

Natural Variability of Allergen Levels in Conventional Soybeans: Assessing Variation across North and South America from Five Production Years

Tao Geng,^{*} Duška Stojšin, Kang Liu, Bruce Schaalje, Cody Postin, Jason Ward,[§] Yongcheng Wang, Zi Lucy Liu, Bin Li, and Kevin Glenn[¶]

Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167, United States

ABSTRACT: Soybean (*Glycine max* L. Merrill) is one of eight major allergenic foods with endogenous proteins identified as allergens. To better understand the natural variability of five soybean allergens (Gly m 4, Gly m 5, Gly m 6, Gly m Bd 28k, and Gly m Bd 30k), validated enzyme-linked immunosorbent assays (ELISAs) were developed. These ELISAs measured allergens in 604 soybean samples collected from locations in North and South America over five growing seasons (2009–2013/2014) and including 37 conventional varieties. Levels of these five allergens varied 5–19-fold. Multivariate statistical analyses and pairwise comparisons show that environmental factors have a larger effect on allergen levels than genetic factors. Therefore, from year to year, consumers are exposed to highly variable levels of allergens in soy-based foods, bringing into question whether quantitative comparison of endogenous allergen levels of new genetically modified soybean adds meaningful information to their overall safety risk assessment.

KEYWORDS: soybean, allergenicity, safety assessment, ELISA, natural variability

INTRODUCTION

Grain from soybean (*Glycine max* L. Merrill) is a good source of nutritional protein for animal and human consumption with protein levels ranging from ~34 to >56% on a dry weight basis.¹ In addition to being a major food source of protein, consumption of soybean has also been associated with numerous health benefits.^{2–6} However, some soybean proteins are associated with food allergies, resulting in soybean being ranked as one of the “big eight” food allergens in the United States,⁷ which together are responsible for >90% of all food allergies.^{8,9}

Up to 15 soybean proteins have been listed as potential allergens,¹⁰ with a recent review identifying 8 of them as being most suitable for measurement and indicating that measuring the levels of endogenous allergens is irrelevant in the assessment of food safety for a variety of reasons.¹¹ However, if measurements are required by government regulatory organizations, five of those eight (Gly m 4, β -conglycinin (Gly m 5), glycinin (Gly m 6), Gly m Bd 28K, and Gly m Bd 30K) have the strongest combination of supporting clinical data and sufficient sequence information to support the development of quantitative methods. Gly m 5 and Gly m 6 are the two major seed storage proteins.¹¹ Gly m 5, or β -conglycinin, is a vicilin-type globulin of the 7S seed storage protein in soybean, whereas Gly m 6, or glycinin, is a legumin-type globulin of the 11S seed storage proteins in soybean. Both belong to a cupin superfamily.¹² By examining patients with double-blind, placebo-controlled food challenge (DBPCFC) and/or positive case histories, Holzhauser et al. concluded that sensitization to Gly m 5 or Gly m 6 is potentially indicative for severe allergic reactions to soybean, because IgE-binding to Gly m 5 or Gly m 6 was found in 86% (6/7) subjects with anaphylaxis to soybean.¹³ To date, several other 7S and 11S globulins from

peanut, sesame, and some tree nuts have been identified as food allergens.¹² By comparison, Gly m Bd 28K is a minor (<0.5% (w/w)) soybean seed protein of 7S globulin.¹⁴ Gly m Bd 28K was isolated and characterized from defatted soybean flakes as a 28 kDa glycosylated protein.¹⁵ It was later demonstrated that IgE antibodies from 10 soybean-allergic sera bind to the purified protein.¹⁶ This protein is highly homologous to a MP27/MP32 allergen from pumpkin seeds and to a globulin-like allergen from carrot.¹⁷ Gly m Bd 28K is one of the three most common soybean seed allergens.¹⁸ Gly m Bd 30K, also known as P34, is also a relatively minor soybean storage vacuole protein (approximately 1% of soy protein).¹⁴ It belongs to the papain superfamily of thiol proteases.¹² Gly m Bd 30K is a major soybean allergen responsible for dermatitis and food allergy.¹⁹ Some of its homologues are responsible for the allergenicity of kiwi, papaya, and pineapple.¹² Gly m 4 is a member of the pathogenesis-related proteins 10 (PR-10) family that is specifically expressed in response to plant pathogen infection, plant hormones, wounding, or environmental factors (salt or cold stresses).²⁰ It belongs to the Bet v 1 homologous superfamily of pollen and food allergens²¹ and thus shows high amino acid sequence identity (>50%) and almost identical three-dimensional structure to several pollen allergens.²¹ Gly m 4 was recognized by IgE in 96% (21/22) of patients with DBPCFC and/or positive case histories.²² In addition, it was found to be responsible for inducing an oral allergy syndrome and severe systemic reactions.²³

Received: October 12, 2016

Revised: December 15, 2016

Accepted: December 20, 2016

Published: December 20, 2016

Considering that soybean is a commonly allergenic food, safety assessments of genetically modified (GM) soybean include measurement of the levels of endogenous allergens compared to the levels in conventional comparators.²⁴ However, as reviewed previously, from a clinical perspective it is unclear that data on potential changes in allergen levels help inform public safety, because individuals with food allergies, such as to foods containing ingredients from soybeans, are typically instructed to carefully read product labels to help them completely avoid eating such foods.^{25,29–31} In addition, two soybean allergy studies showed a wider variation in IgE binding to different conventional soybean varieties than that of GM soybean,^{26,27} suggesting that individuals are exposed to highly variable levels of soybean allergenic proteins, which could be either due to the natural variability of soybean allergen proteins or due to the level of soybean grain in consumed in the diet.

The present study significantly contributes to a better understanding of the natural variability of endogenous allergenic proteins in soybeans and the factors that might influence the variability of five allergen proteins (Gly m 4, β -conglycinin (Gly m 5), glycinin (Gly m 6), Gly m Bd 28K, and Gly m Bd 30K).

MATERIALS AND METHODS

Soybean Field Trials. Thirty-seven conventional soybean varieties were grown at 26 geographically diverse locations in either North or South America during the 2009, 2010, 2012, 2013, and 2013/2014 growing seasons (Table 1). A total of 12 states/provinces were represented by these locations: 10 states in the United States (Arkansas, Illinois, Indiana, Iowa, Kansas, Missouri, Nebraska, North Carolina, Ohio, and Pennsylvania) and two provinces in Argentina (Buenos Aires and Santa Fe).

The 37 conventional soybean varieties were selected to provide some level of genetic diversity and to represent a range of maturity groups (MG 2.7–MG 4.2) suitable for growing in the areas where field trials were conducted.

The trials at each location were planted in a randomized complete block design with four replications. Agronomic practices typical for each region were followed for these trials. Seed was harvested at maturity, the R8 growth stage (when at least 95% of the pods had reached mature pod color), when the grain moisture content was approximately 12–15%. The harvested grain was transported at ambient temperature after harvest, but was stored in cold storage at $-80\text{ }^{\circ}\text{C}$ prior to the enzyme-linked immunosorbent assay (ELISA) analysis.

Weather Data Information. Temperature and precipitation values were obtained for each location from publicly available weather data sources collected by weather stations located in close proximity to the field locations. Irrigation was recorded at each location where it was applied. Midseason minimum and maximum daily temperatures were collected for the month of July for the U.S. locations and for February for Argentina locations and averaged across all locations for a growing season (Table 1). Midseason water availability was calculated as the sum of precipitation and irrigation averaged across sites for each growing season and was collected for the month of July for the U.S. locations and for February for Argentina locations. The months of July and February were chosen to represent midseason weather information because these periods correspond to soybean flower and pod development for these field trials, which are crucial phases in determining grain quantity and quality at harvest.

Protein Extraction from Soybean Seeds. Soybean seed samples were ground using a mega grinder (Monsanto Co., St. Louis, MO, USA) for approximately 1 min. Soybean allergens Gly m 4, Gly m 5, and Gly m 6 were extracted using a KLECO Pulverizer (Kinetic Laboratory Equipment Co., Visalia, CA, USA) with 1.5 mL (1:100 tissue to buffer ratio) of 0.01 M phosphate-buffered saline with 0.05% (v/v) Tween-20, pH 7.4 (PBST), and two 4.8 mm steel balls. Each

Table 1. Growing Seasons, Locations, and Weather Parameters Associated with Field Trials That Provided Soybean Samples

	growing season				
	2009	2010	2012	2013	2013/2014
growing locations ^a	6 U.S.: Indiana, Iowa, Kansas, Missouri, Nebraska, Illinois	8 U.S.: Arkansas, Iowa (3), Illinois, Indiana, Missouri, Pennsylvania	8 U.S.: Arkansas, Iowa (2), Illinois, North Carolina, Nebraska, Pennsylvania	8 U.S.: Arkansas, Iowa, Illinois, North Carolina, Nebraska, Ohio, Pennsylvania	8 Argentina: Buenos Aires (7), Santa Fe
no. of conventional varieties ^b	15	21	19	18	18
midseason water availability ^c (mm)	124	167	136	128	283
min–max midseason temperature ^d ($^{\circ}\text{C}$)	16.7–28.3	19.5–30.3	21.1–33.1	18.5–29.6	17.8–26.6

^aA total of 38 locations with 26 unique locations across all growing seasons. ^bA total of 91 entries with 37 unique conventional soybean varieties across all growing seasons. ^cAverage July water availability for U.S. locations and average February water availability for Argentina locations. Water availability is the sum of precipitation and irrigation averaged across sites. ^dAverage minimum and maximum July temperatures for U.S. locations and average minimum and maximum February temperatures for Argentina locations.

Table 2. Summary of ELISA Format, Protein Standard, and Antibodies for Each Soybean Allergen ELISA

	allergen				
	Gly m 4	Gly m 5	Gly m 6	Gly m Bd 28k	Gly m Bd 30k
ELISA format	indirect sandwich ELISA	indirect sandwich ELISA	indirect sandwich ELISA	indirect sandwich ELISA	indirect sandwich ELISA
protein standard	rGly m 4 ^a	native Gly m 5 ^b	native Gly m 6 ^b	rGly m Bd 28k ^a	rGly m Bd 30k ^a
capture antibody	mouse mAb ^c	goat peptide pAb ^d	goat peptide pAb ^d	goat pAb ^e	rabbit pAb ^e
detection antibody	goat pAb ^e	mouse mAb ^c	mouse mAb ^c	goat pAb ^e	goat pAb ^e

^arGly, recombinant soybean allergens produced from *E. coli*. ^bNative, allergens were purified from soybean seed. ^cmAb, monoclonal antibody. ^dpeptide pAb, common peptides from subunits of Gly m 5 or Gly m 6 were used to produce antibodies. ^epAb, polyclonal antibody.

extract was centrifuged, and the supernatant was collected and stored in a -70 ± 10 °C freezer until analysis.

Gly m Bd 28k and Gly m Bd 30k were extracted from 100 mg of ground sample by using a Harbil Mixer (Fluid Management, Inc., Wheeling, IL, USA) with 10 mL (1:100 ratio of tissue to buffer) of 4 M guanidine with 10 mM DTT, pH 7.5, and eight chrome beads (0.25 in. in diameter) for 10 min at 1500 rpm. Insoluble material was removed from soybean seed extract using a 16 mm × 4 in. serum filter (catalog no. 02-681-51, Fisher Scientific, Pittsburgh, PA, USA).

ELISAs for Five Soybean Allergens. All five soybean allergen ELISAs were developed as previously described.²⁸ Briefly, each allergen protein was either expressed in *Escherichia coli* or purified from soybean seed. Then, the protein concentration and purity were determined by using amino acid analysis and SDS-PAGE gels. Finally, each protein was characterized by N-terminal sequence analysis and mass spectrometry.

Monoclonal and polyclonal antibodies were produced by using fully characterized proteins or specific peptides as immunogens. The antibodies were purified by a protein G column (GE Healthcare, Piscataway, NJ, USA), followed by an allergen or peptide-specific affinity column if needed. Western blots were used to determine the specificity of affinity-purified antibodies to the allergen protein standards and allergens in soybean extract. For antibody pairs of sandwich ELISAs, at least one antibody showed negligible cross-reaction with other soybean proteins in the soybean seed extracts. The ELISA methods are summarized in Table 2.

Each allergen ELISA was optimized for its specificity, linearity, robustness, ruggedness, and stability. ELISAs were subsequently validated for their precision, limits of detection and quantitation (LOD and LOQ), dilutional parallelism, and extraction efficiency as previously described.²⁸ Briefly, the precisions of standard curve points and QC+ samples were determined both within a run (intra-assay precision of technical replicates) and across runs (interassay precision of assay plates over a period of time) in triplicate wells over 36 assay runs carried out on three different days by three different analysts. Because no negative matrix is available due to the endogenous nature of these proteins, the LOD was determined by 21 replicate buffer samples in triplicate wells on three ELISA runs by three different analysts. The LOQ was defined as the lowest concentration on the standard curve with a percentage coefficient of variation (CV) of <15%. Dilutional parallelism, a method used to demonstrate the accuracy of an ELISA assay, was demonstrated by testing four dilutions of a soybean protein extract from five samples respectively, within the quantitative range of the standard curves. Each result was then adjusted accordingly for the appropriate dilution factor and compared to the average concentration of all dilutions for that sample. Extraction efficiency, another method to test the accuracy, was also performed by successive extractions on the same material and analyzing each extract to determine the amount of each allergen that remains after the first extraction. The extraction efficiency was expressed as the percent of an allergen recovered in the first extraction compared to the total recovered from the consecutive extractions of the same sample.

Measurement of Five Allergen Levels from Soybean Seed. The measurements were conducted in a GLP (Good Laboratory Practice) laboratory at Covance Laboratories Inc. (Madison, WI, USA) with GLP compliance using validated ELISA methods. Samples were analyzed in a randomized order to reduce assay bias.

Variance Component Analysis (VCA) for Allergen Levels.

VCA was conducted for all ELISA results from a total of 604 samples to estimate the proportion of random effects contributing to the total variance, based on the following analysis of variance (ANOVA) model by combining all field locations:

$$Y = \mu + E + S + V + E \times V + S \times V + L(SE) + E \times S + M + P + D + \epsilon$$

Y is the unique individual observation; μ is the overall mean; E is the growing season effect; S is the state/province effect; V is the variety effect; $E \times V$ is the interaction between season and variety effect; $S \times V$ is the interaction between state/province and variety effect; $L(SE)$ is the location effect nested within state/province and season effect; $E \times S$ is the interaction between season and state/province effect; M is the maturity range effect; P is the planting time effect; D is the growing season duration effect; and ϵ is the residual error.

In this application, all of the effects in the ANOVA model were set as random. The SAS procedure PROC MIXED was employed to get covariance parameter estimates for each effect in the model (copyright 2002–2012 by SAS Institute, Cary, NC, USA). Finally, the variance component parameters of growing season, state/province, location, planting time, growing season duration, and the interaction between season and state effects were combined as environmental effect; variety and maturity range were combined as the genetic effect. They were both divided by the total variance to get the variance proportions for each.

Principal Components Analysis (PCA). A PCA was conducted to explore the multivariate clustering patterns of the data. The allergen ELISA data were first normalized to have a common mean of 0 and a standard deviation of 1. Graphs of PCA scores and loadings were utilized to visualize the high-dimensional structure of the data. The scores plot projects the five-dimensional observations (the five soy allergens) onto a two-dimensional subspace that captures as much of the variability and clustering structure as possible. The two axes of the scores plot (or “principal components”) are uncorrelated linear combinations of the five allergen values. When the points are labeled by states/provinces and year, the scores plot helps identify and interpret point clusters. The loadings plot provides information about the nature of the principal components. Each vector represents one of the five allergens, and the coordinates of the end-point give the correlations of that allergen with the principal components. For example, vectors that are close to horizontal represent allergens that are highly correlated with the first principal component (PC1). Vectors that are close to vertical represent allergens that determine the second principal component (PC2).

Pairwise Comparisons. To better understand which genetic and environmental components affected the level of the five soybean allergens, pairwise comparisons were conducted for several variables. Two genetic variables were included in this analysis: soybean varieties and their maturity ranges. The pairwise analysis included all 37 soybean varieties, whereas maturity ranges were expressed in half maturity group increments (MG 2.5–2.9, MG 3.0–3.4, MG 3.5–3.9, and MG 4.0–4.4). Four environmental variables were considered for the pairwise comparisons: growing seasons, states/provinces, planting times, and growing season duration. All 5 growing seasons and 12 states/provinces were included in pairwise comparisons. The planting times were determined in relative terms as early, intermediate, and late

planting for each season considering 2 week intervals for the first planting date. Duration of growing season was measured in number of days from planting to harvest. Four duration intervals were considered in this study: 110–119, 120–129, 130–139, and over 140 days.

The SAS PROC MIXED procedure was used to determine the presence of significant differences among the levels for each of the target variables at the 5% level of significance.

The following mixed model was used for variable growing season comparisons for each allergen:

$$Y_{ijk} = \mu + E_i + S_j + V_k + E \times S_{ij} + E \times V_{ik} + \varepsilon_{ijk}$$

Y_{ijk} is the value of allergen level of growing season i , state/province j , and variety k ; μ is the overall mean; E_i is the fixed effect of the i th growing season; S_j is the random effect of the j th state/province; V_k is the random effect of k th variety; $E \times S_{ij}$ is the random effect of the interaction between the i th growing season and j th state/province; $E \times V_{ik}$ is the random effect of the interaction between the i th growing season and k th variety; and ε_{ijk} is the residual error.

The following mixed model was used for variable variety comparisons for each allergen:

$$Y_{ijk} = \mu + V_i + E_j + S_k + V \times E_{ij} + V \times S_{ik} + E \times S_{jk} + \varepsilon_{ijk}$$

Y_{ijk} is the value of allergen level of variety i , growing season j , and state/province k ; μ is the overall mean; V_i is the fixed effect of the i th variety; E_j is the random effect of the j th growing season; S_k is the random effect of k th state/province; $V \times E_{ij}$ is the random effect of the interaction between the i th variety and j th growing season; $V \times S_{ik}$ is the random effect of the interaction between the i th variety and k th state/province; $E \times S_{jk}$ is the random effect of the interaction between the j th growing season and k th state/province; and ε_{ijk} is the residual error.

The following mixed model was used for variable planting time comparisons for each allergen:

$$Y_{ijkl} = \mu + P_i + E_j + S_k + V_l + E \times S_{jk} + E \times V_{jl} + P \times E_{ij} + P \times V_{il} + P \times S_{ik} + S \times V_{kl} + \varepsilon_{ijkl}$$

Y_{ijkl} is the value of allergen level of planting time i , growing season j , state/province k , and variety l ; μ is the overall mean; P_i is the fixed effect of the i th planting time; E_j is the random effect of the j th growing season; S_k is the random effect of the k th state/province; V_l is the random effect of the l th variety; $E \times S_{jk}$ is the random effect of the interaction between the j th growing season and k th state/province; $E \times V_{jl}$ is the random effect of the interaction between the j th growing season and l th variety; $P \times E_{ij}$ is the random effect of the interaction between the i th planting time and j th growing season; $P \times V_{il}$ is the random effect of the interaction between the i th planting time and l th variety; $P \times S_{ik}$ is the random effect of the interaction between the i th planting time and k th state/province; $S \times V_{kl}$ is the random effect of the interaction between the k th state/province and l th variety; and ε_{ijkl} is the residual error.

The following mixed model was used for variable state/province comparisons for each allergen:

$$Y_{ijk} = \mu + S_i + E_j + V_k + S^*E_{ij} + E \times V_{jk} + S \times V_{ik} + \varepsilon_{ijk}$$

Y_{ijk} is the value of allergen level of state/province i , growing season j , and variety k ; μ is the overall mean; S_i is the fixed effect of the i th state/province; E_j is the random effect of the j th growing season; V_k is the random effect of k th variety; $S \times E_{ij}$ is the random effect of the interaction between the i th state/province and j th growing season; $E \times V_{jk}$ is the random effect of the interaction between the j th growing season and k th variety; $S \times V_{ik}$ is the random effect of the interaction between the i th state/province and k th variety; and ε_{ijk} is the residual error.

The following mixed model was used for variable growing season duration comparisons for each allergen:

$$Y_{ijkl} = \mu + D_i + E_j + S_k + V_l + E \times S_{jk} + E \times V_{jl} + D \times E_{ij} + D \times V_{il} + D \times S_{ik} + S \times V_{kl} + \varepsilon_{ijkl}$$

Y_{ijkl} is the value of allergen level of growing season duration i , growing season j , state/province k , and variety l ; μ is the overall mean; D_i is the fixed effect of the i th growing season duration; E_j is the random effect of the j th growing season; S_k is the random effect of the k th state/province; V_l is the random effect of the l th variety; $E \times S_{jk}$ is the random effect of the interaction between the j th growing season and k th state/province; $E \times V_{jl}$ is the random effect of the interaction between the j th growing season and l th variety; $D \times E_{ij}$ is the random effect of the interaction between the i th growing season duration and j th growing season; $D \times V_{il}$ is the random effect of the interaction between the i th growing season duration and l th variety; $D \times S_{ik}$ is the random effect of the interaction between the i th growing season duration and k th state/province; $S \times V_{kl}$ is the random effect of the interaction between k th state/province and l th variety; and ε_{ijkl} is the residual error.

The following mixed model was used for variable maturity range comparisons for each allergen:

$$Y_{ijkl} = \mu + M_i + E_j + S_k + V_l + E \times S_{jk} + E \times V_{jl} + M \times E_{ij} + M \times V_{il} + M \times S_{ik} + S \times V_{kl} + \varepsilon_{ijkl}$$

Y_{ijkl} is the value of allergen level of maturity range i , growing season j , state/province k , and variety l ; μ is the overall mean; M_i is the fixed effect of the i th maturity range; E_j is the random effect of the j th growing season; S_k is the random effect of the k th state/province; V_l is the random effect of the l th variety; $E \times S_{jk}$ is the random effect of the interaction between the j th growing season and k th state/province; $E \times V_{jl}$ is the random effect of the interaction between the j th growing season and l th variety; $M \times E_{ij}$ is the random effect of the interaction between the i th maturity range and j th growing season; $M \times V_{il}$ is the random effect of the interaction between the i th maturity range and l th variety; $M \times S_{ik}$ is the random effect of the interaction between the i th maturity range and k th state/province; $S \times V_{kl}$ is the random effect of the interaction between the k th state/province and l th variety; and ε_{ijkl} is the residual error.

RESULTS

Method Development and Validation of Soybean Allergen ELISAs. For abundant soybean allergens Gly m 5 and Gly m 6, the protein standards for ELISAs were purified from soybean seed, whereas the proteins were expressed in *E. coli* for less abundant soybean allergens (Gly m 4, Gly m Bd 28k, and Gly m Bd 30k). Similar to what was previously described for purification of Gly m 4,²⁸ once each of these five soybean allergens was isolated, its purity level was determined through the calculation of the allergen band(s) intensity over total protein intensity within the same lane on SDS-PAGE gels. All five allergen protein standards had >70% purity. The purity corrected concentration of each protein standard was used to prepare ELISA standard curves. Also similar to previously described methods for Gly m 4,²⁸ the identity of each allergen included N-terminal sequence analysis and mass spectrometry analysis of trypsin-digested fragments. For all five proteins, the N-terminal amino acid sequence exactly matched the predicted amino acid sequence for the respective protein. Also, peptide mass spectrometry fingerprint analysis showed that the unique peptides identified from trypsin digestion of each of the five allergens corresponded to the predicted masses for each fragment and that assembly of a peptide map of each protein confirmed the identity to the predicted result for that protein (data not shown).

Table 3. Summary of Values Associated with Five Endogenous Soybean Allergens (mg/g Fresh Weight) across All Samples

	allergen				
	Gly m 4	Gly m 5	Gly m 6	Gly m Bd 28k	Gly m Bd 30k
mean	0.11	35.21	169	0.20	1.39
minimum	0.02	5.52	78.4	0.10	0.30
maximum	0.33	90.4	393	0.47	3.49
range ^a	19	16	5	5	12

^aRange, fold from minimum to maximum.

The specificity of each allergen ELISA was assessed by using other soybean allergens or proteins. Except targeted allergens, ELISA methods showed less than LOQ for other soybean proteins (data not shown). Taken together, by combination of the specificity of both antibodies and ELISA, each allergen ELISA method was verified for its specificity to measure the target allergen from soybean seed.

Soybean allergen ELISAs were also validated for their precision, limits of detection and quantitation (LOD and LOQ), dilutional parallelism, and extraction efficiency. Assay precision was calculated as the percentage coefficient of variation (CV) using the interpreted concentrations of the QC+ samples and the standards observed over all assay runs. For all five ELISA methods, the CV percentages of intra-assay and interassay precisions of standards and QC+ were <15% (data not shown). Nevertheless, dilutional parallelism and extraction efficiency of five assays were validated to be within 70–130%.

Overall, ELISA methods were developed and validated to measure five allergens from soybean seed in a specific, precise, and accurate manner.

Variability of Field Trial Environments. Thirty-seven conventional soybean varieties were grown in 26 locations over a total of 5 growing seasons from 2009 to 2013/2014, representing a diverse range of environmental conditions across the North and South American regions where soybean is typically grown. The midseason minimum temperatures ranged on average from 16.7 to 21.1 °C (Table 1). The midseason maximum temperatures ranged on average from 26.6 to 33.1 °C. Overall, the 2013/2014 Argentina growing season had the lowest midseason temperatures, whereas the 2012 U.S. growing season was the warmest. Midseason water availability ranged from 124 to 283 mm depending on the season. Water availability for the U.S. growing seasons was fairly consistent, ranging from 124 to 167 mm; however, water availability in the 2013/2014 Argentina growing season was much higher at 283 mm. When midseason temperature and water availability were considered together, the 2012 U.S. growing season had the most drought stress (high temperatures and lower water availability), whereas the 2013/14 Argentina growing season had the most optimal growing conditions (lower temperatures and approximately twice the water availability compared to other seasons). The 2009, 2010, and 2013 U.S. seasons were between these two extremes.

Quantification of Soybean Allergens from Seed Samples. The five ELISA methods described above were used to measure their respective allergen levels in a total of 604 soybean samples collected from 37 varieties grown over 5 growing seasons in North and South America. The results showed that expression levels varied considerably for all of the measured allergens (Table 3). The most variation was observed for Gly m 4 (19-fold), followed by Gly m 5 (16-fold), Gly m Bd 30k (12-fold), Gly m 6 (5-fold), and Gly m Bd 28k (5-fold).

The magnitude of each soybean allergen level is similar to those in previous studies using ELISA, mass spectrometry, or Western blots.^{22,29–36}

Assessments of Variability. The contributions of genetic and environmental factors to the variability of the level of these five endogenous allergens were assessed using VCA. Results indicated that environmental factors contributed considerably more to the total variability of these soybean allergens compared to genetic factors. This was true when assessed as a composite of all allergens (Figure 1A), as well as when analyzed for each of the five individual allergens tested in this study (Figure 1B).

Genetic Factors as Source of Variation. Genetic factors considered in this study were soybean varieties and their

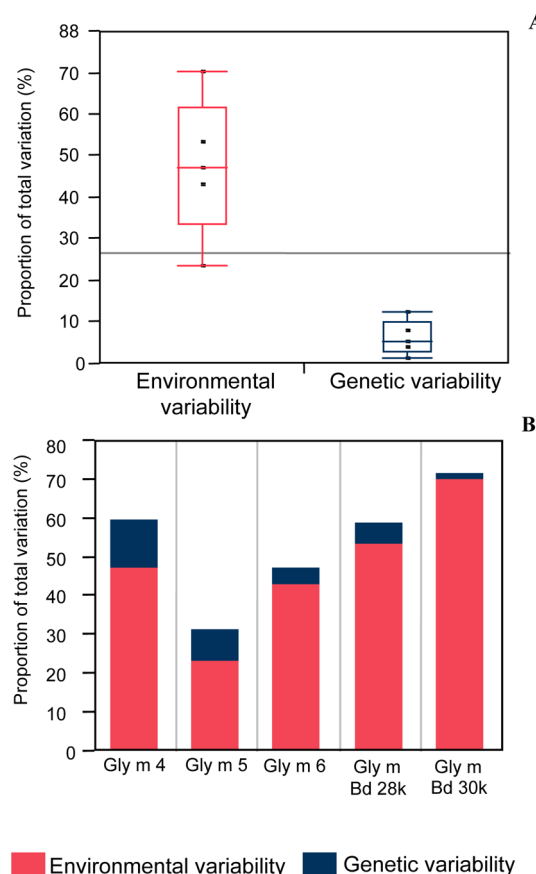


Figure 1. Sources of variation estimates from variance component analysis across growing seasons: (A) overall environmental (red) and genetic (blue) variability for combined soybean allergens; (B) proportion of environmental (red) and genetic (blue) variability for each individual soybean allergen. The remaining proportion of each allergen bar is the residual error from uncontrolled systematic variation.

Table 4. Mean Values of Five Endogenous Allergens (mg/g Fresh Weight) across Evaluated Soybean Varieties^a

soybean variety	maturity group	Gly m 4	Gly m 5	Gly m 6	Gly m Bd 28k	Gly m Bd 30k
Midwest Genetics G2712	2.7	0.05 ab	33.7	178	0.17	2.31
Crows C2804	2.8	0.05 b	43.3	146	0.25	1.64
Dwight	2.9	0.09 b	25.2	141	0.20	1.46
NuPride 2954	2.9	0.13 ab	28.7	197	0.19	1.08
Maverick	3.0	0.14 a	16.9	137	0.20	1.10
NuTech 315	3.1	0.16 ab	26.3	198	0.19	1.24
Schillinger TP31834	3.1	0.10 b	55.1	255	0.20	1.01
Wilken 3316	3.1	0.12 ab	37.8	172	0.19	1.39
NuPride 3202	3.2	0.14 ab	40.3	205	0.19	1.46
NK 32Z3	3.2	0.13 ab	40.1	236	0.20	1.27
NE3202	3.2	0.15 a	35.4	137	0.25	0.82
A3244	3.2	0.10 ab	33.9	141	0.19	1.44
C3211N	3.2	0.09 b	33.3	162	0.18	1.43
Stine 3300-0	3.3	0.13 a	32.4	138	0.23	1.34
Hoegemeyer 333	3.3	0.14 ab	39.5	194	0.17	1.69
Stewart SB3454	3.4	0.10 ab	31.6	145	0.21	1.37
eMerge 348TC	3.4	0.09 ab	33.6	122	0.23	0.85
A3525	3.5	0.10 b	41.6	216	0.18	1.43
Croplan HT3596STS	3.5	0.17 ab	37.3	214	0.16	1.64
FS 3591	3.5	0.13 ab	35.2	224	0.18	1.21
Garst 3585N	3.5	0.09 b	33.0	171	0.20	1.67
LG C3540	3.5	0.10 ab	27.3	160	0.19	1.91
Midland 363	3.6	0.14 a	38.4	152	0.21	1.34
Quality Plus 365C	3.6	0.15 ab	42.3	175	0.18	1.12
Crows C37003N	3.7	0.08 b	31.6	140	0.17	1.52
HOSHEA	3.7	0.16 ab	39.9	187	0.15	1.77
Lewis 372	3.7	0.15 ab	33.6	192	0.22	1.08
Hoffman HS387	3.8	0.13 ab	34.5	162	0.20	1.37
Williams 82	3.8	0.13 a	35.7	149	0.19	1.37
C3884N	3.8	0.14 ab	27.4	155	0.17	1.20
NK S38-T8	3.8	0.15 ab	36.8	232	0.19	1.18
Schillinger 388.TC	3.8	0.16 ab	36.9	201	0.17	1.20
Stewart SB3819	3.8	0.07 ab	37.5	170	0.21	1.92
Lewis 391	3.9	0.03 ab	38.8	128	0.17	2.48
Crows C3908	3.9	0.06 ab	34.9	132	0.20	2.35
Hoffman H419	4.1	0.12 ab	36.5	138	0.23	0.87
Gateway 427	4.2	0.09 ab	38.8	151	0.30	0.93
difference ^b		*	ns	ns	ns	ns

^aDifferent letters in the same column indicate statistical difference among the values. ^bStatistical difference at the 0.05 significance level.

Table 5. Mean Values of Five Endogenous Allergens (mg/g Fresh Weight) across Seasons Associated with Soybean Field Trials^a

growing season	Gly m 4	Gly m 5	Gly m 6	Gly m Bd 28k	Gly m Bd 30k
2009	0.13 a	37.9 a	218 a	0.17 b	1.23 c
2010	0.15 a	38.5 a	212 a	0.18 b	1.38 b
2012	0.05 b	35.0 a	158 b	0.18 b	2.34 a
2013	0.13 a	27.0 b	141 b	0.17 b	0.99 de
2013/2014	0.11 a	38.3 ab	128 b	0.28 a	0.84 e
difference ^b	*	*	*	*	*

^aDifferent letters indicate statistical difference among values of the same column. ^bStatistical difference at the 0.05 significance level.

associated maturity groups. The 37 conventional soybean varieties were genetically diverse as they were developed by independent breeding programs (e.g., universities, seed companies) and were released/commercialized in different years representing different breeding eras. The soybean varieties included in the study showed no significant differences for four of five evaluated soybean allergens (Table 4). Only Gly

m 4 showed some variability depending on soybean variety. Specifically, A3525, Crows C2804, Crows C37003N, C3211N, Dwight, Garst 3585N, and Schillinger TP31834 were all statistically significantly different with relatively low values (0.05–0.10 mg/g fw), compared to Maverick, Midland 363, NE3202, Stine 3300-0, and Williams 82, which had relatively high levels (0.13–0.15 mg/g fw) of Gly m 4. However, the rest

of the varieties (68%) did not significantly differ from either of the two groups.

Environmental Factors as Sources of Variation.

Environmental factors associated with the soybean field trials in this study were growing seasons, locations (grouped by states/provinces), planting time, and duration of each growing season. Soybean samples were collected during five growing seasons, and there was a significant seasonal effect observed on the levels of all five soybean allergens (Table 5; Figure 2A). The

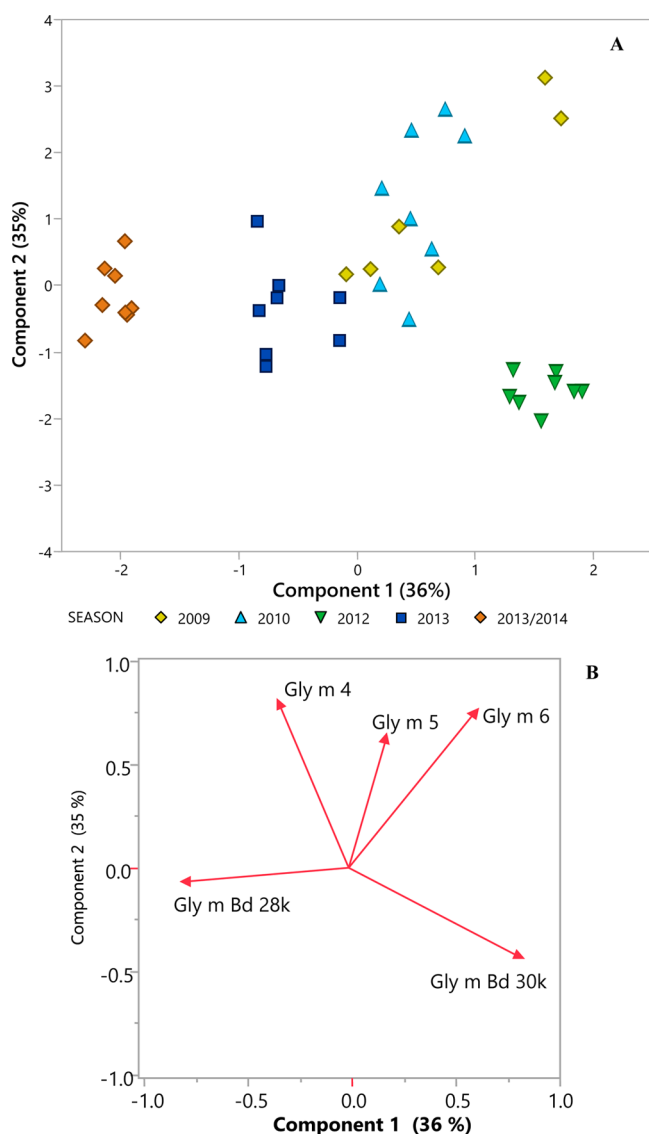


Figure 2. Results of principal component analysis: (A) component scores for soybean allergen profiles labeled by growing seasons (each point within growing season represents specific location); (B) correlations of individual soybean allergens with the principal components.

PCA scores plot reveals that, when looking at levels of all five allergens together, the composite values are mostly clustered by growing seasons (Figure 2A). Most notable, the 2012 U.S. locations were tightly clustered on one end of the graph, whereas the 2013/2014 Argentina locations formed a compact cluster on the opposite side of the scores plot. This suggests that environmental conditions for these two seasons had unique effects on allergen levels, distinct from the environ-

mental conditions for the other three seasons (2009, 2010, and 2013) that showed allergen levels between these two extreme for growing seasons.

Significant differences among the seasons were observed for each of the measured allergens; however, they were not affected in the same way (Table 5). A severe drought in the 2012 U.S. season resulted in a level of Gly m 4 (0.05 mg/g fw) that was approximately 2–3-fold lower than the Gly m 4 levels observed in the other four growing seasons (ranging from 0.11 to 0.15 mg/g fw). This is not surprising because Gly m 4 is associated with response to environmental stress.²¹ By comparison, levels of Gly m 6 showed a different trend as they decreased from the 2009 to the 2013/2014 season.

Locations (as grouped by states/provinces) had much less effect on variability of the levels of allergens evaluated in this study compared to that of seasons. Only Gly m Bd 28k showed a significant difference across locations (Table 6). The level of Gly m Bd 28k from samples grown in the two locations in Argentina (0.26–0.29 mg/g fw) was close to double the level for the other 10 states, all in the United States (0.17–0.19 mg/g fw). This can be due in part to the already discussed seasonal effect, as all of the samples from Argentina were collected in the 2013/2014 season. All of the other pairwise comparisons among the U.S. locations showed no significant difference for Gly m Bd 28k levels except for Arkansas and Kansas, but the observed differences were not exceptional.

The PCA loading plot reflects the relative contributions of the five soybean allergen proteins to principal components 1 and 2 (PC1 and PC2) (Figure 2B). For example, Gly m 5 and Gly m 6, the two major soybean seed storage proteins, have similar contributions. Both Gly m 5 and Gly m 6 are in the upper positive quadrant, reflecting that increasing levels of these proteins contribute positively to the values of both PC1 and PC2. Conversely, the other three soybean allergen proteins are not major seed storage contributors (e.g., Gly m 4 is a pathogenesis-related protein); therefore, it is not surprising that they make negative contributions to at least one, if not both, principal components. Specifically, an increase in the level of Gly m Bd 28k contributes negatively to both PC1 and PC2 (i.e., lower negative quadrant). An increase in the level of Gly m 4 contributes positively to PC2, but negatively to PC1 (i.e., upper negative quadrant). An increase in the level of Gly m Bd 30k has the opposite effect: it contributes positively to PC1 but negatively to PC2.

DISCUSSION

There has been a wide adoption of genetically modified (GM) soybeans by farmers globally, but especially in North and South America, over the past 20 years.³⁷ To enable commercialization of GM crops, a comprehensive safety assessment is done to evaluate possible food/feed safety and environmental risks, including an evaluation of their potential allergenicity derived from either the introduced protein or any food or feed derived from the GM crops.^{38,39} Assessments of allergy risks are to ensure (1) that any introduced protein is unlikely to be allergenic in the diet and (2) that the process of introducing a gene into a plant has not inadvertently increased the levels of endogenous allergens, especially in crops such as soybean that are considered commonly allergenic foods.³⁸ Understanding the natural variability of allergen levels is critical to being able to assess whether levels of endogenous allergens have been inadvertently affected during development of new GM crops. This is an especially irrelevant question given that individuals

Table 6. Mean Values of Five Endogenous Allergens (mg/g Fresh Weight) across States/Provinces Associated with Soybean Field Trials^a

state/province	Gly m 4	Gly m 5	Gly m 6	Gly m Bd 28k	Gly m Bd 30k
Arkansas	0.08	30.6	157	0.17 d	1.53
Illinois	0.13	33.4	184	0.18 cd	1.56
Indiana	0.10	36.4	206	0.18 cd	1.30
Iowa	0.12	42.5	206	0.18 cd	1.61
Kansas	0.09	29.7	152	0.19 c	1.53
Missouri	0.12	28.7	193	0.18 cd	1.23
Nebraska	0.12	36.3	184	0.17 cd	1.55
North Carolina	0.10	27.4	139	0.17 cd	1.75
Ohio	0.12	26.0	167	0.17 cd	1.04
Pennsylvania	0.13	33.9	174	0.17 cd	1.51
Buenos Aires	0.11	39.5	128	0.29 a	0.83
Santa Fe	0.13	30.3	124	0.26 b	0.89
difference ^b	ns	ns	ns	*	ns

^aDifferent letters indicate statistical difference among values of the same column. ^bStatistical difference at the 0.05 significance level.

that are allergic to a specific food are typically instructed to completely avoid eating foods containing ingredients from that allergenic source.^{40–42} If there is a large natural variability observed in conventional varieties, the value for extensive measurements of these allergens in GM crops is questionable, especially given that, to date, the effect of plant transformation on allergen levels is small and typically not statistically significantly different from levels in the conventional comparators.⁴³

This study demonstrated that endogenous allergen levels can vary from 5- to 19-fold for samples from 5 growing seasons in North and South America. For this study, 604 soybean seed samples from 37 different conventional varieties were analyzed for the levels of 5 endogenous allergens. The allergen results are consistent with previous studies on compositional differences, which showed large natural variability of compositional components.^{43–47} Gly m 4 showed the highest variation, which could be related to the pathogenesis-resistant property of this protein²⁰ and differential environmental stresses associated with the various growing seasons. By comparison, Gly m 6 levels showed less variability, which makes sense given that Gly m 6 is a major proportion of all stored soybean protein. VCA showed that the environmental factors associated with the various growing seasons for these samples were the major contributor to the natural variability of most soybean allergens (Figure 1). More specifically, growing season was the environmental factor that had the strongest effect on levels of the tested allergens (Table 5). This is not a surprise, because year-to-year differences occur due to many environmental factors: abiotic stresses, such as drought, high/low temperatures, high/low precipitation, and biotic stresses, such as damage due to plant pathogens, insect pests, or weed pressure that each varies year-to-year. PCA identified that endogenous allergen levels for samples collected from two growing seasons (2012 in the United States and 2013/2014 in Argentina) clustered distinctly and differently from the allergen levels in samples from the other three growing seasons (Figure 2A). Examining the environmental conditions for the 2012 and 2013/2014 growing seasons (Table 1) reveals that the two seasons were on the extreme ends of midseason temperatures. In the United States in 2012, both the minimum and maximum midseason temperatures were the highest across all five seasons, whereas the maximum temperature in the Argentina 2013/

2014 season was the lowest. Furthermore, the 2013/2014 growing season received much more water than the other four growing seasons. Therefore, temperature and water may have contributed to the allergen levels in samples from these two seasons clustering separately from the other seasons.

Other genetic and environmental factors, such as maturity groups, planting time, and duration of growth season, could also affect grain yield.^{48,49} The present results indicated that variability of allergen levels was not associated with either of these two environmental factors (data not shown). Others have also showed that different planting dates did not significantly affect Gly m 5 (β -conglycinin) and Gly m 6 (glycinin) levels.⁵⁰

In summary, the present study shows that environmental factors across the five growing seasons from which soybean samples were collected result in up to 19-fold differences in endogenous allergen levels in the 37 conventional soybean varieties. Justifiably, there are no regulatory requirements to measure allergens in conventionally cultivated soybeans from each growing season, which is understandable given the long history of safe consumption of soybean-derived foods despite the large natural variability of the allergens in the food supply. By comparison, the effect of plant transformation on allergen levels when GM crops are compared to their conventional comparators is minimal and typically not even statistically significantly different, bringing into question whether quantitative comparison of endogenous allergen levels of new GM plant adds any meaningful information to the overall risk assessment of the allergenic potential of the new GM varieties.

AUTHOR INFORMATION

Corresponding Author

*(T.G.) E-mail: tao.geng@monsanto.com.

ORCID

Tao Geng: 0000-0002-7367-4693

Kevin Glenn: 0000-0002-7600-2550

Present Address

§(J.W.) Royal Canin USA, 500 Fountain Lakes Boulevard, Suite 100, St. Charles, MO 63301, USA.

Funding

This research received financial support from Monsanto Co.

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Wilson, R. F. Seed composition. In *Soybeans: Improvement, Production and Uses*, 3rd ed.; Boerma, H. R., Specht, J. E., Eds.; ASA, CSSA, SSSA: Madison, WI, USA, 2004.
- (2) Kim, H. K.; Nelson-Dooley, C.; Della-Fera, M. A.; Yang, J. Y.; Zhang, W.; Duan, J. H.; Hartzell, D. L.; Hamrick, M. W.; Baile, C. A. Genistein decreases food intake, body weight, and fat pad weight and causes adipose tissue apoptosis in ovariectomized female mice. *J. Nutr.* **2006**, *136*, 409–414.
- (3) Omoni, A. O.; Aluko, R. E. Soybean foods and their benefits: potential mechanisms of action. *Nutr. Rev.* **2005**, *63*, 272–283.
- (4) Ali, A. A.; Velasquez, M. T.; Hansen, C. T.; Mohamed, A. I.; Bhatena, S. J. Effects of soybean isoflavones, probiotics, and their interactions on lipid metabolism and endocrine system in an animal model of obesity and diabetes. *J. Nutr. Biochem.* **2004**, *15*, 583–590.
- (5) Friedman, M.; Brandon, D. L. Nutritional and health benefits of soy proteins. *J. Agric. Food Chem.* **2001**, *49*, 1069–1086.
- (6) Carroll, K. K.; Kurowska, E. M. Soy consumption and cholesterol reduction – review of animal and human studies. *J. Nutr.* **1995**, *125*, S594–S597.
- (7) FDA. Food Allergen Labeling and Consumer Protection Act of 2004; Food and Drug Administration: Washington, DC, USA, 2004.
- (8) Hefle, S. L.; Nordlee, J. A.; Taylor, S. L. Allergenic foods. *Crit. Rev. Food Sci. Nutr.* **1996**, *36*, S69–S89.
- (9) FAO. *Food Allergies*; Report of the Technical Consultation of the Food and Agriculture Organization of the United Nations; Food and Agriculture Organization of the United Nations: Rome, Italy, 1995.
- (10) OECD. *Revised Consensus Document on Compositional Considerations for New Variety of Soybean*; Organization of Economic Cooperation and Development 2012; ENV/JM/MONO(2012)24.
- (11) Herman, E. M.; Burks, A. W. The impact of plant biotechnology on food allergy. *Curr. Opin. Biotechnol.* **2011**, *22*, 224–230.
- (12) Breiteneder, H. 2 – Classifying food allergens. In *Detecting Allergens in Food*; Koppelman, S. J., Hefle, S. L., Eds.; Woodhead Publishing: 2006; pp 21–61.
- (13) Holzhauser, T.; Wackermann, O.; Ballmer-Weber, B. K.; Bindslev-Jensen, C.; Scibilia, J.; Perono-Garoffo, L.; Utsumi, S.; Poulsen, L. K.; Vieths, S. Soybean (*Glycine max*) allergy in Europe: Gly m 5 (β -conglycinin) and Gly m 6 (glycinin) are potential diagnostic markers for severe allergic reactions to soy. *J. Allergy Clin. Immunol.* **2009**, *123*, 452–458.e4.
- (14) Herman, E. M.; Helm, R. M.; Jung, R.; Kinney, A. J. Genetic modification removes an immunodominant allergen from soybean. *Plant Physiol.* **2003**, *132*, 36–43.
- (15) Tsuji, H.; Bando, N.; Hiemori, M.; Yamanishi, R.; Kimoto, M.; Nishikawa, K.; Ogawa, T. Purification and characterization of soybean allergen Gly m Bd 28K. *Biosci., Biotechnol., Biochem.* **1997**, *61*, 942–947.
- (16) Hiemori, M.; Bando, N.; Ogawa, T.; Shimada, H.; Tsuji, H.; Yamanishi, R.; Terao, J. Occurrence of IgE antibody-recognizing N-linked glycan moiety of a soybean allergen, Gly m Bd 28K. *Int. Arch. Allergy Immunol.* **2000**, *122*, 238–245.
- (17) Tsuji, H.; Hiemori, M.; Kimoto, M.; Yamashita, H.; Kobatake, R.; Adachi, M.; Fukuda, T.; Bando, N.; Okita, M.; Utsumi, S. Cloning of cDNA encoding a soybean allergen, Gly m Bd 28K. *Biochim. Biophys. Acta, Gene Struct. Expression* **2001**, *1518*, 178–182.
- (18) Ogawa, T.; Bando, N.; Yamanishi, R.; Tsuji, H.; Hiemori, M. Characterization of Gly m Bd 28K, an allergenic glycoprotein from soybean. *J. Allergy Clin. Immunol.* **1998**, *101*, S90–S91.
- (19) Ogawa, T.; Tsuji, H.; Bando, N.; Kitamura, K.; Zhu, Y.-L.; Hirano, H.; Nishikawa, K. Identification of the soybean allergenic protein, Gly m Bd 30K, with the soybean seed 34-kDa oil-body-associated protein. *Biosci., Biotechnol., Biochem.* **1993**, *57*, 1030–1033.
- (20) Radauer, C.; Breiteneder, H. Evolutionary biology of plant food allergens. *J. Allergy Clin. Immunol.* **2007**, *120*, S18–S25.
- (21) Berkner, H.; Neudecker, P.; Mittag, D.; Ballmer-Weber, B. K.; Schweimer, K.; Vieths, S.; Rosch, P. Cross-reactivity of pollen and food allergens: soybean Gly m 4 is a member of the Bet v 1 superfamily and closely resembles yellow lupine proteins. *Biosci. Rep.* **2009**, *29*, 183–192.
- (22) Mittag, D.; Vieths, S.; Vogel, L.; Becker, W. M.; Rihs, H. P.; Helbling, A.; Wuthrich, B.; Ballmer-Weber, B. K. Soybean allergy in patients allergic to birch pollen: clinical investigation and molecular characterization of allergens. *J. Allergy Clin. Immunol.* **2004**, *113*, 148–154.
- (23) Kleine-Tebbe, J.; Wangorsch, A.; Vogel, L.; Crowell, D. N.; Hausteiner, U. F.; Vieths, S. Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1-related PR-10 protein in soybean, SAM22. *J. Allergy Clin. Immunol.* **2002**, *110*, 797–804.
- (24) EC. Commission Implementing Regulation (EU) No. 503/2013 on applications for authorisation of genetically modified food and feed in accordance with Regulation (EC) No. 1829/2003 of the European Parliament and of the Council and amending Commission Regulations (EC) No. 641/2004 and (EC) No. 1981/2006. *Off. J. Eur. Union* **2013**, *L57*, 1–48.
- (25) Ladics, G. S.; Budziszewski, G. J.; Herman, R. A.; Herouet-Guicheney, C.; Joshi, S.; Lipscomb, E. A.; McClain, S.; Ward, J. M. Measurement of endogenous allergens in genetically modified soybeans – short communication. *Regul. Toxicol. Pharmacol.* **2014**, *70*, 75–79.
- (26) Sten, E.; Skov, P. S.; Andersen, S. B.; Torp, A. M.; Olesen, A.; Bindslev-Jensen, U.; Poulsen, L. K.; Bindslev-Jensen, C. A comparative study of the allergenic potency of wild-type and glyphosate-tolerant gene-modified soybean cultivars. *Apmis* **2004**, *112*, 21–28.
- (27) Codina, R.; Arduoso, L.; Lockey, R. F.; Crisci, C.; Medina, I. Allergenicity of varieties of soybean. *Allergy* **2003**, *58*, 1293–1298.
- (28) Geng, T.; Liu, K.; Frazier, R.; Shi, L.; Bell, E.; Glenn, K.; Ward, J. M. Development of a sandwich ELISA for quantification of Gly m 4, a soybean allergen. *J. Agric. Food Chem.* **2015**, *63*, 4947–4953.
- (29) Liu, B.; Teng, D.; Wang, X. M.; Yang, Y. L.; Wang, J. H. Expression of the soybean allergenic protein P34 in *Escherichia coli* and its indirect ELISA detection method. *Appl. Microbiol. Biotechnol.* **2012**, *94*, 1337–1345.
- (30) Tsuji, H.; Okada, N.; Yamanishi, R.; Bando, N.; Kimoto, M.; Ogawa, T. Measurement of Gly-M-Bd-30k, a major soybean allergen, in soybean products by a sandwich enzyme-linked-immunosorbent assay. *Biosci., Biotechnol., Biochem.* **1995**, *59*, 150–151.
- (31) Liu, B.; Teng, D.; Wang, X. M.; Wang, J. H. Detection of the soybean allergenic protein Gly m Bd 28K by an indirect enzyme-linked immunosorbent assay. *J. Agric. Food Chem.* **2013**, *61*, 822–828.
- (32) Bando, N.; Tsuji, H.; Hiemori, M.; Yoshizumi, K.; Yamanishi, R.; Kimoto, M.; Ogawa, T. Quantitative analysis of Gly m Bd 28K in soybean products by a sandwich enzyme-linked immunosorbent assay. *J. Nutr. Sci. Vitaminol.* **1998**, *44*, 655–664.
- (33) Ma, X.; Sun, P.; He, P. L.; Han, P. F.; Wang, J. J.; Qiao, S. Y.; Li, D. F. Development of monoclonal antibodies and a competitive ELISA detection method for glycinin, an allergen in soybean. *Food Chem.* **2010**, *121*, S46–S51.
- (34) Liu, B.; Teng, D.; Yang, Y. L.; Wang, X. M.; Wang, J. H. Development of a competitive ELISA for the detection of soybean alpha subunit of beta-conglycinin. *Process Biochem.* **2012**, *47*, 280–287.
- (35) You, J. M.; Li, D.; Qiao, S. Y.; Wang, Z. R.; He, P. L.; Ou, D. Y.; Dong, B. Development of a monoclonal antibody-based competitive ELISA for detection of beta-conglycinin, an allergen from soybean. *Food Chem.* **2008**, *106*, 352–360.
- (36) Julka, S.; Kuppannan, K.; Karnoup, A.; Dielman, D.; Schafer, B.; Young, S. A. Quantification of Gly m 4 protein, a major soybean allergen, by two-dimensional liquid chromatography with ultraviolet and mass spectrometry detection. *Anal. Chem.* **2012**, *84*, 10019–10030.
- (37) James, C. Global status of commercialized biotech/gm crops: 2014. *ISAAA Brief.* **2014**, *49*, 303–309.
- (38) Codex. Foods derived from modern biotechnology. Codex Alimentarius Commission, Joint FAO/WHO Food Standards Programme, Rome, Italy, 2009.
- (39) EFSA, European Food Safety Agency. Guidance document on the scientific panel on genetically modified organisms for the risk

assessment of genetically modified plants and derived food and feed. *EFSA J.* **2004**, *99*, 1–94.

(40) Goodman, R. E.; Vieths, S.; Sampson, H. A.; Hill, D.; Ebisawa, M.; Taylor, S. L.; van Ree, R. Allergenicity assessment of genetically modified crops—what makes sense? *Nat. Biotechnol.* **2008**, *26*, 73–81.

(41) Goodman, R. E.; Panda, R.; Ariyaratna, H. Evaluation of endogenous allergens for the safety evaluation of genetically engineered food crops: review of potential risks, test methods, examples and relevance. *J. Agric. Food Chem.* **2013**, *61*, 8317–8332.

(42) Herman, R. A.; Ladics, G. S. Endogenous allergen upregulation: transgenic vs. traditionally bred crops. *Food Chem. Toxicol.* **2011**, *49*, 2667–2669.

(43) Harrigan, G. G.; Lundry, D.; Drury, S.; Berman, K.; Riordan, S. G.; Nemeth, M. A.; Ridley, W. P.; Glenn, K. C. Natural variation in crop composition and the impact of transgenesis. *Nat. Biotechnol.* **2010**, *28*, 402–404.

(44) Harrigan, G. G.; Skogerson, K.; MacIsaac, S.; Bickel, A.; Perez, T.; Li, X. Application of ^1H NMR profiling to assess seed metabolomic diversity. A case study on a soybean era population. *J. Agric. Food Chem.* **2015**, *63*, 4690–4697.

(45) Berman, K. H.; Harrigan, G. G.; Nemeth, M. A.; Oliveira, W. S.; Berger, G. U.; Tagliaferro, F. S. Compositional equivalence of insect-protected glyphosate-tolerant soybean MON 87701 $\times$ MON 89788 to conventional soybean extends across different world regions and multiple growing seasons. *J. Agric. Food Chem.* **2011**, *59*, 11643–11651.

(46) Zhou, J.; Berman, K. H.; Breeze, M. L.; Nemeth, M. A.; Oliveira, W. S.; Braga, D. P.; Berger, G. U.; Harrigan, G. G. Compositional variability in conventional and glyphosate-tolerant soybean (*Glycine max* L.) varieties grown in different regions in Brazil. *J. Agric. Food Chem.* **2011**, *59*, 11652–11656.

(47) Zhou, J.; Harrigan, G. G.; Berman, K. H.; Webb, E. G.; Klusmeyer, T. H.; Nemeth, M. A. Stability in the composition equivalence of grain from insect-protected maize and seed from glyphosate-tolerant soybean to conventional counterparts over multiple seasons, locations, and breeding germplasms. *J. Agric. Food Chem.* **2011**, *59*, 8822–8828.

(48) Heatherly, L. G.; Elmore, R. W. *Soybeans: Improvement, Production and Uses*, 3rd ed.; ASA, CSSA, SSA: Madison, WI, USA, 2004.

(49) Piper, E. L.; Boote, K. J. Temperature and cultivar effects on soybean seed oil and protein concentrations. *J. Am. Oil Chem. Soc.* **1999**, *76*, 1233–1241.

(50) Jenkinson, J. E.; Fehr, W. R. Influence of locations and planting dates on protein composition of soybean lines with modified beta-conglycinin and glycinin concentration. *Crop Sci.* **2010**, *50*, 1805–1810.