples for AQUA-MRM

Protein was dissolved in a chaotrope-containing buffer (8 M urea, 50 mM Tris-HCl, pH 8.0) followed by quantification using the Bradford assay (Bio-Rad Laboratories Inc, Hercules, CA, USA) with bovine gamma globulin standards (2 mg/mL, Bio-Rad). Maize protein was portioned (20 μ g/sample) and spiked with AQUA peptides (dissolved in 50% (v/v) acetonitrile (ACN), 1% (v/v) formic acid) at 50 fmol/ μ g, prior to chemical treatment. Protein was reduced by adding dithiothreitol (DTT) to 10 mM and incubated for 1 h at RT. Alkylation was performed by adding

iodoacetamide (IAA) to 50 mM and incubated at RT in the dark for 1 h. Urea was diluted and IAA neutralized by adding 3 volumes of a solution containing 10 mM ammonium bicarbonate and 10 mM DTT. Trypsin (Promega, Madison, WI, USA) was added to sample protein at a 1:49 ratio and digestion proceeded for 18 h at 37 °C. After digestion, samples were frozen at -80 °C, and evaporated to dryness using a Centrivap (ThermoFisher Scientific, Waltham, MA, USA). Dried protein digests were dissolved immediately prior to mass spectrometry analysis to a concentration of 0.40 µg/µL with 5% (v/v) ACN, 0.1% formic acid (v/v) in water.

Verification of complete protein digestion

Trypsin digests of proteins was qualitatively assessed by SDS-PAGE and staining with colloidal Coomassie Brilliant Blue dye (CBB). Briefly, Digested protein remaining from the samples used for the AQUA-MRM analysis was combined from the five sample replicates and lyophilized to dryness. Remaining undigested protein from samples prepared for AQUA-MRM analysis, was also pooled from the five sample replicates. Both undigested and digested samples were diluted or dissolved, respectively, with SDS-PAGE sample buffer to 1 µg/µL, based upon calculations prior to digestion. Twenty-five microliters (25 µg) was loaded into wells of Any kD™ Mini-PROTEAN® TGX PAGE gels (Bio-Rad Laboratories Inc, Hercules, CA, USA). PageRuler Plus (Thermo Scientific, Carlesbad, CA, USA) was used a protein standard. Gels were resolved as recommended by the manufacturer, until the dye front was ~1.0 cm from the bottom of the gel. Gels were stained overnight with CBB, and destained in water for 6 h before imaging.

Sample preparation for discovery analysis

To assess the LTP component in maize kernels, four representative kernel varieties were assayed using two complementary un-targeted “discovery” proteomics workflows: GeLC and in-solution digestion. Protein for these experiments was extracted, dissolved, and assayed for

concentration as described in *Protein extraction* and *Verification of complete protein digestion*. Fifty µg of protein from each variety was either digested in-solution (without AQUA addition) or resolved by SDS-PAGE. SDS-PAGE and in-solution digestion were performed as described in *Verification of complete protein digestion* and *Preparation of samples for AQUA-MRM*, but using Precision Plus protein standard for SDS-PAGE (Bio-Rad Laboratories Inc, Hercules, CA, USA). For in-gel digestion, the stained region between 20 kDa and 10 kDa was excised and processed as in Hajduch et al. (24)

Discovery mass spectrometry

Discovery LC-MS/MS was performed with an Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, DE). Resolution settings were 30,000 for MS1 (m/z range of 300-2000) and unit resolution for MS2 acquisitions. Data-dependent acquisition was employed using one MS scan followed by nine MS2 scans for the nine most abundant ions, with dynamic exclusion enabled. Peptides were concentrated and desalted in-line using a C₈ CapTrap, peptide trapping cartridge (Michrom Bioresources, Auburn, CA USA) prior to MS analysis. HPLC separations were performed using in-house fused silica nanospray needles, 10.5 cm length, (360 µm outer diameter, 150 µm inner diameter; Polymicro Technologies, Phoenix, AZ USA), that were packed with "Magic C18AQ" (200 Å, 5 µm particles; Michrom Bioresources) with an ACN gradient from 2-45% ACN over 80 min, followed by 17 min for column washing and equilibration. RAW files were searched using SEQUEST (v 1.20) via Proteome Discoverer (v 1.3.0.339) with three databases: UniProt/SwissProt (November, 2012), B73 genome (5b.60, maizesequence.org) and Zea maize UniGene translated transcriptome (Zm166, via NCBI) with carbamidomethylation as a fixed modification for cysteines, methionine oxidation as a variable modification, a parent ion mass tolerance of 20 ppm and a fragment ion mass tolerance of 0.80 Da. Results were filtered using mass accuracy and confidence filters of 10 ppm and high, respectively.

182 *Targeted AQUA-MRM mass spectrometry*

183 Targeted mass spectrometric analysis was performed on column with 2 µg of protein digestion,
184 5 µL loading volume. The linear range of the endogenous analyte was assessed from the
185 maximum possible loading amount, 2.4 µg/injection - to 0.0038 µg/injection by 5-fold serial
186 dilutions (Supplemental Figure 1). Autosampling, chromatography, and mass spectrometry
187 were performed as described in Stevenson et al.¹⁸ while using the application-specific transition
188 ions for the LTP AQUA-MRM (Supplementary Table 2). Analysis was conducted within the pre-
189 determined linear range of quantification for the AQUA peptides, which was determined for
190 individual transition ions (four most intense). Each LTP AQUA peptide was used as an internal
191 control at 200 fmol/injection, which is within the linear quantitative range.

192

193 *Determination of LTP concentration in biological samples*

194 Concentration of each native LTP isoform and variants therein was calculated using the
195 response ratio (native/AQUA peptide) for a single transition ion, multiplied by the concentration
196 of the AQUA internal standard. The concentration was expressed as µg LTP/mg of total protein
197 extracted from maize kernel for each LTP in each maize extract and was calculated as follows:
198 Calculation 1: [Response ratio from mass spectrometer × fmol of AQUA peptide/ 1µg of protein
199 sample = fmol of native peptide/µg of protein sample]; Calculation 2: [moles of native LTP × FW
200 (g/mol) of LTP = g LTP/µg of protein sample]; Calculation 3: [g LTP × (1e6 µg LTP/g LTP) ×
201 (1000 µg protein sample/1 mg protein sample) = µg LTP/mg protein sample]

202

203 **Results**

204 *LTP discovery proteomics*

205 Although the Zea m 14 allergen has been described as the major maize grain allergen, the
206 recognition that LTP proteins would likely be expressed as a family of isoforms provided the
207 impetus to survey and characterize the “total LTP load” that a maize variety may express. An

initial discovery experiment was conducted to assess the diversity of LTPs in representative samples. Orbitrap mass spectrometry was used to accurately identify measurable peptides for the targeted and expressed LTPs in four representative varieties of maize. Two complementary peptide preparation methodologies were used; one method utilized fractionation by SDS-PAGE while the other used un-fractionated “total” protein. Both peptide preparations were subjected to LC-MS/MS and protein identification workflows searching three databases: UniProt/Swiss-Prot, *Zea mays* B73 inbred translated genome and the *Zea mays* UniGene transcriptome. Putative LTPs were extracted from the list of discovered proteins using molecular weight filters (20 kDa) and local BLAST analysis via BioEdit software using Zea m 14 as a query sequence; 20 kDa was the upper limit as a rational filter considering the LTP proteins are not likely to exceed ~13-15 kDa. In total, thirteen LTP sequence listings (accessions) were identified from the four biological samples with high confidence, including Zea m 14 (P19656 in UniProt, Table 1). Two of the thirteen accessions, GRMZM2G010868 and GRMZM5G898755, each had two splice variants annotated in the maize genome database (www.maizesequence.org). One of them, P19656, had an isoform annotated in UniProt, P19656-2. (<http://www.uniprot.org/uniprot/>). Three others, Q2QYL2, Q7XJ39, and Q43194, were from non-maize sources, and were excluded.

LTP homology assessment

To create an AQUA-MRM assay that quantified all selected LTP paralogs to allergen Zea m 14, the ten maize LTPs discovered during our proteomic screen of various maize proteomes were grouped by their phylogenetic relationships (Figure 1A). A clear relationship was apparent between Zea m 14 and maize gene loci GRMZM2G101958, GRMZM2G010868 and GRMZM5G898755, including both presumptive splice variants for the two latter loci. In addition, sequence identity was assessed by pair-wise sequence alignments between each LTP and Zea m 14 (Figure 1B). The percent identity scores for the seven proteins that grouped closest to

Zea m 14 in the phylogram, also shared the highest sequence identity with greater than 42% identity. It should be noted that none of the maize B73 loci were 100% identical to Zea m 14, although GRMZM2G101958 was closest at 87%. Taken together, these results show that of the maize LTP candidate proteins GRMZM2G101958, GRMZM2G010868, and GRMZM5G898755 are most similar to Zea m 14, and reasonably could contribute to food allergenicity based on a level of shared identity and function homology that may lead to cross-reactivity/cross-sensitization; for well described allergens such as those from pollen this would generally be 50% shared identity or more ²⁵. To simplify nomenclature the isoforms and splice variants for LTPs with greater than 42% identity to Zea m 14 were grouped and named LTPa, b and c (Figure 1B).

Sequence assessment of LTPs

The feasibility of using AQUA-MRM as a quantitation method for these small and highly similar proteins was assessed using sequence alignment and manual sequence inspection. All isoforms and splice variants of Zea m 14, LTPa, LTPb and LTPc were aligned using Clustal Omega (Figure 2). Possible tryptic peptides were critiqued for their selectivity and appropriateness for use as AQUA-MRM standards (i.e., labeled peptides to be included in each sample analysis ²⁶). The N-terminus of all LTPs had very few tryptic cleavage sites. LTPa has a peptide that starts at residue 48 and an analogous peptide in Zea m 14 starts at residue 47. These peptides differ in the C-terminal cleavage site, with LTPa having a Lys and Zea m 14 having an Arg (GQGSGPSAGCCSGVK and GQGSGPSAGCCSGVR). Neither peptide has neighboring downstream cleavage sites or double cleavage sites that could lead to “ragged ends”, as was the case for the analogous peptide in LTPb which has a C-terminal arginine with a lysine two residues away (GTGSGPSASCCSGVRNLK). LTPb also has only one available peptide at residue 87 (GVSGLNAGNAASIPSK). LTPc has only one available peptide which starts at residue 88 (LGGGVSMANAANIPSK). LTPa and Zea m 14 both have a second

available peptide which is shared between them, which starts at residue 84 in LTPb and 83 in Zea m 14 (NAAAGVSGLNAGNAASIPSK). This second peptide is an improvement on the first peptide, as it is free from modifiable amino acids which negatively affect protein quantitation by AQUA-MRM. In addition, this second peptide can be used in parallel to measure both of these proteins (Zea m 14s and LTPa) together, while their first peptide can provide the quantities of each separately. Because AQUA-MRM quantitation yields molar quantities for target analytes, the quantities of Zea m 14 and LTPa as measured by their independent peptides can be directly compared to the signal of the second, conserved peptide (NAAAGVSGLNAGNAASIPSK; Figure 2, green). Theoretically, the quantity of the second peptide, which is shared between them, should equal the sum of the concentration of the two isoform-specific peptides (Figure 2, gray and yellow peptides). This assay design was pursued to reveal whether Zea m 14 (P19656-1) and its isovariant (P19656-2) were *bona fide* proteins, as neither is a perfect match to B73 gene locus GRMZM2G101958 (LTPa). A matrix presentation of the LTP sequences and their representative peptides is shown in Table 2. To ensure that each of these peptides is protein specific, BLAST analysis of the translated B73 genome and the *Zea mays* (Zm166) UniGene transcriptome was performed. None of the peptides had an exact match to any protein other than the expected LTPs. Finally, none of these peptides discriminate between the putative splice variants observed for each of the three gene loci, which differ at the C-terminus (Figure 2). However, resolving these variants was not necessary to devise a comprehensive, absolute quantitative assay for each of the classes of endogenous LTP allergens in maize.

Bioanalytical method validation for native and AQUA peptides

Bioanalytical method validation requires the determination of many parameters including specificity, precision, linear quantitative range, and carryover. Carryover between analytical runs was assessed during the AQUA peptide dilution series experiments, which used a fixed

amount of maize grain protein as matrix and the cognate endogenous peptide therein as the control to calculate a response ratio. A blank run between each dilution in the series was assessed for residual native peptide signal and compared to that of the previous analytical runs. Each analytical run contained 2 μg of native peptides. Background was subtracted for both analytical runs and blank runs. The resulting signal for the blank runs were divided by the sum of the blank signal and the average area from the analytical runs, and multiplied by 100 to find the percent of total signal present in the blank run. Each peptide was assessed individually. On average, all were below 0.9% except the LTPc peptide which was at 1.30%.

The linear range of native (endogenous) peptides used in these assay was determined by a six point, five-fold dilution series from 2.4 μg to 0.0038 $\mu\text{g}/\text{injection}$ (including a zero control), with duplicate injections. AQUA peptides were spiked at 40 fmol/ μL . When the log5 value of the response ratios (RRs, endogenous/AQUA) were plotted against the log5 values of the matrix amounts, R^2 values for all transition ions were between 0.9842 and 1.000, and slope values were between 1.05 and 0.885 when the zero control was excluded. The experimentally-verified linear dynamic range was nearly four orders of magnitude for each peptide, although the current instrument and column capabilities likely offer a broader range. Duplicate injections provided percent relative standard deviation (%RSD) values that averaged less than 11% across all dilutions and all peptides except LTPc which was 36% (Supplemental Figure 1).

The linear range of AQUA peptides was assessed by seven-point, five-fold dilution series from 0.5 pmol to 0.16 fmol/injection including a zero control with duplicate injections (Supplemental Figure 2). For all peptides and transition ions therein, except LTPc, this six point dilution series was linear with R^2 values ranging from 0.996 to 0.9999, when excluding the zero control. Therefore the LLOQ was 0.16 fmol for each of these AQUA peptides. For these assays, the LLOD and ULOQ exceeded the six point dilution series. However, for the LTPc peptide the LLOD was 0.16 fmol and the ULOQ was 100 fmol. The LLOQ of 0.8 fmol and linear range from 0.8 – 100 fmol offered R^2 values ranging between 0.991-0.9996 when excluding the

zero control. Considering the linear range for all AQUA peptide dilution series, the slopes ranged from 0.8873 to 1.0038 with the majority being less than 5% under 1.00 (without zero control). Duplicate injections yielded response ratio %RSDs that averaged less than 8% across all dilutions. AQUA-MRM assay parameters for all peptides are listed in Supplemental Table 2. The lower and upper limits of quantification (LLOQ/ULOQ) were assessed in terms of their surrogate proteins, by converting the molar quantity for each peptide into the gram quantity for their respective protein (Table 3). For LTPa, the LLOQ was 9.39 pg/injection (0.8 fmol). For LTPb, the LLOQ was 1.88 pg/injection (0.16 fmol). The LLOQ for LTPc was 10.66 pg/injection (0.8 fmol). For Zea m 14 and LTPa/Zea m 14, the LLOQ was 2.03 pg/injection (0.16 fmol) for both.

LTP AQUA-MRM analysis

Absolute measures of the concentration in maize grain extracts for all LTPs in each variety are graphed in Figure 3. Figure 3 (top panel) shows the higher abundance LTPa and Zea m 14, while Figure 3 (bottom) contains the quantities of the lower abundance LTPb and LTPc isoforms. In most samples, both LTPa and Zea m 14 were present at similar concentrations. However, in four of the samples (variety 1, 2, 14 and 15), one of those LTPs is below either detection or quantification limits. The third series in Figure 3 (top), labeled "LTPa/Zea m 14", shows the result which simultaneously quantifies both LTPa and Zea m 14 together via a shared peptide. In all cases where either LTPa or Zea m 14 proteins are low or undetected, the quantity of the single, representative LTP is similar to the concentration of the peptide that is common to both LTPa and Zea m 14 (i.e., LTPa/Zea m 14). In addition, for the samples where both Zea m 14 and LTPa are present in similar concentrations, the additive value of the two independent assays (each protein represented by its unique peptide) is similar to the LTPa/Zea m 14 value. When the sums of the independent LTPa and Zea m 14 are plotted with the shared peptide LTPa/Zea m 14 values, it is evident that the shared peptide represents the sum of the

individual peptides (Figure 4). The average percent error for the sum of LTPa and Zea m 14 versus the value measured by the LTPa/Zea m 14 is only 5.3%. LTPb and LTPc isoforms are less abundant with LTPb being the more abundant species (Figure 3, bottom panel). LTPc was below detection or quantification limits in five varieties, numbers 1, 2, 5, 11, and 13.

Discussion

The premise behind developing analytical approaches for measuring food allergens is to gain a better understanding of their precise expression levels (referred to as absolute) in various plant organs and to extend this technology towards their identification in processed foods and the environment. Traditionally, allergen detection and quantification methods have relied primarily on antibody-based approaches, employing either IgE antibodies derived from allergic patient sera or IgG antibodies raised in animals against individual allergens. However, it is generally understood that most antibody-based methods provide semi-quantitative, relative estimates of content and that specificity and reproducibility can be affected by matrix components in the grain and foods in which allergens are usually monitored. Using a comparative approach to determine feasible assays for soybean allergens, Pedersen et al.²⁷ found that matrix component interference could be minimized. By using a general ranking of abundance of the targeted soy within a given food type, calculated soy levels were consistent across assays (higher abundance equaled a higher level of measured soy), however "...the measured soy levels given by each of these assays were not in correlation with each other, allowing only semi-quantitative estimations of the soy content". In contrast, absolute measures of protein allergens, such as has been developed for LTP by mass spectrometry, when paired with an understanding of genetic/bioinformatic background can provide reliable concentration values for the total abundance of related isoforms *in planta*.

The sequence of the B73 maize genome, coupled with in-depth transcriptomics, provided predictive gene models for the LTP family of proteins. Amino acid sequence alignment

revealed three genes sharing between 43% to 87% amino acid identity with Zea m 14. Three additional genes shared between 15% to 29% sequence identity and clustered phylogenetically apart from the others, and were therefore not considered close homologs to Zea m 14. The most closely related gene to Zea m 14 was GRMZM2G101958 (LTPa), which due to a presumptive alternative translation termination site shared 87% identity. Therefore, it is likely that the Zea m 14 allergen arose from this gene. This is supported by the observation that one of the two reported Zea m 14 transcript variants is identical to the LTPa gene at the variable C-terminus. Three of the peptides for this multiplexed maize LTP assay were developed to verify the existence of both the reported sequence for Zea m 14 as well as the LTPa transcript. The two discriminating peptides GQGSGPSAGCCSGV(R/K) differed only in the terminating residue. A third peptide was developed to a conserved sequence for LTPa and the two variants for Zea m 14 (NAAAGVSGLNAGNAASIPSK). Quantitative analysis of 21 maize varieties revealed the existence of both the LTPa and Zea m 14 proteins, albeit with varying abundance among them (Figure 3). The levels of LTPa varied from undetectable to almost 20 μg LTPa/mg maize grain protein among the 21 varieties studied. Zea m 14 ranged from undetectable to 30 μg LTP/mg maize grain protein. To confirm the accuracy of these results the third peptide was designed against a domain conserved in all three transcripts and was included in the multiplexed assay performed for each sample. This peptide overlapped with the multiple gene variants documented for LTPa and was included in each analysis as a “spike”. The results confirmed the values for the two independent peptides, with an average percent error of 5.3%. The results from this peptide demonstrate signals from the two specific peptides were authentic and accurate as the sum was equivalent to the conserved peptide representing all three gene products in all varieties (Fig. 4). These proteins are considered to represent the bulk of LTP expression in maize grain.

LTP homologs encoded by genes GRMZM2G010868 (LTPb) and GRMZM5G898755 (LTPc) were 74% and 43% identical to Zea m 14, respectively, at the amino acid level. Each of these two gene assemblies were represented by two transcript variants exclusively variable at the C-terminus (Figure 2). Two peptides were designed to discern these two genes but not the transcript variants that arose from them (Figure 2). Analysis of the 21 maize varieties revealed both proteins are present within maize grain at variable levels. Expression of LTPb ranged from 0.5 to 2.5 μg LTPb/mg maize grain protein. LTPc exhibited a low level of expression, varying from undetectable to 0.5 μg LTPc/mg maize grain protein. Although the levels of LTPb and LTPc gene products were lower than LTPa, this is dependent upon the specific maize variety being analyzed (Figure 3). Therefore, to accurately assess the total LTP allergen “load” in maize grain it is essential to monitor all three LTP isoforms related to Zea m 14.

To our knowledge, only one other non-antibody-based method for quantifying maize LTPs has been described, which uses LC-UV/MS²⁸. In this method, LTPs were selectively extracted and analyzed intact utilizing liquid chromatography coupled to UV absorbance for quantitation. Quantitation was performed using external calibration curves with a prepared mix of LTPs standardized by amino acid analysis. Analysis of maize grain revealed an LTP concentration ranging from 58-678 $\mu\text{g/g}$ grain fresh weight from fourteen maize varieties. Using the AQUA-MRM approach we determined total LTPa (Zea m 14) concentration to be 9 to 32 μg LTP/mg protein among 21 maize varieties. Assuming maize grain is 10% protein by weight, the values from Kuppannan et al.²⁸ convert to 0.58-6.78 μg LTP/mg protein. Two LTP variants were observed in 2010 by Kuppannan et al.²⁸ and it is likely variant A is the Lys-containing version of Zea m 14 (GQGSGPSAGCCSGVK/R) denoted as LTPa in Figure 2, as determined by MS analysis of this tryptic peptide. Thus, for direct comparison to the values reported here for Zea m 14 (LTPa) it is necessary to sum the values for “LTP” and “LTP variant A” in Kuppannan et al.²⁸. This produced a range of 2.23-8.14 μg LTP/mg protein, which is close to

values reported here using AQUA-MRM. The slightly lower estimate of maize grain LTP concentration by Kuppannan et al. could be due to varietal or methodological differences between these two approaches. Variant B observed by Kuppannan et al.²⁸ was up to 20-fold less abundant than the other LTP variants, therefore it would be reasonable to conclude this represented one of the related LTP genes (i.e. LTPb or LTPc in our study) based upon values reported here. The method presented here has the advantage of added selectivity by utilizing signal derived from individual peptide ions present in each chromatographic peak, so that even co-eluting species are distinguished during quantitative analysis. In addition, each assay utilizes internal reference standards that are isotope-labeled and analyzed simultaneously which increases quantitative accuracy and reproducibility.

In summary, the maize allergen Zea m 14 is part of a family of LTPs with related gene isoforms and putative splice or translation variants. Additionally, LTPs are small, 7-10 kDa proteins and therefore contain few proteotypic peptides for quantitation in a typical bottom-up AQUA-MRM workflow. Despite these limitations we demonstrated a robust, multiplexed assay for three LTP genes and the 2-3 variants of each of these genes which varied by as little as a single amino acid. We rigorously validated this multiplexed assay for absolute quantitation from extracts of maize grain in order to provide reliable and repeatable concentration measurements in a complex seed matrix. As a control, a redundant peptide that represents the theoretical sum of two specific peptides was employed as part of the methodology. The quantitative results from this redundant peptide support the theoretical sum and thus verify the accuracy of these predictive gene models as well as the AQUA-MRM method (Figure 4). In total, this quantitative study design accounted for all LTPs related to Zea m 14 (above 42% amino acid identity). It provided the specificity, accuracy, reproducibility, sensitivity, and linear quantitative range to be employed for comprehensive Zea m 14 allergen assessment in maize with the possibility of validation with other food matrices. In the future, experimental field designs that help elucidate the natural variability in LTP expression may be useful in providing a baseline. It is becoming

well established that environmental factors and the genetic background of various crop varieties are the basis for largest percentage of variability. This variability will need to be better studied in order to work towards an understanding of exposure risk for the LTP allergen.

Abbreviations:

LC-MS/MS, liquid chromatography tandem mass spectrometry; LTP, lipid transfer protein; MRM, multiple reaction monitoring; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; kDa, kilodalton; DTT, dithiothreitol; IAA, iodoacetamide; RT, room temperature; ACN, acetonitrile; CBB, Coomassie Brilliant Blue

Acknowledgments

This research was funded by a grant from Syngenta Crop Sciences. The authors wish to acknowledge Trent Stevens for reference management.

Funding/Conflict of interest statement

The research described here was conducted through financial support to JJT from Syngenta Crop Sciences.

References

- (1) Frisner, H.; Rosendal, A.; Barkholt, V., Identification of immunogenic maize proteins in a casein hydrolysate formula. *Pediatr Allergy Immunol* **2000**, *11*, 106-110.
- (2) Metcalfe, D.; Sampson, H.; Simon, R., *Food allergy: adverse reactions to foods and food additives*. 3rd ed.; Blackwell Publishing: Malden, MA, 2003.
- (3) Pastorello, E. A.; Farioli, L.; Pravettoni, V.; Ispano, M.; Scibola, E.; Trambaioli, C.; Giuffrida, M. G.; Ansaloni, R.; Godovac-Zimmermann, J.; Conti, A.; Fortunato, D.; Ortolani, C., The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. *J Allergy Clin Immunol* **2000**, *106*, 744-751.
- (4) Tanaka, L. G.; El-Dahr, J. M.; Lehrer, S. B., Double-blind, placebo-controlled corn challenge resulting in anaphylaxis. *J Allergy Clin Immunol* **2001**, *107*, 744.
- (5) Pauls, J. D.; Cross, D., Food-dependent exercise-induced anaphylaxis to corn. *J Allergy Clin Immunol* **1998**, *101*, 853-854.
- (6) Ispano, M.; Colafrancesco, M.; Ansaloni, R.; Vighi, G.; Ortolani, C., Comparison of the results of skin prick tests, CAP system and ENEA system in the diagnosis of food allergy. *Monogr Allergy* **1996**, *32*, 181-186.
- (7) Jones, S. M.; Magnolfi, C. F.; Cooke, S. K.; Sampson, H. A., Immunologic cross-reactivity among cereal grains and grasses in children with food hypersensitivity. *J Allergy Clin Immunol* **1995**, *96*, 341-351.
- (8) Venter, C.; Skypala, I.; Dean, T., Maize allergy: what we have learned so far. *Clin Exp Allergy* **2008**, *38*, 1844-1846.
- (9) Pasini, G.; Simonato, B.; Curioni, A.; Vincenzi, S.; Cristaudo, A.; Santucci, B.; Peruffo, A. D.; Giannattasio, M., IgE-mediated allergy to corn: a 50 kDa protein, belonging to the Reduced Soluble Proteins, is a major allergen. *Allergy* **2002**, *57*, 98-106.

- 480 (10) Pastorello, E. A.; Robino, A. M., Clinical role of lipid transfer proteins in food allergy. *Mol*
481 *Nutr Food Res* **2004**, *48*, 356-362.
- 482 (11) Kader, J. C., LIPID-TRANSFER PROTEINS IN PLANTS. *Annu Rev Plant Physiol Plant*
483 *Mol Biol* **1996**, *47*, 627-654.
- 484 (12) Henrissat, B.; Popineau, Y.; Kader, J. C., Hydrophobic-cluster analysis of plant protein
485 sequences. A domain homology between storage and lipid-transfer proteins. *Biochem J* **1988**,
486 *255*, 901-905.
- 487 (13) Tchang, F.; This, P.; Stiefel, V.; Arondel, V.; Morch, M. D.; Pages, M.; Puigdomenech, P.;
488 Grellet, F.; Delseny, M.; Bouillon, P.; et al., Phospholipid transfer protein: full-length cDNA and
489 amino acid sequence in maize. Amino acid sequence homologies between plant phospholipid
490 transfer proteins. *J Biol Chem* **1988**, *263*, 16849-16855.
- 491 (14) Van Loon, L.; Pierpoint, W.; Boller, T.; Conejero, V., Recommendations for naming plant
492 pathogenesis-related proteins. *Plant Mol Biol Report* **1994**, *12*, 245-264.
- 493 (15) Ariyaratna, H.; Pramod, S.; Goodman, R., The abundance of lipid transfer protein (LTP),
494 the major food allergen of corn, varies between hybrids and growth conditions. *J Allergy Clin*
495 *Immunol* **2009**, *123*, S28.
- 496 (16) Goodman, R. E.; Panda, R.; Ariyaratna, H., Evaluation of endogenous allergens for the
497 safety evaluation of genetically engineered food crops: review of potential risks, test methods,
498 examples and relevance. *J Agric Food Chem* **2013**, *61*, 8317-8332.
- 499 (17) Goodman, R. E.; Vieths, S.; Sampson, H. A.; Hill, D.; Ebisawa, M.; Taylor, S. L.; van Ree,
500 R., Allergenicity assessment of genetically modified crops--what makes sense? *Nat Biotechnol*
501 **2008**, *26*, 73-81.
- 502 (18) Stevenson, S. E.; Woods, C. A.; Hong, B.; Kong, X.; Thelen, J. J.; Ladics, G. S.,
503 Environmental effects on allergen levels in commercially grown non-genetically modified
504 soybeans: assessing variation across north america. *Front Plant Sci* **2012**, *3*, 196,1-13.

- 505 (19) Borges, J. P.; Barre, A.; Culerrier, R.; Granier, C.; Didier, A.; Rougé, P., Lipid transfer
506 proteins from Rosaceae fruits share consensus epitopes responsible for their IgE-binding cross-
507 reactivity. *Biochem Biophys Res Commun* **2008**, *365*, 685-690.
- 508 (20) Su, M.; Venkatachalam, M.; Teuber, S. S.; Roux, K. H.; Sathe, S. K., Impact of γ -irradiation
509 and thermal processing on the antigenicity of almond, cashew nut and walnut proteins. *J Sci*
510 *Food Agric* **2004**, *84*, 1119-1125.
- 511 (21) Roux, K. H.; Teuber, S. S.; Robotham, J. M.; Sathe, S. K., Detection and stability of the
512 major almond allergen in foods. *J Agric Food Chem* **2001**, *49*, 2131-2136.
- 513 (22) Pastorello, E. A.; Pompei, C.; Pravettoni, V.; Farioli, L.; Calamari, A. M.; Scibilia, J.;
514 Robino, A. M.; Conti, A.; Iametti, S.; Fortunato, D.; Bonomi, S.; Ortolani, C., Lipid-transfer
515 protein is the major maize allergen maintaining IgE-binding activity after cooking at 100 degrees
516 C, as demonstrated in anaphylactic patients and patients with positive double-blind, placebo-
517 controlled food challenge results. *J Allergy Clin Immunol* **2003**, *112*, 775-783.
- 518 (23) Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J. F.;
519 Guindon, S.; Lefort, V.; Lescot, M.; Claverie, J. M.; Gascuel, O., Phylogeny.fr: robust
520 phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **2008**, *36*, W465-469.
- 521 (24) Hajduch, M.; Ganapathy, A.; Stein, J. W.; Thelen, J. J., A systematic proteomic study of
522 seed filling in soybean. Establishment of high-resolution two-dimensional reference maps,
523 expression profiles, and an interactive proteome database. *Plant Physiol* **2005**, *137*, 1397-1419.
- 524 (25) Radauer, C.; Breiteneder, H., Pollen allergens are restricted to few protein families and
525 show distinct patterns of species distribution. *J Allergy Clin Immunol* **2006**, *117*, 141-147.
- 526 (26) Houston, N. L.; Lee, D. G.; Stevenson, S. E.; Ladics, G. S.; Bannon, G. A.; McClain, S.;
527 Privalle, L.; Stagg, N.; Hérouet-Guicheney, C.; MacIntosh, S. C.; Thelen, J. J., Quantitation of
528 soybean allergens using tandem mass spectrometry. *J Proteome Res* **2011**, *10*, 763-773.

- 529 (27) Pedersen, M. H.; Holzhauser, T.; Bisson, C.; Conti, A.; Jensen, L. B.; Skov, P. S.;
530 Bindslev-Jensen, C.; Brinch, D. S.; Poulsen, L. K., Soybean allergen detection methods--a
531 comparison study. *Mol Nutr Food Res* **2008**, *52*, 1486-1496.
- 532 (28) Kuppanan, K.; Albers, D. R.; Schafer, B. W.; Dielman, D.; Young, S. A., Quantification
533 and characterization of maize lipid transfer protein, a food allergen, by liquid chromatography
534 with ultraviolet and mass spectrometric detection. *Anal Chem* **2011**, *83*, 516-524.
- 535

Accession	Description	Organism	MM (kDa)
GRMZM2G071575_T01	LTP	maize	18.89
GRMZM5G898755_T01	nonspecificLTP4pre	maize	12.29
GRMZM5G898755_T02	nonspecificLTP4pre	maize	13.32
GRMZM2G101958_T01	phosphoLTP1	maize	11.74
GRMZM2G010868_T01	LTP	maize	11.72
GRMZM2G010868_T02	LTP	maize	11.51
P19656	NLTP_MAIZE phosphoLTP1	maize	11.70
GRMZM2G137329_T01	nonspecific_LTPpre	maize	9.86
GRMZM2G081464_T01	nonspecific_LTP2	maize	10.30
P83506	NLTP2	maize	7.31
Q2QYL2	NLT2B_ORYSJ nLTP2b1pre	rice	11.44
Q7XJ39	NLT2A_ORYSJ nLTP2a	rice	11.49
Q43194	NLTP2_SORBI nLTP2	sorghum	11.95

Table 1 LTPs identified during the discovery proteomics experiment. All proteins that shared homology with Zea m 14 based on local BLAST analysis against the list of all identified proteins, and were less than 20 kDa, are listed here. Descriptions were provided by maizesequence.org or UniProt.org websites. Molecular mass (MM) for each entry is noted.

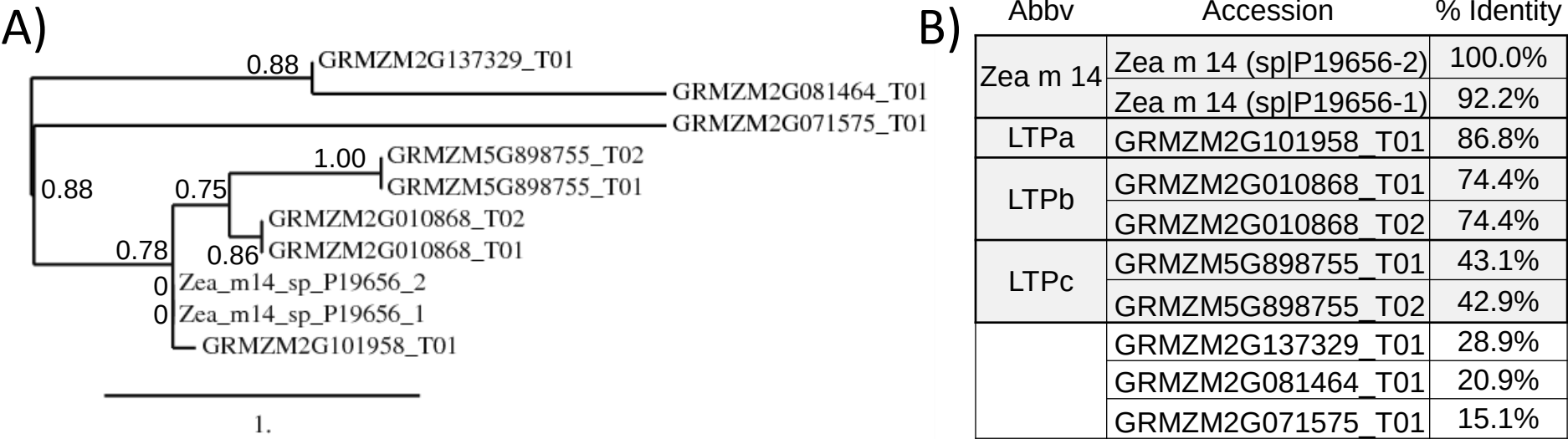


Figure 1 Sequence homology of LTPs found by LC-MS/MS showing relatedness to Zea m 14 (sp|P19656). A) Phylogram of the top maize proteins (translated gene sequences) identified during the discovery experiment. Zea_m14_sp_P19656_1 and 2 represent the two isoforms of UniProt entry P19656, both of which are recognized as LTP allergen Zea m 14. Other entries containing T01 and T02 indicate putative, annotated splice variants. Numbers at branch points are branch support values. B) Pairwise Clustal Omega percent identity scores from alignments performed between Zea m 14 isovariant 2 and the other listed accessions. LTPs in grey have greater than 42% sequence identity with Zea m 14-2 and were given an abbreviation (abbrev column) to simplify nomenclature. These same LTPs were also clustered together phylogenetically (shown in A).

Figure 2 LTP protein alignment showing peptides for AQUA analysis. Ideal, quantitative tryptic peptides are highlighted. Specificity is indicated by color. Putative splice variants are paired by brackets. Each LTP has at least one peptide that distinguishes it, and two of them share one peptide (green). Trypsin cleavage sites are heavy underscored. Amino acid positions are indicated at the ends of each line.

	Tryptic peptides				
LTP protein accessions	GQSGPSAGCCSGVK	GQSGPSAGCCSGVR	NAAAGVSGLNAGNAASIPSK	GVSGLNAGNAASIPSK	LGGGVSMANAAANIPSK
GRMZM2G101958 (LTPa)	*		*		
GRMZM2G010868 (LTPb)				*	
GRMZM5G898755 (LTPc)					*
Zea m 14		*	*		

Table 2 LTP protein/peptide matrix. Tryptic peptides are in columns, and the LTP accessions in rows. Asterisks are placed at the intersections of peptides and LTP accessions to which they belong. The peptide highlighted in grey is shared by two LTPs (the same as the green peptide in Figure 2). All other peptides are unique to the respective LTP isoform.

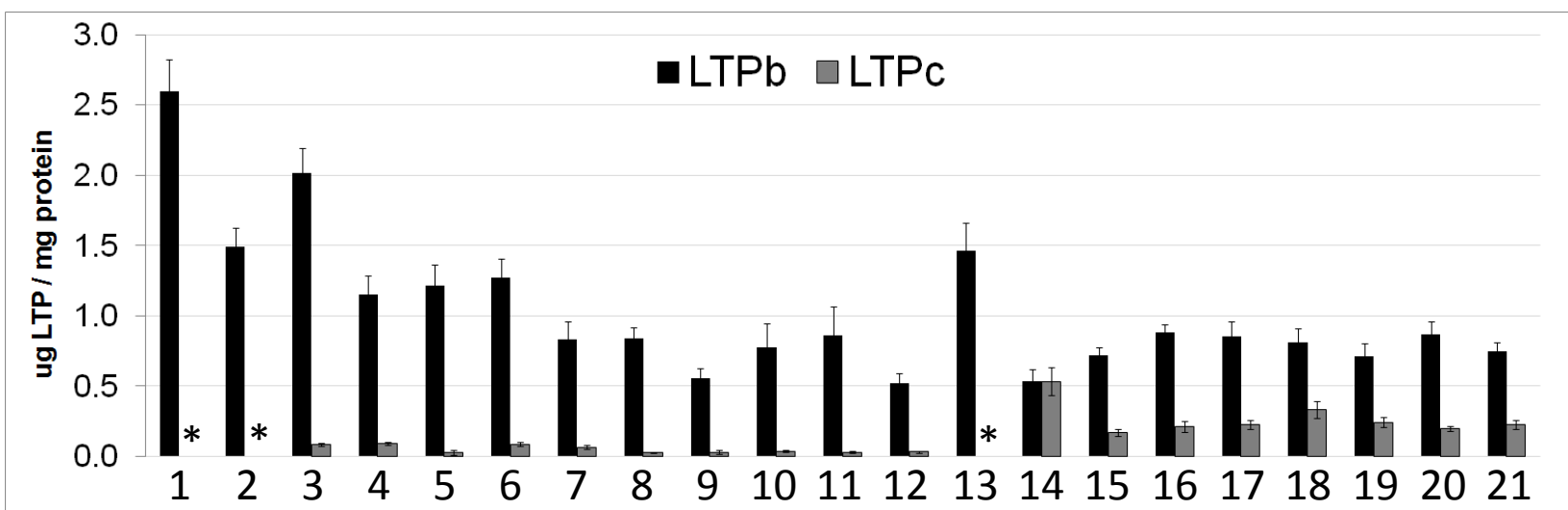
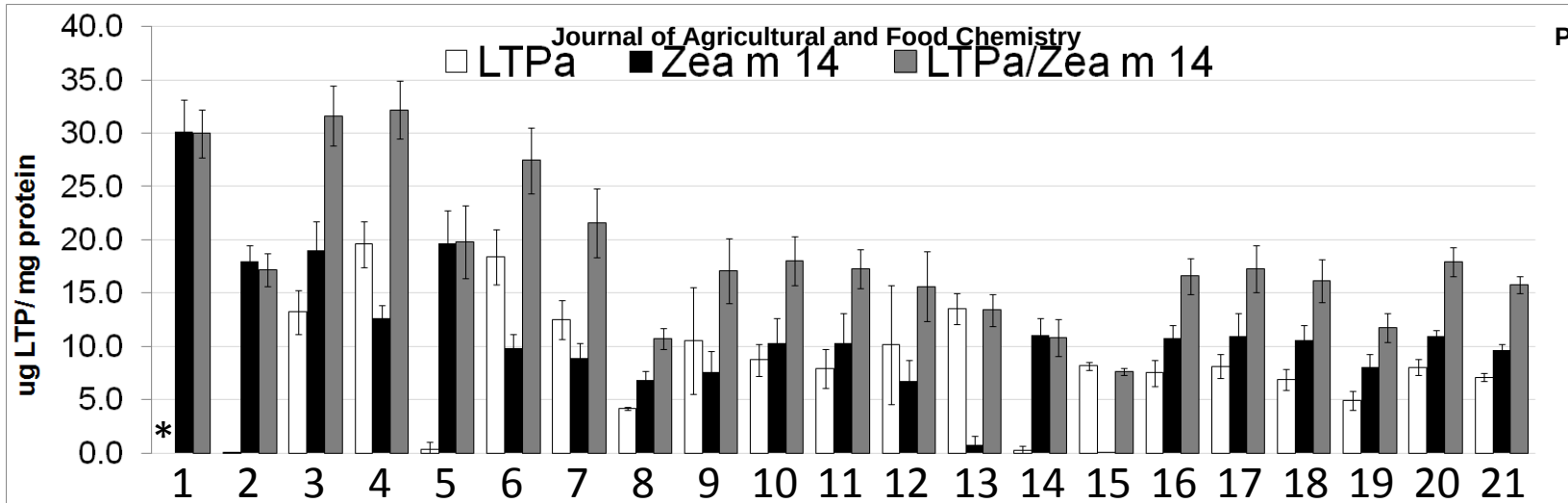


Figure 3 Absolute quantities of LTPs in each variety. LTPa/Zea m 14 represents the quantity provided by the peptide assay which monitors both LTPa and Zea m 14 simultaneously using a peptide that is shared by both proteins. Error bars represent standard deviation for five biological replicates. Numbers below the x-axis are variety identifiers. Asterisks indicate samples below detection limits.

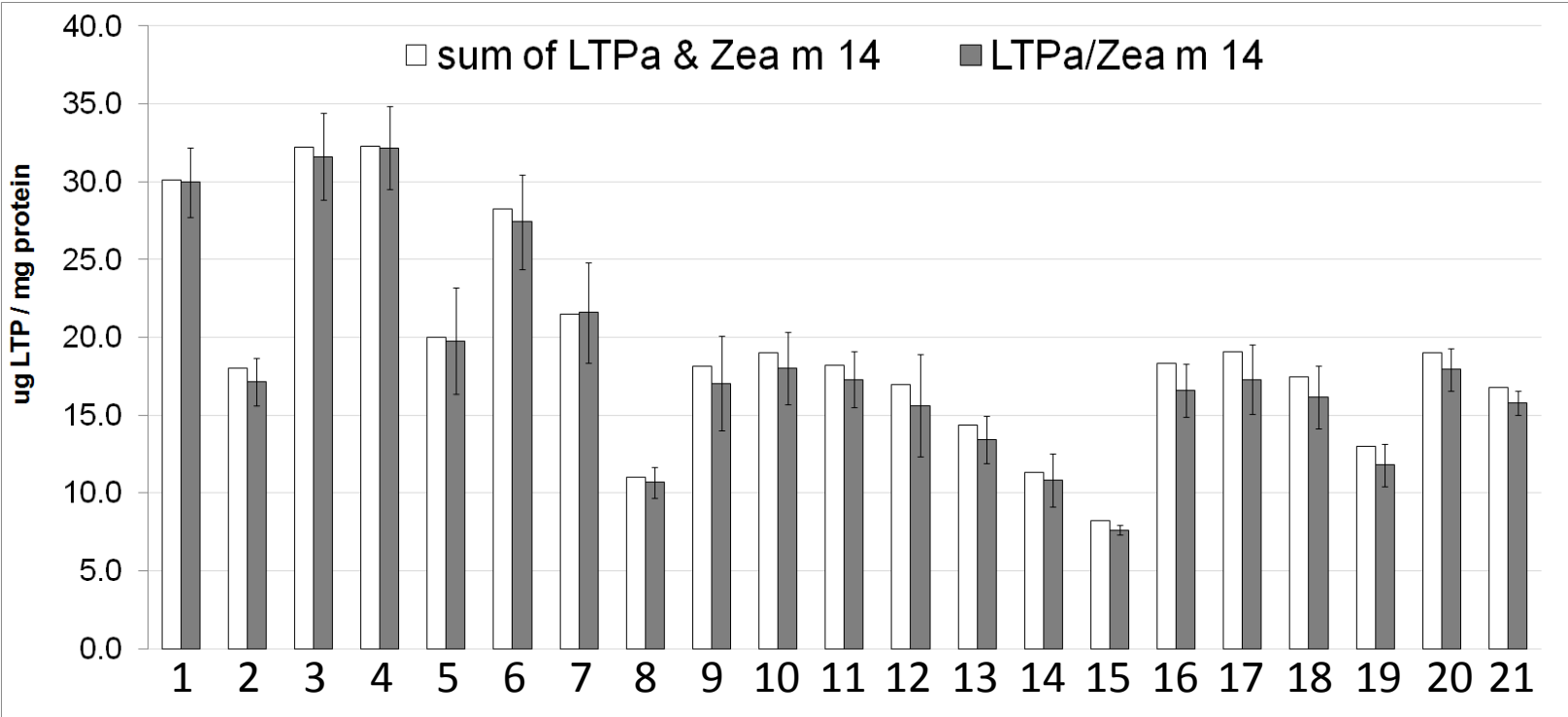


Figure 4 Stoichiometric agreement between three separate LTP AQUA-MRM assays. Error bars represent standard deviation for five biological replicates. Numbers below the x-axis are variety identifiers.

Prot Abbv	IUIS allergen nomenclature	MM (g/mol)	AQUA peptide sequence	AQUA assay LLOQ (pg/injection)	AQUA assay ULOQ (ng/injection)
LTPa	none	11742.84	GQGSGPSAGCCSGVK	9.39	5.87
LTPb	none	11718.82	GVSGLNAGNAASIPSK	1.88	5.86
LTPc	none	13319.46	LGGGVSMANAAANIPSK	10.66	6.66
Zea m 14	Zea m 14	12659.23	GQGSGPSAGCCSGVR	2.03	6.33
LTPa / Zea m 14	Zea m 14	12659.23	NAAAGVSGLNAGNAASIPSK	2.03	6.33

Table 3 Summary table for proteins and peptides. LLOQ is lower limit of quantitation and represents the lowest point at which we demonstrated quantitative accuracy during our assay development. ULOQ is the highest amount tested with AQUA peptides, though assays with native peptides showed linearity of response beyond this value

TOC Graphic

