Application to Food Standards Australia New Zealand for the Inclusion of Corn MON87403 in *Standard 1.5.2 - Food Derived from Gene Technology*

Submitted by:

Monsanto Australia Limited
Level 12 / 600 St Kilda Road,
Melbourne Victoria 3004

3 June 2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

UNPUBLISHED REPORTS BEING SUBMITTED

[redacted]. 2015. Assessment of ATHB17113 Protein Levels in Leaf, Root, Forage, and Grain Tissues Collected from Maize MON 87403 Produced in United States Field Trials during 2012. **MSL0026598**. Monsanto Company.

[redacted] 2014. Amended Report for MSL0025447: Characterization of the ATHB17Δ113 Protein Purified from the Maize Leaf of MON 87403 and Comparison of the Physiochemical Properties of the Plant-Produced and *Escherichia coli* (*E. coli*)-Produced ATHB17Δ113 Proteins. **MSL0025829**. Monsanto Company.

[redacted]. 2015. Characterization of the DNA Binding Activity of the *Escherichia coli*-produced ATHB17Δ113 protein. **MSL0026645**. Monsanto Company.

[redacted]. 2015. Characterization of the MON 87403 produced ATHB17Δ113 protein by Mass Spectrometry. **MSL0026621**. Monsanto Company.

[redacted]. 2014. Amended Report for MSL0025316: Molecular Characterization of MON 87403. **MSL0025909.** Monsanto Company.

[redacted]. 2013. Bioinformatics Evaluation of the ATHB17Δ113 Protein Utilizing the AD_2013, TOX_2013 and PRT_2013 Databases. **MSL0025242**. Monsanto Company.

[redacted]. 2014. Bioinformatics Evaluation of DNA Sequences Flanking the 5' and 3' Junctions of Inserted DNA in MON 87403: Assessment of Putative Polypeptides. **MSL0025733**. Monsanto Company.

[redacted]. 2014. Bioinformatics Evaluation of the Transfer DNA Insert in MON 87403 Utilizing the AD_2014, TOX_2014 and PRT_2014 Databases. **MSL0025648**. Monsanto Company.

[redacted]. 2015. Assessment of R1 Ear Biomass of Maize MON 87403 in 2012 U.S. Field Trials. **MSL0026473**. Monsanto Company.

[redacted]. 2013. Segregation of the T-DNA Insert in MON 87403 Across Three Generations. **MSL0024676**. Monsanto Company.

[redacted] 2013. An Acute Toxicity Study of *E. coli*-produced ATHB17Δ113 protein by Oral Gavage in Mice. **CRO-2013-0121**. Monsanto Company.

[redacted]. 2014. Effect of Heat Treatment on the Functional Activity of *Escherichia coli* (*E. coli*)-Produced ATHB17Δ113 Protein. **MSL0025364**. Monsanto Company.

[redacted]. 2015. Amended Report for MSL0025076: Compositional Analyses of Maize Forage and Grain from MON 87403 Grown in the United States during 2012. **MSL0025787**. Monsanto Company.

[redacted]. 2014. Assessment of the *in vitro* Digestibility of ATHB17Δ113 Protein in Simulated Gastric and Simulated Intestinal Fluids. **MSL0025516**. Monsanto Company.

CHECKLIST

General Requirements (3.1)	Reference
3.1.1 Form of application	
☑ Executive Summary	*Executive Summary*
☑ Relevant sections of Part 3 identified	
☑ Pages sequentially numbered	
☑ Electronic + 2 hard copies	
☑ Electronic and hard copies identical	
☑ Hard copies capable of being laid flat	
☑ All references provided	
3.1.2 Applicant details	*Page 1*
3.1.3 Purpose of the application	*Page 1*
3.1.4 Justification for the application	*Page 2*
3.1.5 Information to support the application	*Volumes 1 - 100*
3.1.6 Assessment procedure	*Page 3*
☑ General	
☐ Major	
☐ Minor	
3.1.7 Confidential Commercial Information	
☑ Confidential material separated in both electronic and hard copy	
☑ Justification provided	
3.1.8 Exclusive Capturable Commercial Benefit	*Page 3*
3.1.9 International and Other National Standards	*Page 4*
3.1.10 Statutory Declaration	*Page 116*
3.1.11 Checklist/s provided with Application	
☑ Checklist	
☑ Any other relevant checklists for Sections 3.2 – 3.7	*Checklist 3.5.1*

Foods Produced using Gene Technology (3.5.1)	
☑ A.1 Nature and identity of GM food	*Page 5*
☑ A.2 History of use of host and donor organisms	*Page 8*
☑ A.3 Nature of genetic modification	*Page 12*
☑ A.4 Analytical method for detection	*Page 53*
☑ B.1 Equivalence studies	*Page 54*
☑ B.2 Antibiotic resistance marker genes (if used)	*Page 69*
☑ B.3 Characterisation of novel protein(s)/substances	*Page 70*
☑ B.4 Potential toxicity of novel protein(s)/substances	*Page 76*
☑ B.5 Potential allergenicity of novel protein(s)	*Page 90*
☑ B.6 Toxicity of novel herbicide metabolites	*Page 93*
☑ B.7 Compositional Analyses	*Page 94*
☑ C.1 Nutritional impact of GM food	*Page 115*
☑ C.2 Animal feeding studies (if available)	*Page 115*

ABBREVIATIONS AND DEFINITIONS[1]

AA	amino acid
ADF	acid detergent fiber
APHIS	Animal and Plant Health Inspection Service
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
bp	base pair
bwt	body weight
CAB	chlorophyll a/b-binding
CaMV	cauliflower mosaic virus
CFR	Code of Federal Regulations
COA	certificate of analysis
CV	coefficient of variation
DAP	days after planting
DDI	daily dietary intake
DF	dilution factor
DNA	deoxyribonucleic acid
DTT	dithiothreitol
DWCF	dry weight conversion factor
dwt	dry weight
E. coli	*Escherichia coli*
EAR	ERF-associated amphilic repression
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
E-score	expectation score
FA	fatty acid
FDA	Food and Drug Adminstration (U.S.)
FSANZ	Food Standards Australia and New Zealand
fwt	fresh weight
GLP	good laboratory practice
GRAS	generally recognized as safe
HD	homeodomain
HD-Zip	homeodomain-leucine zipper
HPLC	high-performance liquid chromatography
IgG	immunoglobulin G
ILSI	International Life Science Institute
JSC	junction sequence class
kb	kilobase
kDa	kilodalton
LB	loading buffer
LOD	limit of detection
LOQ	limit of quantitation

[1] Alred, G.J., C.T. Brusaw, and W.E. Oliu. 2003. Handbook of Technical Writing, 7th edn., pp. 2-7. Bedford/St. Martin's, Boston, MA.

LZ	leucine zipper
mRNA	messenger RNA
MALDI-TOF MS	matrix assisted laser desorption/ionization - time of flight mass spectrometry
MOE	margin of exposure
MW	molecular weight
MWM	molecular weight marker
NDF	neutral detergent fiber
NGS/JSA	Next-Generation Sequencing / Junction Sequence Analysis
NOAEL	No Observable Adverse Effect Level
OECD	Organisation for Economic Co-operation and Development
OPA	O-phthalaldehyde
ORF	open reading frame
PCR	polymerase chain reaction
ppm	parts per million
ROP	Repressor of primer
RT	room temperature
RT-PCR	reverse transcription polymerase chain reaction
SD	standard deviation
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	standard error
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
T-DNA	Transfer DNA
TDF	total dietary fiber
TFA	trifluoroacetic Acid
U	units
USDA	United States Department of Agriculture
UV	ultraviolet
v/v	volume to volume ratio
w/v	weight to volume ratio

PART 1 GENERAL INFORMATION

1.1 Applicant Details

(a) Applicant's name/s	[redacted]
(b) Company/organisation name	Monsanto Australia Limited
(c) Address (street and postal)	Level 12 / 600 St Kilda Road, Melbourne, Victoria, 3004 PO Box 6051, St Kilda Road Central, Victoria, 8008
(d) Telephone number	+61 3 9522 7101
(e) Email address	[redacted]
(f) Nature of applicant's business	Technology Provider to the Agricultural and Food Industries
(g) Details of other individuals, companies or organisations associated with the application	

1.2 Purpose of the Application

This application is submitted to Food Standards Australia New Zealand by Monsanto Australia Limited on behalf of Monsanto Company.

The purpose of this submission is to make an application to vary **Standard 1.5.2** – *Food Produced Using Gene Technology* of the *Australia New Zealand Food Standards Code* to seek the addition of maize line MON 87403 and products containing maize line MON 87403 (hereafter referred to as MON 87403) to the Table to Clause 2 (see below).

Food derived from gene technology	Special requirements
Food derived from maize line MON 87403	None

1.3 Justification for the Application

1.3(a) The need for the proposed change

Monsanto Company has developed biotechnology-derived maize MON 87403 that has increased ear biomass at an early reproductive stage (R1) compared to conventional control maize. Insertion of the coding region of the Arabidopsis *ATHB17* gene results in production of a truncated ATHB17 protein (ATHB17Δ113) in MON 87403. ATHB17 is a member of the HD-Zip family of plant transcription factors, which are proteins that bind to specific DNA sequences and regulate gene expression. The HD-Zip family of proteins is found broadly across plant species and specific HD-Zip proteins have been shown to play important roles in the modulation of plant growth and development. Increased ear biomass in MON 87403 is associated with increased partitioning of dry matter (photosynthate) from the source (vegetative) tissue to the sink (ear) tissue.

1.3(b) The advantages of the proposed change over the status quo, taking into account any disadvantages

The early reproductive stages in maize are a critical period of maize growth when the maximum ear biomass (sink size) is determined by a combination of genetics and environmental conditions. Dry matter (photosynthate) produced by the plant during reproductive stages is allocated to the ear for its growth after the sink size is established. Thus, ear biomass, which is set during early reproductive stages, is considered an important determinant of future reproductive success such that a larger ear biomass at early reproductive stages is associated with increased grain yield opportunity at harvest.

MON 87403 will be combined with other deregulated biotechnology-derived traits through traditional breeding methods to create commercial products with increased yield opportunity as well as protection against maize insect pests and tolerance to herbicides. These next generation combined-trait maize products will continue to offer growers a broad choice of trait combinations and continued pest control durability.

1.4 Regulatory Impact Information

1.4(a) Costs and benefits

If the draft variation to permit the sale and use of food derived from MON 87403 is approved, possible affected parties may include consumers, industry sectors and government. The consumers who may be affected are those particularly concerned about the use of biotechnology. Industry sectors affected may be food importers and exporters, distributors, processors and manufacturers. Lastly, government enforcement agencies may be affected.

A cost/benefit analysis quantified in monetary terms is difficult to determine. In fact, most of the impacts that need to be considered cannot be assigned a dollar value. Criteria would need to be deliberately limited to those involving broad areas such as trade, consumer information and compliance. If the draft variation is approved:

Consumers:

- There would be benefits in the broader availability of corn products.

- There is unlikely to be any significant increase in the prices of foods if manufacturers are able to use comingled corn products.
- Consumers wishing to do so will be able to avoid GM corn products as a result of labeling requirements and marketing activities.

Government:

- Benefit that if corn MON 87403 was detected in food products, approval would ensure compliance of those products with the Code. This would ensure no potential for trade disruption on regulatory grounds.
- Approval of corn MON 87403 would ensure no potential conflict with WTO responsibilities.
- In the case of approved GM foods, monitoring is required to ensure compliance with the labeling requirements, and in the case of GM foods that have not been approved, monitoring is required to ensure they are not illegally entering the food supply. The costs of monitoring are thus expected to be comparable, whether a GM food is approved or not.

Industry:

- Sellers of processed foods containing corn derivatives would benefit as foods derived from corn MON 87403 would be compliant with the Code, allowing broader market access and increased choice in raw materials. Retailers may be able to offer a broader range of corn products or imported foods manufactured using corn derivatives.
- Possible cost to food industry as some food ingredients derived from corn MON 87403 would be required to be labelled

1.4(b) Impact on international trade

If the draft variation to permit the sale and use of food derived from MON 87403 was rejected it would result in the requirement for segregation of any corn derived products containing MON 87403 from those containing approved corn, which would be likely to increase the costs of imported corn derived foods.

It is important to note that if the draft variation is approved, corn MON 87403 will not have a mandatory introduction. The consumer will always have the right to choose not to use/consume this product.

1.5 Assessment Procedure

Monsanto Australia is submitting this application in anticipation that it will fall within the General Procedure category.

1.6 Exclusive Capturable Commercial Benefit

This application is likely to result in an amendment to the Code that provides exclusive benefits and therefore Monsanto intends to pay the full cost of processing the application.

1.7 International and Other National Standards

1.7(a) International standards

Monsanto makes all efforts to ensure that safety assessments are aligned, as closely as possible, with relevant international standards such as the Codex Alimentarius Commission's *Principles for the Risk Analysis of Foods Derived from Modern Biotechnology* and supporting *Guideline for the Conduct of Food Safety Assessment of Foods Derived from Recombinant-DNA Plants* (Codex Alimentarius, 2009).

In addition, the composition analysis is conducted in accordance with OECD guidelines and includes the measurement of OECD-defined corn nutrients and anti-nutrients based on conventional commercial corn varieties (OECD, 2002a).

1.7(b) Other national standards or regulations

Monsanto has submitted a food and feed safety and nutritional assessment summary for MON 87403 to the United States Food and Drug Administration (FDA) and has also requested a Determination of Nonregulated Status for MON 87403, including all progenies derived from crosses between MON 87403 and other corn, from the Animal and Plant Health Inspection Service (APHIS) of the U.S. Department of Agriculture (USDA).

Applications have also been and will be submitted to Canadian Food Inspection Agency (CFIA) and Health Canada (HC), Korea's Ministry of Food and Drug Safety (MFDS) for food, and Rural Development Administration (RDA) for feed use, and Japan's Ministry of Health, Labour, and Welfare (MHLW) for food use.

Regulatory submissions will be made to countries that import significant corn or food and feed products derived from countries where MON 87403 corn will be grown and have functional regulatory review processes in place. This will result in submissions to a number of additional governmental regulatory agencies including, but not limited to Ministry of Agriculture, People's Republic of China; Japan's Ministry of Agriculture, Forestry, and Fisheries, as well as to regulatory authorities in other corn importing countries with functioning regulatory systems.

PART 2 SPECIFIC DATA REQUIREMENTS FOR SAFETY ASSESSMENT

A. TECHNICAL INFORMATION ON THE GM FOOD

A1 Nature and Identity of the Genetically Modified Food

A1(a) A description of the new GM organism

Maize is one of the largest U.S. crops based on acreage and quantity harvested each year. In 2013 maize was planted on 35.48 million ha in the United States (USDA-FAS, 2014). Because of its importance, plant breeders and producers continuously strive to improve commercial maize yield. Initial improvements in maize yield were due to the domestication of varieties with desirable traits like larger ear biomass. Maize ears became larger over time during the domestication era (University of Utah, 2014). During the hybrid era (1939 to present), commercial maize yield in the U.S. increased nearly six-fold with an average 99 kg ha^{-1} increase every year (Lee and Tollenaar, 2007). The major factor that contributed to yield increase during the hybrid era is a favorable response of hybrids to increased plant population density which resulted in an increase in the number of ears per hectare (Bruns and Abbas, 2003). However, at the individual plant level, yield increase was associated with a decrease in ear biomass. Commercial varieties with MON 87403 can provide increased yield opportunity as this trait increases ear biomass during early reproductive stages.

Monsanto Company has developed biotechnology-derived maize MON 87403 that has increased ear biomass at an early reproductive growth stage (R1) compared to conventional control maize. MON 87403 was produced through insertion of the coding region of the full-length *Arabidopsis thaliana ATHB17* gene through *Agrobacterium*-mediated transformation. ATHB17 is a member of the HD-Zip family of plant transcription factors, which are proteins that bind to specific DNA sequences and regulate gene expression. The HD-Zip family of proteins is found broadly across plant species and specific HD-Zip proteins have been shown to play an important role in the modulation of plant growth and development. The HD-Zip family consists of four subfamilies and ATHB17 is a member of the class II subfamily. HD-Zip II proteins form either homodimers or heterodimers with other HD-Zip II proteins within the same subfamily and function as repressors of gene expression. In MON 87403, maize-specific splicing of the *ATHB17* mRNA transcript results in a truncated protein, ATHB17Δ113, which is missing the first 113 N-terminal amino acids that are present in *Arabidopsis thaliana*. The ATHB17Δ113 protein retains the ability to form homo- and hetero-dimers and bind to target DNA sequences like the full-length protein (Rice *et al*., 2014). The ATHB17Δ113 protein is, however, unable to function as a transcriptional repressor because the protein lacks a functional repression domain. By a dominant-negative mechanism, ATHB17Δ113 can alter the activity of endogenous maize HD-Zip II proteins, which are predominantly expressed in ear tissue (Rice *et al*., 2014). Thus, the ATHB17Δ113 protein likely modulates HD-Zip II-regulated pathways in the ear, which leads to increased ear growth at an early reproductive stage.

Given the expected growth in human population, increased agricultural output will be required to come from increased productivity (i.e., yield per unit area) as opposed to an increase in area under production (OECD-FAO, 2008). Agricultural biotechnology is one of

the tools that can be utilized to help address the increasing demand for food and feed due to this population growth. Augmenting gains in yield opportunity with biotechnology traits as well as with continual gains from traditional breeding would have a positive impact on the U.S. economy. For example, increasing the average rate of yield gain by just 1 bu/ac/year, over and above historical average yield gains (~1.6 bu/ac/year) for the next ten years, would have a net economic impact of $16B USD in the U.S. (Leibman *et al*., 2014).

Early reproductive stages in maize have been identified as a crucial phase during maize growth when a determination of maximum ear biomass is determined by both plant genetics and environmental conditions (Borrás and Westgate, 2006; Jones *et al*., 1996). Dry matter produced by the plant during reproductive stages is allocated to the ear for its growth (Ritchie *et al*., 1997) after the sink size is determined. Thus an increase in ear biomass that is determined during the early reproductive stages is considered an important determinant of reproductive success (Borrás *et al*., 2004; Zaidi *et al*., 2003). Published literature suggests that a larger ear biomass at early reproductive stages can result in increased kernels per hectare (Fisher and Palmer, 1983; Severini *et al*., 2011) thus potentially increasing agronomic benefit to farmers.

MON 87403 will be combined with other deregulated biotechnology-derived traits through traditional breeding methods to create commercial products with increased yield opportunity as well as protection against maize pests and tolerance to multiple herbicides. These next generation combined-trait maize products will continue to offer broader grower choice and continued pest control. Adoption of improved maize hybrids with increased yield opportunity that results in incremental increases in national average grain yield can positively impact production, exports and economic welfare.

MON 87403 Results in Increased Ear Biomass

Efficacy of MON 87403 was demonstrated by directly comparing its R1 (silking stage) ear weight to a conventional control with the same genetic background. Ear weight is defined as the total weight of the primary ear including the husk, shank, cob, silk and ovules at the R1 stage. MON 87403 and the control were grown at 13 field locations within U.S. maize production regions in 2012. MON 87403 and the control were planted (at a density of ~7 seeds/m and row spacing of 30 inches) at each site in randomized complete block designs with four replications. Measurements were collected from all of the plants at the R1 stage in a 1 m length of the designated row. Statistical comparisons were made between MON 87403 and the control across all 13 sites (combined-site analyses). The level of statistical significance for all statistical comparisons was predetermined to be 5% (α=0.05).

There was a statistically significant increase in R1 ear weight between MON 87403 and the control in the combined site analysis (Table 1). The R1 ear weight in MON 87403 was 11.7% higher than the control (Table 1).

Table 1. Differences in R1 Ear Weight between MON 87403 and the Conventional Control in 2012 U.S. Field Trials[1]

Characteristic (units)	MON 87403 (Mean ± SE)	Control (Mean ± SE)	Change (%)	p-value
R1 ear weight[2] (g)	144.50 (±8.47)	129.30 (±8.13)	11.7	0.004*

[1] Locations included in the combined-site analysis: Jackson, Arkansas; Vermilion, Illinois; Warren, Illinois; Boone, Indiana; Greene, Iowa; Jefferson, Iowa; Pawnee, Kansas; (2 sites) Polk, Nebraska; York, Nebraska; Perquimans, North Carolina; Berks, Pennsylvania; Lehigh, Pennsylvania.

[2] R1 ear weight is the dry weight of the entire primary ear at silking stage from all plants in the 1 m sampling.

*Denotes statistical difference between MON 87403 and the control ($\alpha = 0.05$). N = 51.

For details please refer to [redacted], 2015 (MSL0026473).

A1(b) Name, number or other identifier of each new line or strain

In accordance with OECD's "Guidance for the Designation of a Unique Identifier for Transgenic Plants" MON 87403 has been assigned the unique identifier MON-874Ø3-1.

A1(c) The name the food will be marketed under (if known)

Maize containing the transformation event MON 87403 will be produced in North America. There are currently no plans to produce this product in Australia and New Zealand. A commercial trade name for the product has not been determined at the time of this submission and will be available prior to commercial launch of the product in North America.

A1(d) The types of products likely to include the food or food ingredient

Maize is widely used for a variety of food and feed purposes, and it is intended that MON 87403 will be utilized in the same manner and for the same uses as conventional maize. Maize grain and its processed products are consumed in a multitude of human food and animal feed products. Maize forage (as silage) is extensively consumed as an animal feed by ruminants.

A2 History of Use of the Host and Donor Organisms

A2(a) Description of all donor organism(s)

A2(a)(i) Common and scientific names and taxonomic classification

The *ATHB17* gene is derived from *Arabidopsis thaliana*, an annual plant commonly referred to as thale or mouse-ear cress. The taxonomy of *A. thaliana* is:

Kingdom	Plantae
Phylum	Magnoliophyta
Class	Magnoliopsida
Order	Capparales
Family	Brassicaceae
Genus	Arabidopsis

A2(a)(ii) Information on pathogenicity, toxicity, allergenicity

Arabidopsis thaliana is generally not considered an allergenic or toxic source organism. Although *Arabidopsis thaliana* contains homologs of proteins previously described as allergens in other plant species (e.g., germins, lipid transfer protein, profilins, and small molecular weight calcium binding proteins), no Arabidopsis proteins have been reported in a peer-reviewed database of known allergens (FARRP, 2013). Only one case of occupational asthma has been reported in a laboratory worker due to exposure to Arabidopsis pollen (Yates *et al*., 2008).

A2(a)(iii) History of use of the organism in food supply or human exposure

Arabidopsis thaliana is a small weed that is a member of the mustard family and has a broad natural distribution throughout Europe, Asia, and North America (Meinke *et al*., 1998). *Arabidopsis thaliana* is not purposely consumed as a food source by humans, and there is no documented consumption by animals. However, certain populations of a close relative, *Arabidopsis lyrata*, have been reported to be subject to sheep grazing (Sandring *et al*., 2007). *Arabidopsis thaliana* is in the Brassicaceae family (Meinke *et al*., 1998), which contains well-known food and oilseed crops such as broccoli, cauliflower, cabbage, and canola/rapeseed. *Camelina sativa*, an emerging oilseed crop, is in the same family as and reported to be the cultivated species most closely related to Arabidopsis (Flannery *et al*., 2006). *Camelina sativa* leaves are consumed as fresh greens by humans in the country of Georgia (Facciola, 1998), the meal can be used as a component of livestock feed (AAFCO, 2011), and the plants in a crop setting are grazed by roaming wildlife (Pilgeram *et al*., 2007). The safe consumption of near relatives of *Arabidopsis thaliana* by humans and animals supports the safety of this organism.

A2(b) Description of the host organism

A2(b)(i) Phenotypic information

Maize (*Zea mays* L.) is a member of the tribe Maydae, which is included in the subfamily Panicoideae of the grass family Gramineae.

Family - Gramineae
- Subfamily - Panicoideae
  - Tribe - Maydae
    - Western Hemisphere:
      - Genus - *Zea*
        - A. Subgenus - *Luxuriantes*
          1. *Zea luxurians* (2n = 20)
          2. *Zea perennis* (2n = 40)
          3. *Zea diploperennis* (2n = 20)
        - B. Subgenus - *Zea*
          1. *Zea mays* (2n = 20)
             Subspecies
             1. *Z. mays parviglumis* (2n = 20)
             2. *Z. mays huehuetenangensis* (2n = 20)
             3. *Z. mays mexicana* (Schrad.) (2n = 20)

The genera included in the tribe Maydae include *Zea* and *Tripsacum* in the Western Hemisphere, and *Coix*, *Polytoca*, *Chionachne*, *Schlerachne*, and *Trilobachne* in Asia. Although some researchers have implicated the Asian genera in the origin of maize, the evidence for them is not as extensive and convincing as for the genera located in the Western Hemisphere.

The genus *Zea* includes two sub-genera: *Luxuriantes* and *Zea*. Maize (*Zea mays* L.) is a separate species within the subgenus *Zea,* along with three subspecies. All species within the genus *Zea*, except maize, are different species of teosinte. Until recently, the teosinte species were included in the genus *Euchlaena* rather than the genus *Zea*.

Maize is grown in nearly all areas of the world and is the largest cultivated crop in the world followed by wheat (*Triticum* sp.) and rice (*Oryza sativa* L.) in total global metric ton production. From 2009 to 2013, worldwide maize grain production averaged approximately 875 million metric tons (MMT) per year (USDA-FAS, 2013b). During this same period, the top maize grain producers were the United States, China, Brazil, and the European Union (EU), accounting for 71% of average annual global maize production (USDA-FAS, 2013a). Also during this period, maize production trended upwards from 825 MMT in 2009 to over 960 MMT in 2013 (USDA-FAS, 2013b).

In industrialized countries maize has two major uses: (1) as animal feed in the form of grain, forage or silage; and (2) as a raw material for wet- or dry-milled processed products such as high fructose maize syrup, oil, starch, glucose, dextrose and ethanol. By-products of the wet- and dry- mill processes are also used as animal feed. These processed products are used as ingredients in many industrial applications and in human food products. Most maize produced in industrialized countries is used as animal feed or for industrial purposes, but

maize remains an important food staple in many developing regions, especially sub-Saharan Africa and Central America, where it is frequently the mainstay of human diets (Morris, 1998).

Maize is a very familiar plant that has been rigorously studied due to its use as a staple food/feed and the economic opportunity it brings to growers. The domestication of maize likely occurred in southern Mexico between 7,000 and 10,000 years ago (Goodman, 1988). While the putative progenitor species of maize have not been recovered, it is likely that teosinte played an important role in contributing to the genetic background of maize. Although grown extensively throughout the world, maize is not considered a persistent weed or a plant that is difficult to control. Maize, as we know it today, cannot survive in the wild because the female inflorescence (the ear) is covered by a husk thereby restricting seed dispersal, it has no seed dormancy, and is a poor competitor in an unmanaged ecosystem. The transformation from a wild, weedy species to one dependent on humans for its survival most likely evolved over a long period of time through plant breeding by the indigenous inhabitants of the Western Hemisphere. Today, virtually all the maize grown in the U.S. is a hybrid, a production practice that started in the 1930's (Wych, 1988). Maize hybrids are developed and used based on the positive yield increases and plant vigor associated with heterosis, also known as hybrid vigor.

Conventional plant breeding results in desirable characteristics in a plant through the unique combination of genes already present in the plant. However, there is a limit to genetic diversity with conventional plant breeding. Biotechnology, as an additional tool to conventional breeding, offers access to greater genetic diversity than conventional breeding alone, resulting in expression of highly desirable traits that are profitable to growers.

A2(b)(ii) How the organism is propagated for food use

Maize is wind-pollinated, and the distances that viable pollen can travel depend on prevailing wind patterns, humidity, and temperature. Pollen is shed from the tassel and is viable for approximately 10 to 30 minutes as it is rapidly desiccated in the air (Kiesselbach, 1980). Maize plants shed pollen for up to 14 days.

The reproductive phase begins when one or two auxiliary buds, present in the leaf axils, develop and form the pistillate inflorescence of female flower. The auxiliary bud starts the transformation to form a long 'cob' on which the flowers will be borne. From each flower a style begins to elongate towards the tip of the cob in preparation for fertilization. These styles form long threads, known as silks. The base of the silk is unique, as it elongates continuously until fertilization occurs. Styles may reach a length of 30 centimetres, the longest known in the plant kingdom. Individual maize kernels, or fruit, are unique in that mature seed is not covered by floral bracts (glumes, lemmas, and paleas) as in most other grasses, but rather the entire structure is enclosed and protected by large modified leaf bracts, collectively referred to as the ear. The mature female flowers will remain ready for fertilization for up to two weeks, at which point if fertilization has not occurred, the nucleus will de-organize and fertilization will no longer be possible.

The pollen of maize, a protandrous plant, matures before the female flower is receptive. This may have been an ancient mechanism to ensure cross-pollination, but is no longer considered conducive to modern agricultural practices. However, decades of conventional selection and improvement have produced many maize varieties with similar maturities for both male and female flowers, to ensure seed set for agricultural purposes.

Under natural conditions, maize reproduces only by seed production. Pollination occurs with the transfer of pollen from the tassels to the silks of the ear. About 95% of the ovules are cross-pollinated and about 5% are self-pollinated, although plants are completely self-compatible. Maize, as a thoroughly domesticated plant, has lost all ability to disseminate its seeds and relies entirely on the aid of man for its distribution**.**

A2(b)(iii) What part of the organism is used for food

Maize has been a staple of the human diet for centuries, and its processed fractions are consumed in a multitude of food and animal feed products. In the U.S. the demand for maize is driven by the demand for feed and fuel. In 2013, animal feed accounted for almost 40% of maize consumption, 30% for ethanol, approximately 11% for food and industrial uses, 8% was used as dried distillers grain with the remainder being exported (NCGA, 2014). Recent increases in meat consumption in emerging economic countries, particularly China, coupled with biofuels production has increased global demand for maize (Edgerton, 2009).

Food uses include sweet corn, popcorn, and processed field maize, which are all varieties/hybrids of *Zea mays* subsp. *mays*. The majority of maize used for food and industrial purposes is processed by wet milling to produce starch and sweetener products (e.g., high fructose corn syrup) for use in foodstuffs. Non-food products such as industrial starches, maize gluten feed and maize gluten meal are also manufactured through the wet mill process (May, 1987; Watson, 1988). The primary products derived from the dry milling process are corn meal, corn flour, and ethanol.

Due to its high starch and low fiber contents, maize is considered a valuable energy source in animal feed for livestock such as cattle, pigs and poultry. Whole maize is usually ground and mixed with a high-protein feed compound and with vitamin and mineral supplements to balance the ratio according to the nutritional requirements of the animals being fed (Leath and Hill, 1987). Maize is also used for processing and the production of derivatives, which have a wide range of food, feed and industrial applications. Some of the processed fractions are used for animal feed, such as maize gluten, a resource that is rich in maize protein. Ethanol production from the dry mill process provides dried distiller's grain solubles (DDGS) which are another source of animal feed (RFA, 2010).

A2(b)(iv) Whether special processing is required to render food safe to eat

Maize grain contains 82% endosperm, 12% germ, 5% bran, and 1% tip cap. In addition, 2.2% of the bran fraction is made up of crude fiber (Earle and Curtis, 1946; Perry, 1988). Food uses include sweet corn, popcorn, and processed field maize, which are all varieties/hybrids of *Zea mays* subsp. *mays*.

Maize processing methods include wet milling, dry milling, and fermentation. The milling process separates the maize kernel into three basic parts; endosperm, pericarp, and the germ(Watson, 1988).

Products from wet milling: The majority of the maize used for food and industrial purposes is processed by wet milling to produce starch and sweetener products for use in foodstuffs. Starch is used as a food ingredient in: dairy and ice cream; batters and breading; baked goods; soups, sauces and gravies; salad dressings; meat and poultry; confections; and, in drinks. Starch can also be converted to a variety of sweetener and fermentation products including high fructose maize syrup and ethanol (Watson, 1988).

Products from dry milling: The primary food products derived from the dry milling process are maize grits, maize meal, and maize flours. Maize grits are derived from endosperm of the maize kernel, with less than 1 % oil content. Maize grits are consumed in the U.S. as side dish for breakfast. Maize meal, however, has larger particles than maize grits and is often enriched with thiamine, riboflavin, niacin, and iron to produce baked products such as maize bread and muffins. Maize flour consists of fine endosperm particles, and is often used as a binder in processed meats, as well as in producing several snack foods (Rooney and Serna-Saldivar, 1987).

Products from fermentation: Starch produced from the wet milling process can also be used in producing ethanol and distilled beverages through fermentation (Rooney and Serna-Saldivar, 1987).

A2(b)(v) The significance to the diet in Australia and New Zealand of the host organism

Maize has been a staple of the human diet for centuries, and its processed fractions are consumed in a multitude of food. Estimates of maize consumption are available from the WHO Global Environmental Monitoring System - Food Contamination Monitoring and Assessment Programme (GEMS/Food) (www.who.int/foodsafety/chem/gems). The GEMS/Food programme has developed 13 Cluster Diets which are considered to be representative of the major food consumption patterns exhibited by regional and cultural groups around the world. Australia is included in Cluster M, along with United States and Canada and several other countries.

A3 The Nature of the Genetic Modification

A3(a) Method used to transform host organism

MON 87403 was developed through *Agrobacterium*-mediated transformation of immature maize embryos based on the method described by Sidorov and Duncan (2009), utilizing PV-ZMAP5714. Immature embryos were excised from a post-pollinated maize ear of LH244. After co-culturing the excised immature embryos with *Agrobacterium* carrying the plasmid vector, the immature embryos were placed on selection medium containing glyphosate and carbenicillin disodium salt in order to inhibit the growth of untransformed plant cells and excess *Agrobacterium*, respectively. Once transformed callus developed, the

callus was placed on media conducive to shoot and root development. The rooted plants (R0) with normal phenotypic characteristics were selected and transferred to soil for growth and further assessment.

The tandem T-DNA approach employed by PV-ZMAP5714 uses a single right border and a single left border to achieve two separate T-DNA insertions (Huang *et al*., 2004). In this system, separate, unlinked insertions of the *ATHB17*-containing T-DNA as well as the selectable marker, which is located in the plasmid backbone, are initiated from the same set of border sequences. After initial transformant selection, the unlinked, *cp4 epsps*-containing insert is segregated away using conventional breeding (Huang *et al*., 2004). The R_0 plants generated through the transformation process described above had already been exposed to glyphosate in the selection medium and demonstrated glyphosate tolerance. The R_0 plants containing the *cp4 epsps* expression cassette from the plasmid vector backbone, as well as the *ATHB17*-containing T-DNA, were self-pollinated to produce R_1 seed and R_1 plants (Huang *et al*., 2004). Subsequently, R_1 plants that were positive for the *ATHB17*-containing T-DNA and negative for the *cp4 epsps* expression cassette were identified by a polymerase chain reaction (PCR) based analysis (Huang *et al*., 2004). The R_1 plants determined to be homozygous for the *ATHB17*-containing T-DNA were selected for further development and their progenies were subjected to further molecular and phenotypic assessments. As is typical of the event production and selection process intended to yield potential commercial events, hundreds of different transformation events (regenerants) were generated in the laboratory using PV-ZMAP5714. After many months of careful selection and evaluation of these hundreds of events in the laboratory, greenhouse and field, MON 87403 was selected as the lead event based on superior agronomic, phenotypic, and molecular characteristics. Studies on MON 87403 were initiated to further characterize the genetic insertion and the expressed product, and to establish the food, feed, and environmental safety relative to commercial maize. The major steps involved in the development of MON 87403 are depicted in Figure 1.

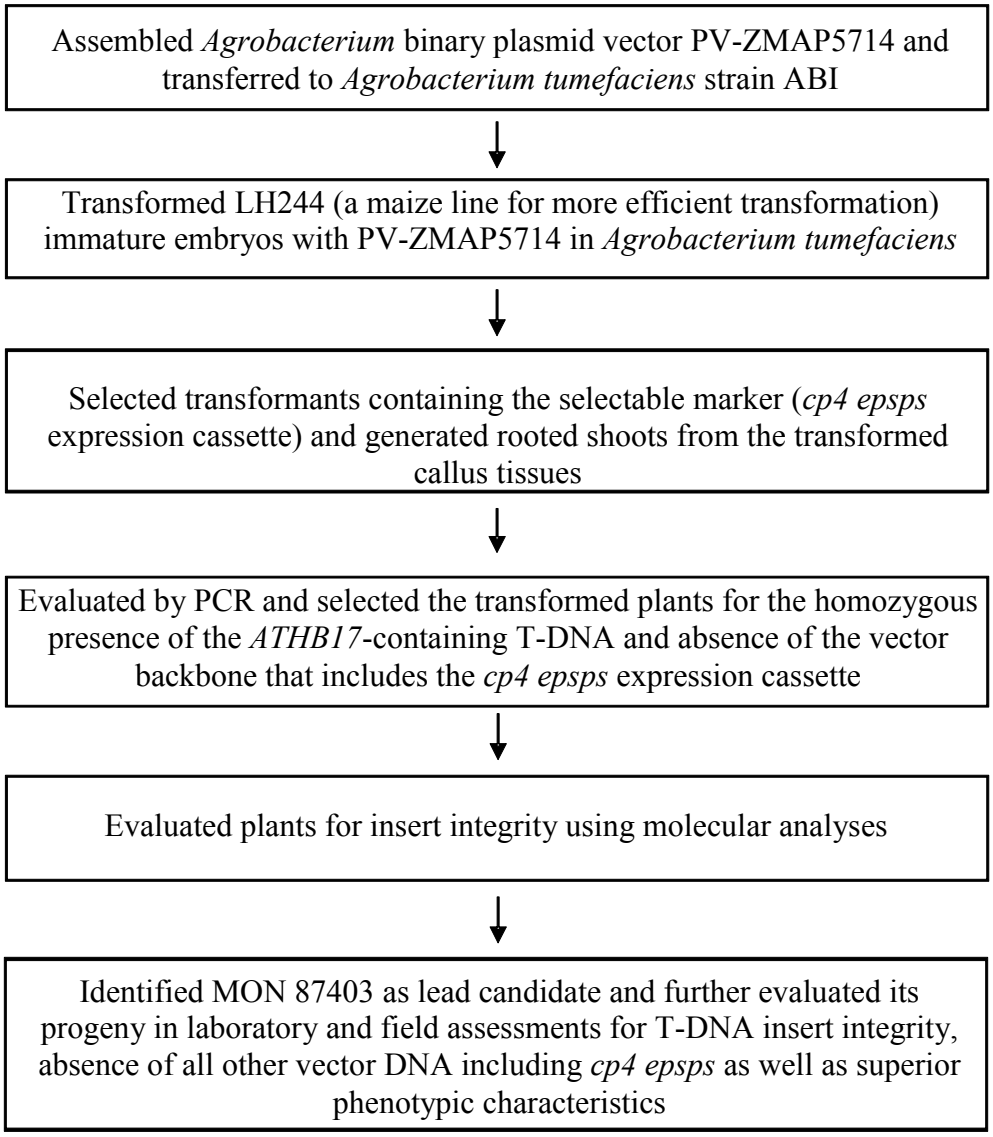


Figure 1. Schematic of the Development of MON 87403

A3(b) Intermediate hosts (*e.g.* bacteria)

A disarmed strain of *Agrobacterium tumefaciens* was the intermediate host used to transfer the plasmid PV-ZMAP5714 into maize cells. PV-ZMAP5714 contains one T-DNA containing the *ATHB17* expression. Following transformation, self-pollination, breeding, and segregation methods were used to produce MON 87403.

A3(c) Gene construct including size, source and function of all elements

A3(c)(i) Gene construct including size, source and function of all elements

MON 87403 was developed through *Agrobacterium tumefaciens*-mediated transformation of maize immature embryos from line LH244 utilizing PV-ZMAP5714.

PV-ZMAP5714

Plasmid vector PV-ZMAP5714 was used for the transformation of maize to produce MON 87403 and its plasmid map is shown in Figure 3. The elements included in this plasmid vector are described in Table 2. PV-ZMAP5714 is approximately 11.7 kb and contains three expression cassettes: one T-DNA, delineated by Left and Right Border regions, contains the *ATHB17* expression cassette and the plasmid backbone contains the *cp4 epsps* selectable marker cassette and the *aadA* expression cassette. PV-ZMAP5714 employs a tandem T-DNA approach to generate marker-free plants (Huang *et al*., 2004). In this tandem T-DNA approach, a single right border and a single left border were used to achieve separate, unlinked insertions of the T-DNA as well as the *cp4 epsps* selectable marker expression cassette. After initial selection of transformants using the *cp4 epsps* marker, conventional breeding and molecular analysis were used to select plants that contained only the *ATHB17* T-DNA and lacked plasmid backbone sequences.

The T-DNA contains the *ATHB17* expression cassette. The *ATHB17* expression cassette is regulated by the *e35S/Ract1* chimeric promoter from the 35S RNA of cauliflower mosaic virus (CaMV) and the *act1* gene from *Oryza sativa* , the 5' untranslated sequence of the *Cab* gene from *Triticum aestivum,* the *act1* intron from *Oryza sativa,* and the 3' untranslated region of *Hsp17* from *Triticum aestivum*. The plasmid backbone contains the *cp4 epsps* and *aadA* expression cassettes. The *cp4 epsps* expression cassette is regulated by the *act1* promoter from *Oryza sativa*, the *act1* intron from *Oryza sativa*, the *CTP2* targeting sequence from *Arabidopsis thaliana*, and the *nos* 3' untranslated region from *Agrobacterium tumefaciens*. The *aadA* expression cassette is regulated by the bacterial promoter and 3' untranslated region of an aminoglycoside-modifying enzyme, 3"(9)-*O*-nucleotidyltransferase from the transposon Tn7. During transformation, both the *ATHB17* T-DNA and the *cp4 epsps* expression cassette were inserted into the maize genome. Subsequently, traditional breeding, segregation, selection and screening were used to isolate those plants that contain the *ATHB17* expression cassette (T-DNA) and do not contain the *cp4 epsps* cassette.

The *ATHB17* Coding Sequence and the ATHB17 Protein (T-DNA)

The *ATHB17* coding sequence is from *Arabidopsis thaliana* that encodes the ATHB17 protein (Figure 2). For additional information about the truncated ATHB17 protein produced in MON 87403 see Section B1(a).

The *cp4 epsps* Coding Sequence and the CP4 EPSPS Protein (backbone)

The *cp4 epsps* expression cassette that is not present in MON 87403 encodes a 47.6 kDa CP4 EPSPS protein consisting of a single polypeptide of 455 amino acids (Padgette *et al*., 1996). The *cp4 epsps* coding sequence is from the *aroA* gene of *Agrobacterium* sp. strain CP4 encoding the CP4 EPSPS protein (Barry *et al*., 2001; Padgette *et al*., 1996). The CP4 EPSPS protein is similar and functionally identical to endogenous plant EPSPS enzymes, but has a much reduced affinity for glyphosate, the active ingredient in Roundup agricultural herbicides, relative to endogenous plant EPSPS (Barry *et al*., 2001; Padgette *et al*., 1996). The presence of this protein renders the plant tolerant to glyphosate and was used to select transformed tissues on glyphosate-containing tissue culture medium.

```
  1 MIKLLFTYIC TYTYKLYALY HMDYACVCMY KYKGIVTLQV CLFYIKLRVF
 51 LSNFTFSSSI LALKNPNNSL IKIMAILPEN SSNLDLTISV PGFSSSPLSD
101 EGSGGGRDQL RLDMNRLPSS EDGDDEEFSH DDGSAPPRKK LRLTREQSRL
151 LEDSFRQNHT LNPKQKEVLA KHLMLRPRQI EVWFQNRRAR SKLKQTEMEC
201 EYLKRWFGSL TEENHRLHRE VEELRAMKVG PTTVNSASSL TMCPRCERVT
251 PAASPSRAVV PVPAKKTFPP QERDR
```

Figure 2. Deduced Amino Acid Sequence of the ATHB17 Protein
The amino acid sequence of the ATHB17 protein was deduced from the full-length coding nucleotide sequence present in PV-ZMAP5714. The amino-terminal 113 amino acids (highlighted in gray) are predicted to not be translated in MON 87403 as a result of mRNA splicing. For additional information about the truncated ATHB17 protein produced in MON 87403, see Section B1(a).

Regulatory Sequences

The *ATHB17* coding sequence in MON 87403 is under the regulation of the *e35S/Ract1* promoter, the *Cab* leader, the *Ract1* intron, and the *Hsp17* 3' untranslated region. The *e35S/Ract1* promoter, which directs transcription in plant cells, is a chimeric promoter consisting of the duplicated enhancer region from the cauliflower mosaic virus 35S RNA promoter (CaMV) (Kay *et al*., 1987) combined with the promoter of the *act1* gene from *Oryza sativa* that encodes Actin 1 (McElroy *et al*., 1990). The *Cab* leader is the 5' untranslated region from the chlorophyll a/b-binding (CAB) protein of *Triticum aestivum* and is involved in regulating gene expression (Lamppa *et al*., 1985). The *Ract1* intron is the intron and flanking untranslated sequence from the *act1* gene from *Oryza sativa* (McElroy *et al*., 1990). The *Hsp17* 3' untranslated region is the 3′ untranslated region from the heat shock protein, Hsp17, of *Triticum aestivum* (McElwain and Spiker, 1989) that directs polyadenylation of the mRNA.

T-DNA Borders

PV-ZMAP5714 contains Left and Right Border regions (Figure 3 and Table 2) that were derived from *A. tumefaciens* plasmids. The border regions each contain a 24-25 bp nick site that is the site of DNA exchange during transformation (Barker *et al*., 1983; Depicker *et al*., 1982; Zambryski *et al*., 1982). The border regions separate the T-DNA from the plasmid backbone region and are involved in the efficient transfer of T-DNA into the maize genome.

Genetic Elements Outside of the T-DNA Borders

Genetic elements that exist outside of the T-DNA border regions are those that are essential for the maintenance or selection of PV-ZMAP5714 in bacteria and allows for selection of transformed plants and are referred to as plasmid backbone. The origin of replication, *ori V*, is required for the maintenance of the plasmid in *Agrobacterium* and is derived from the broad host plasmid *RK2* (Stalker *et al*., 1981). The origin of replication, *ori-pUC*, is required for the maintenance of the plasmid in *E. coli* and is derived from the plasmid vector pUC (Vieira and Messing, 1987). Coding sequence *rop* encodes the repressor of primer

(ROP) protein which is necessary for the maintenance of plasmid vector copy number in *E. coli* (Giza and Huang, 1989). The backbone also contains the *cp4 epsps* expression cassette that codes for the CP4 EPSPS protein (conferring tolerance to glyphosate) that was used as the selectable marker during transformation (Huang *et al*., 2004). The *cp4 epsps* expression cassette is regulated by the *Ract1* promoter, *Ract1* intron, *CTP2* chloroplast targeting sequence, and *nos* 3' untranslated region. The *Ract1* promoter, which directs transcription in plant cells is the promoter and leader of the *act1* gene from *Oryza sativa* encoding the Actin 1 protein (McElroy *et al*., 1990). The *Ract1* intron is the intron and flanking untranslated sequence of the *act1* gene from *Oryza sativa* encoding the Actin 1 protein (McElroy *et al*., 1990) which is involved in regulateing gene expression. The *CTP2* targeting sequence is the targeting sequence of the *ShkG* gene from *Arabidopsis thaliana* encoding the EPSPS transit peptide region that directs transport of the protein to the chloroplast (Herrmann, 1995; Klee *et al*., 1987). The *nos* 3′ untranslated region is the 3' UTR sequence of the *nopaline synthase* (*nos*) gene from *Agrobacterium tumefaciens* pTi encoding NOS (Bevan *et al*., 1983; Fraley *et al*., 1983), that directs polyadenylation of the mRNA. The absence of the backbone and other unintended plasmid sequence in MON 87403 was confirmed by sequencing and bioinformatic analyses (see Section A3(d)(ii)).

Table 2. Summary of Genetic Elements in PV-ZMAP5714

Genetic Element	Location in Plasmid	Function (Reference)
T-DNA		
B-Right Border Region	1-357	DNA region from *Agrobacterium tumefaciens* containing the right border sequence used for transfer of the T-DNA (Depicker *et al*., 1982; Zambryski *et al*., 1982)
Intervening Sequence	358-375	Sequence used in DNA cloning
P-*e35S/Ract1*	376-1556	Chimeric promoter consisting of the duplicated enhancer region from the cauliflower mosaic virus 35S RNA promoter (CaMV) (Kay *et al*., 1987) combined with the promoter of the *act1* gene from *Oryza sativa* that encodes Actin 1 (McElroy *et al*., 1990) that directs transcription in plant cells
Intervening Sequence	1557-1561	Sequence used in DNA cloning
L-*Cab*	1562-1622	5' UTR leader sequence from chlorophyll a/b-binding (CAB) protein of *Triticum aestivum* (wheat) that is involved in regulating gene expression (Lamppa *et al*., 1985)
Intervening Sequence	1623-1638	Sequence used in DNA cloning
I-*Ract1*	1639-2118	Intron and flanking UTR sequence of the *act1* gene from *Oryza sativa* (rice) encoding rice Actin 1 protein. This sequence is involved in regulating gene expression (McElroy *et al*., 1990)
Intervening Sequence	2119-2130	Sequence used in DNA cloning
CS-*ATHB17*	2131-2958	Coding sequence of the *ATHB17* gene from *Arabidopsis thaliana* encoding a member of the class II homeodomain-leucine zipper gene family (HD-Zip II) that is thought to act as a transcription factor (Ariel *et al*., 2007)
Intervening Sequence	2959-2971	Sequence used in DNA cloning
T-*Hsp17*	2972-3181	3' UTR sequence from a heat shock protein, Hsp17, of *Triticum aestivum* (wheat) (McElwain and Spiker, 1989) that directs polyadenylation of the mRNA
Intervening Sequence	3182-3234	Sequence used in DNA cloning
B-Left Border Region	3235-3676	DNA region from *Agrobacterium tumefaciens* containing the left border sequence used for transfer of the T-DNA (Barker *et al*., 1983)

Table 2. Summary of Genetic Elements in PV-ZMAP5714 (continued)

Genetic Element	Location in Plasmid	Function (Reference)
Backbone		
Intervening Sequence	3677-3682	Sequence used in DNA cloning
P-*Ract1*	3683-4603	Promoter and leader of the *act1* gene from *Oryza sativa* (rice) encoding the rice Actin 1 protein (McElroy *et al*., 1990) that directs transcription in plant cells
I-*Ract1*	4604-5081	Intron and flanking UTR sequence of the *act1* gene from *Oryza sativa* (rice) encoding rice Actin 1 protein (McElroy *et al*., 1990). This sequence is involved in regulating gene expression
Intervening Sequence	5082-5090	Sequence used in DNA cloning
TS-*CTP2*	5091-5318	Targeting sequence of the *ShkG* gene from *Arabidopsis thaliana* encoding the EPSPS transit peptide region that directs transport of the protein to the chloroplast (Herrmann, 1995; Klee *et al*., 1987)
CS-*cp4 epsps*	5319-6686	Coding sequence of the *aroA* gene from *Agrobacterium* sp. strain CP4 encoding the CP4 EPSPS protein that provides glyphosate tolerance (Barry *et al*., 2001; Padgette *et al*., 1996)
Intervening Sequence	6687-6701	Sequence used in DNA cloning
T-*nos*	6702-6954	3' UTR sequence of the *nopaline synthase* (*nos*) gene from *Agrobacterium tumefaciens* pTi encoding NOS (Bevan *et al*., 1983; Fraley *et al*., 1983), that directs polyadenylation of the mRNA
Intervening Sequence	6955-7005	Sequence used in DNA cloning
OR-*ori V*	7006-7402	Origin of replication from the broad host range plasmid *RK2*, used for maintenance of plasmid in *Agrobacterium* (Stalker *et al*., 1981)
Intervening Sequence	7403-8910	Sequence used in DNA cloning
CS-*rop*	8911-9102	Coding sequence for repressor of primer protein from the ColE1 plasmid for maintenance of plasmid copy number in *E. coli* (Giza and Huang, 1989)
Intervening Sequence	9103-9529	Sequence used in DNA cloning
OR-*ori-pUC*	9530-10118	Origin of replication from plasmid pUC for maintenance of plasmid in *E. coli* (Vieira and Messing, 1987)

Table 2. Summary of Genetic Elements in PV-ZMAP5714 (continued)

Genetic Element	Location in Plasmid	Function (Reference)
Intervening Sequence	10119-10648	Sequence used in DNA cloning
aadA	10649-11537	Bacterial promoter, coding sequence, and 3' UTR for an aminoglycoside-modifying enzyme, 3"(9)-*O*-nucleotidyltransferase from the transposon Tn7 (Fling *et al*., 1985). This sequence confers spectinomycin and streptomycin resistance
Intervening Sequence	11538-11673	Sequence used in DNA cloning

B, Border
P, Promoter
L, Leader
I, Intron
CS, Coding Sequence
T, Transcription Termination Sequence
TS, Targeting Sequence
OR, Origin of Replication

A3(c)(ii) Detailed map of the location and orientation of all genetic elements

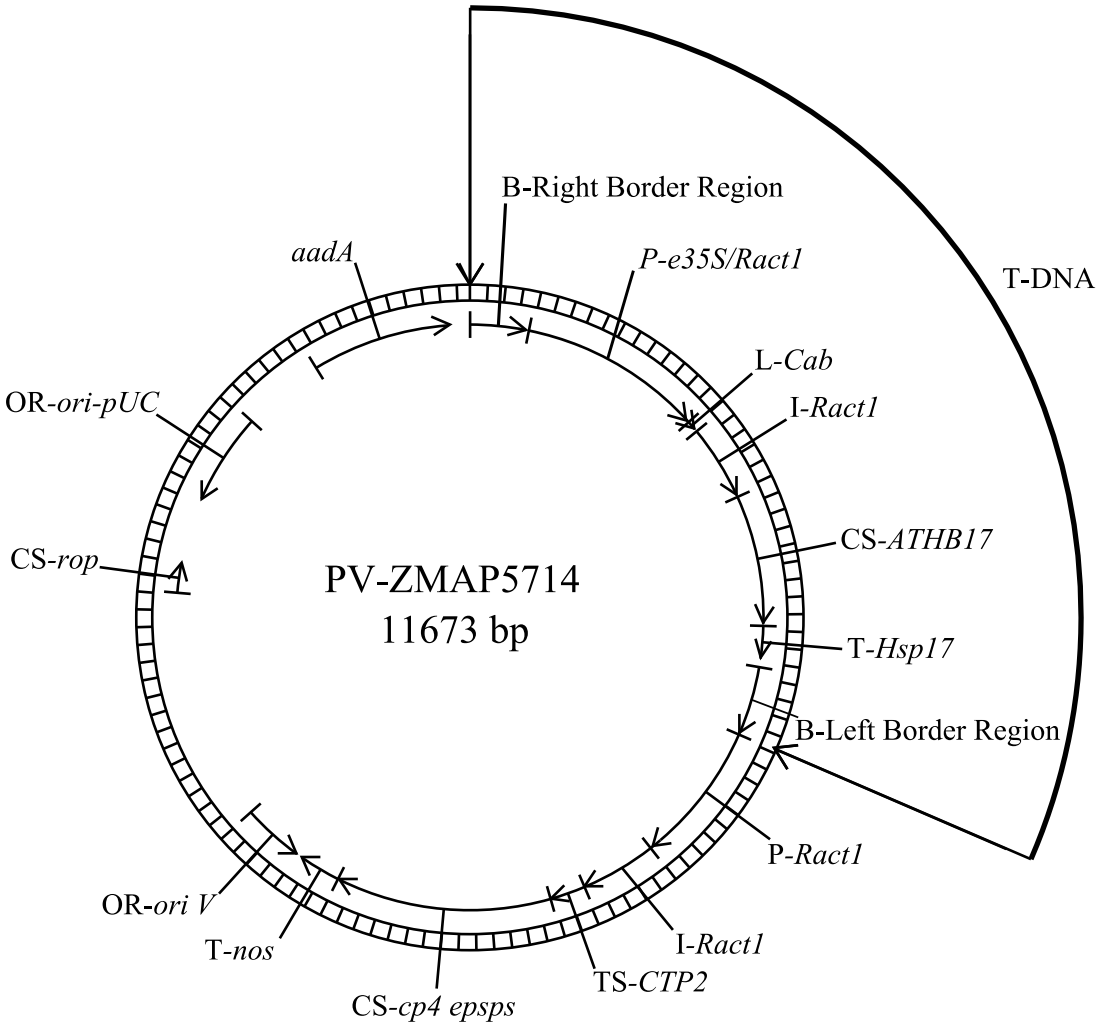


Figure 3. Circular Map of PV-ZMAP5714

PV-ZMAP5714 contains a single T-DNA. Genetic elements are shown on the exterior of the map.

A3(d) Full characterisation of the genetic modification

A3(d)(i) Identification of GM elements

This section contains a comprehensive molecular characterisation of the genetic modification present in MON 87403. It provides information on the DNA insertion(s) into the plant genome of MON 87403, and additional information relative to the arrangement and stability of the introduced genetic material. The information provided in this section addresses the relevant factors in Codex Plant Guidelines, Section 4, paragraphs 30, 31, 32, and 33 (Codex Alimentarius, 2009).

Characterisation of the DNA insert in MON 87403 was conducted using a combination of sequencing, PCR, and bioinformatics. The results of this characterisation demonstrate that MON 87403 contains one copy of the transfer DNA (T-DNA) containing the *ATHB17* expression cassette that is stably integrated at a single locus and is inherited according to Mendelian principles over multiple generations. These conclusions are based on several lines of evidence:

1. Molecular characterisation of MON 87403 by Next Generation Sequencing and Junction Sequence Analysis (NGS/JSA) demonstrated that MON 87403 contained a single intended DNA insert. NGS/JSA provided a comprehensive analysis of MON 87403 to determine the presence and identity of sequences derived from PV-ZMAP5714 (DuBose *et al*., 2013; Kovalic *et al*., 2012). This analysis demonstrated that MON 87403 contained a single T-DNA insert with no detectable backbone sequences.
2. Directed sequencing (locus-specific PCR, DNA sequencing and analyses) performed on MON 87403 was used to determine the complete sequence of the single DNA insert from PV-ZMAP5714, the adjacent flanking DNA, and the 5' and 3' insert-to-flank junctions. This analysis confirmed that the sequence and organization of the inserted DNA is identical to the corresponding region in the PV-ZMAP5714 T-DNA. Furthermore, the genomic organization at the insertion site was assessed by comparing the sequences flanking the T-DNA insert in MON 87403 to the sequence of the insertion site in conventional maize. This analysis determined that no major DNA rearrangement occurred at the insertion site in MON 87403 upon DNA integration.
3. Generational stability analysis by NGS/JSA demonstrated that the single PV-ZMAP5714 T-DNA insert in MON 87403 has been maintained through five breeding generations, thereby confirming the stability of the T-DNA in MON 87403.
4. Segregation analysis corroborates the insert stability demonstrated by NGS/JSA and independently establishes the nature of the T-DNA as a single chromosomal locus.

Taken together, the characterisation of the genetic modification in MON 87403 demonstrates that a single copy of the intended T-DNA was stably integrated at a single locus of the maize genome and that no plasmid backbone sequences are present in MON 87403.

A schematic representation of the NGS/JSA methodology and the basis of the characterisation using NGS/JSA and PCR sequencing are illustrated in Figure 4 below.

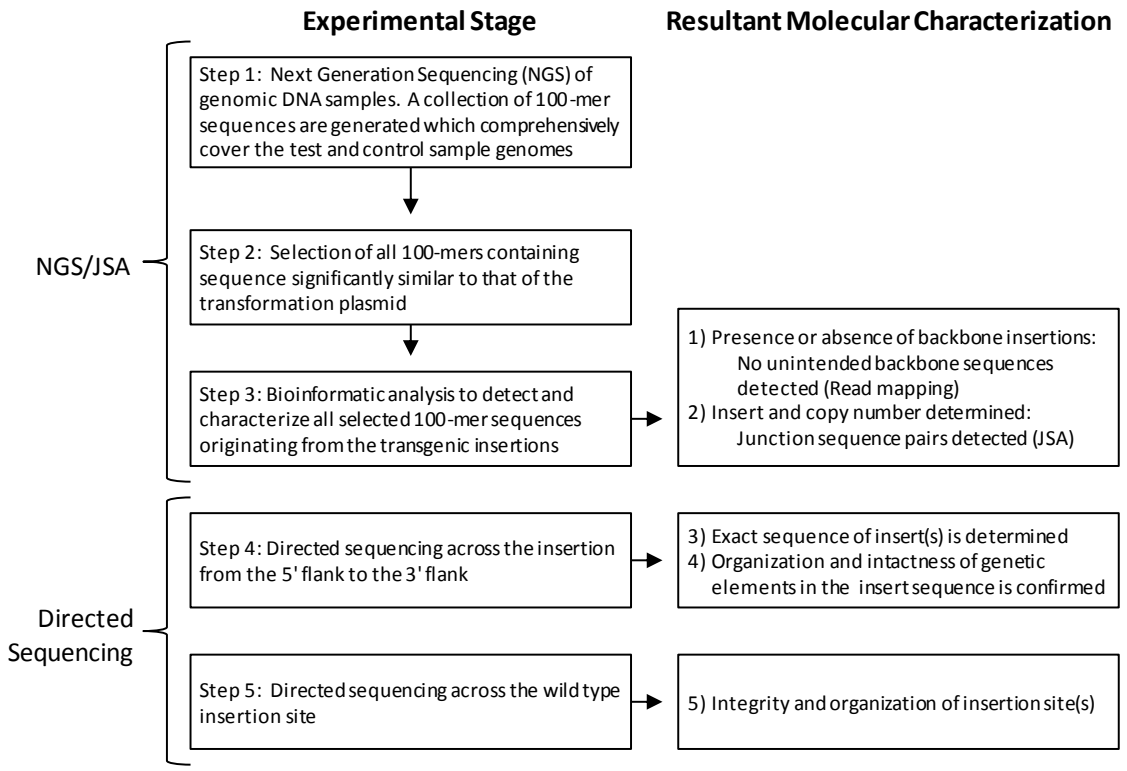


Figure 4. Molecular Characterisation using Sequencing and Bioinformatics

Genomic DNA from the MON 87403 test and the conventional control was sequenced using technology that produces a large set of short, randomly distributed sequence reads (each approximately 100 bp long) that comprehensively cover both genomes (Step 1). Utilising these genomic sequence reads, bioinformatics search tools were used to select all sequence reads that were significantly similar to the transformation plasmid (Step 2) for use in read mapping to determine the presence/absence of backbone sequences and Junction Sequence Analysis (JSA) bioinformatics to determine the insert and copy number (Step 3). Overlapping PCR products are also produced which span any inserts and their wild type locus (Step 4 and Step 5 respectively); these overlapping PCR products are sequenced to allow for detailed characterisation of the inserted DNA and insertion site(s).

The NGS/JSA method characterized the genomic DNA from MON 87403 and the conventional control using short (~100 bp) randomly distributed sequence fragments (sequencing reads) generated in sufficient number to ensure comprehensive coverage of the sample genomes. It has previously been demonstrated that 75× coverage of the soybean genome is adequate to provide comprehensive coverage and ensure detection of inserted DNA and similarly 75× coverage provides comprehensive coverage of the maize genome (Kovalic *et al*., 2012). To confirm sufficient sequence coverage of the genome, the 100-mer sequence reads are analyzed to determine the coverage of a known single-copy endogenous gene, this demonstrates the depth of coverage (the median number of times each base of the genome is independently sequenced). The level of sensitivity of this method was demonstrated by detection of a positive control spiked at 1 and $1/10^{th}$ copy-per-genome equivalent, thus confirming the method's ability to detect any sequences derived from the transformation plasmid. Bioinformatics analysis was then used to select sequencing reads that contained sequences similar to the transformation plasmid, and these were analysed in depth to determine the number and the identity of sequence in the DNA insert(s). NGS/JSA was run on all five generations of MON 87403 samples and the conventional controls. Results of NGS/JSA are shown in Sections A3(d) and A3(f).

The number of DNA inserts was determined by analyzing sequence reads for novel junctions, while the identity of the inserted DNA and absence of backbone sequence was assessed by sequence read mapping. The junctions of the DNA insert and the flanking DNA are unique for each insertion (Kovalic *et al*., 2012). An example is shown in Figure 5. Therefore, insertion sites can be recognized by analyzing for sequence reads containing such junctions.

Each insertion will produce two unique junction sequence classes characteristic of the genomic locus, one at the 5' end of the insert (Figure 6, named junction sequence class A, or JSC-A, in this case) and one at the 3' end of the insert (junction sequence class B, JSC-B) (Kovalic *et al*., 2012).

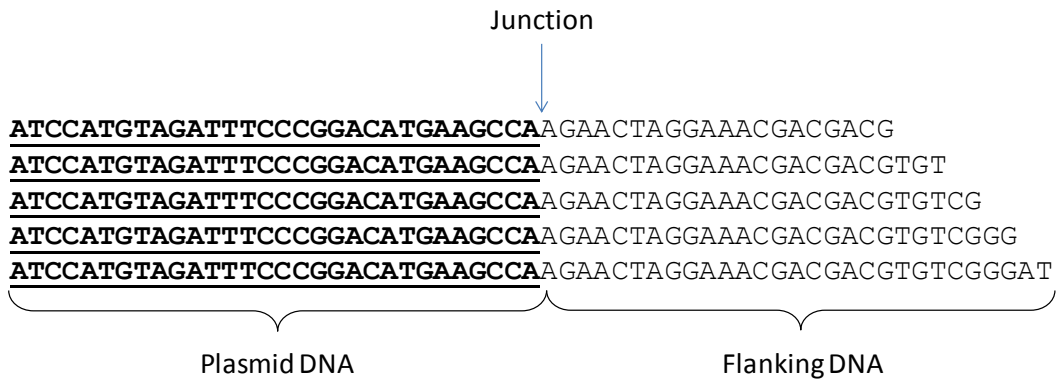


Figure 5. Junctions and Junction Sequences

Depicted above are five example junction sequences formatted and labeled to indicate the plasmid/flanking DNA portions of the sequences and with the junction point indicated (plasmid DNA is shown in bold, underlined text and flank DNA is shown in plain text). Junctions are detected by examining the NGS data for sequences having portions of plasmid sequences that span less than the full read. A group of junction sequences which share the same junction point and common flanking sequence is called a Junction Sequence Class (or JSC).

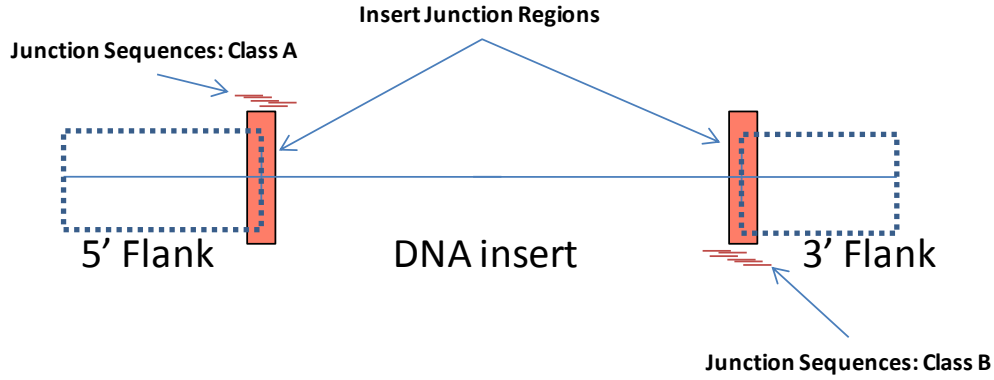


Figure 6. Two Unique Junction Sequence Classes are Produced by the Insertion of a Single Plasmid Region

A schematic representation of a single DNA insertion within the genome showing the inserted DNA, the 5' and 3' flanks (depicted as areas bounded by dotted lines), and the two distinct regions spanning the junctions between inserted DNA and flanking DNA (shaded boxes). The group of ~100-mer sequences in which each read contains sequences from both the DNA insert and the adjacent flanking DNA at a given junction is called a Junction Sequence Class. In this example, two distinct junction sequence classes (in this case: Class A at the 5' end and Class B at the 3' end) are represented.

By evaluating the number of unique junction classes detected, the number of insertion sites of DNA derived from the plasmid can be determined. For a single insert, two junction sequence classes are expected, each originating from one end of the insert, both containing portions of plasmid DNA insert and flanking sequence.

Directed sequencing (locus-specific PCR and DNA sequencing analyses, Figure 4, Step 5) complements the NGS/JSA analyses. Sequencing of the insert and flanking genomic DNA determined the complete sequence of the insert and flanks. This analysis evaluates if the sequence of the insert is identical to the corresponding sequence from the T-DNA in PV-ZMAP5714, if each genetic element in the insert is intact, if the T-DNA sequence is inserted as a single copy, and further confirms that no vector backbone or other unintended plasmid sequences were present within the T-DNA insert in MON 87403. Results are described in Sections A3(d)(i) and A3(d)(ii).

The stability of the T-DNA present in MON 87403 across multiple generations was evaluated by NGS/JSA analyses as described above. This information was used to determine the number and identity of insertion sites in multiple generations. For a single insert, two junction sequence classes are expected; each one originates from either end of the insert, both containing portions of DNA insert and flanking sequence. All integrated sequences are expected to align to the T-DNA region of the plasmid. Results are described in Section A3(f)(i).

Segregation analysis of the T-DNA was conducted to determine the inheritance and stability of the insert in MON 87403. Segregation analysis corroborates the insert stability demonstrated by NGS/JSA and independently establishes the genetic behavior of the T-DNA. Results and methods are described in Section A3(f)(ii).

A3(d)(ii) Determination of Number and Identity of DNA Inserts in MON 87403

The number of insertion sites of PV-ZMAP5714 DNA in MON 87403 was assessed by performing NGS/JSA on MON 87403 genomic DNA. A plasmid map of PV-ZMAP5714 is shown in Figure 3. Table 3 provides descriptions of the genetic elements present in MON 87403. A schematic representation of the insert and flanking sequences in MON 87403 is shown in Figure 7.

Next Generation Sequencing of MON 87403 and Conventional Control Genomic DNA

Genomic DNA from five generations of MON 87403 (Figure 8) and the appropriate conventional control was isolated from seed and prepared for sequencing according to the manufacturer's protocol (Illumina, TruSeq library protocol). These genomic DNA libraries were used to generate short (~100 bp) randomly distributed sequence fragments (sequencing reads) of the maize genome (see Figure 4, Step 1).

To demonstrate sufficient sequence coverage, the 100-mer sequence reads were analyzed by mapping all reads to a known single-copy endogenous gene (*Zea mays* pyruvate decarboxylase (*pdc3*), GenBank accession version: AF370006.2). The analysis of sequence

coverage plots showed that the depth of coverage (i.e., the median number of times any base of the genome is expected to be independently sequenced) was 75× or greater for the five generations of MON 87403 (R_3, R_4, R_5, R_4F_1, and R_5F_1) and the conventional control. It has previously been demonstrated that 75× coverage of the soybean genome is adequate to provide comprehensive coverage and ensure detection of inserted DNA and similarly 75× coverage provides comprehensive coverage of the maize genome (Kovalic *et al*., 2012)

To demonstrate the method's ability to detect any sequences derived from the PV-ZMAP5714 transformation plasmid, a sample of conventional control DNA spiked with PV-ZMAP5714 DNA at 1 and $1/10^{th}$ genome equivalent was analyzed by NGS and bioinformatics. The level of sensitivity of this method was demonstrated to a level of 1 genome equivalent, 100% nucleotide identity was observed over 100% of PV-ZMAP5714 . This result demonstrates that all nucleotides of the transformation plasmid are observed by the sequencing and bioinformatic assessments performed. Also, observed coverage was adequate at a level $1/10^{th}$ genomic equivalent (98.83% coverage at 99.97% identity for the $1/10^{th}$ genome equivalent spiked control sample,and, hence, a detection level of at most $1/10^{th}$ genome equivalent was achieved for the plasmid DNA sequence assessment.

Characterisation of Insert Number in MON 87403 using Bioinformatic Analysis

The number of insertion sites of DNA from PV-ZMAP5714 in MON 87403 was assessed by performing NGS/JSA on MON 87403 genomic DNA using the R_3 generation (Figure 8).

Table 3. Summary of Genetic Elements in MON 87403

Genetic Element[1]	Location in Sequence[2]	Function (Reference)
5' Flanking DNA	1-1345	DNA sequence flanking the 5' end of the insert
B-Right Border Region[r1]	1346-1369	DNA region from *Agrobacterium tumefaciens* containing the right border sequence used for transfer of the T-DNA (Depicker *et al*., 1982; Zambryski *et al*., 1982)
Intervening Sequence	1370-1387	Sequence used in DNA cloning
P-*e35S/Ract1*	1388-2568	Chimeric promoter consisting of the duplicated enhancer region from the cauliflower mosaic virus 35S RNA promoter (CaMV) (Kay *et al*., 1987) combined with the promoter of the *act1* gene from *Oryza sativa* that encodes Actin 1 (McElroy *et al*., 1990) that directs transcription in plant cells
Intervening Sequence	2569-2573	Sequence used in DNA cloning
L-*Cab*	2574-2634	5' UTR leader sequence from chlorophyll a/b-binding (CAB) protein of *Triticum aestivum* (wheat) that is involved in regulating gene expression (Lamppa *et al*., 1985)
Intervening Sequence	2635-2650	Sequence used in DNA cloning
I-*Ract1*	2651-3130	Intron and flanking UTR sequence of the *act1* gene from *Oryza sativa* (rice) encoding Actin 1 protein. This sequence is involved in regulating gene expression (McElroy *et al*., 1990)

Table 3. Summary of Genetic Elements in MON 87403 (continued)

Genetic Element[1]	Location in Sequence[2]	Function (Reference)
Intervening Sequence	3131-3142	Sequence used in DNA cloning
CS-*ATHB17*	3143-3970	Coding sequence of the *ATHB17* gene from *Arabidopsis thaliana* encoding a member of the class II homeodomain-leucine zipper gene family (HD-Zip II) that is thought to act as a transcription factor (Ariel *et al*., 2007)
Intervening Sequence	3971-3983	Sequence used in DNA cloning
T-*Hsp17*	3984-4193	3' UTR sequence from a heat shock protein, Hsp17, of *Triticum aestivum* (wheat) (McElwain and Spiker, 1989) that directs polyadenylation of the mRNA
Intervening Sequence	4194-4246	Sequence used in DNA cloning
B-Left Border Region[r1]	4247-4477	DNA region from *Agrobacterium tumefaciens* containing the left border sequence used for transfer of the T-DNA (Barker *et al*., 1983)
3' Flanking DNA	4478-5744	DNA sequence flanking the 3' end of the insert

[1]Although flanking sequences and intervening sequences are not functional genetic elements; they comprise a portion of the sequence.
[2]Numbering refers to the sequence of the insert in MON 87403 and adjacent DNA.
[r1]Superscript in Left and Right Border Regions indicate that the sequence in MON 87403 was truncated compared to the sequences in PV-ZMAP5714.

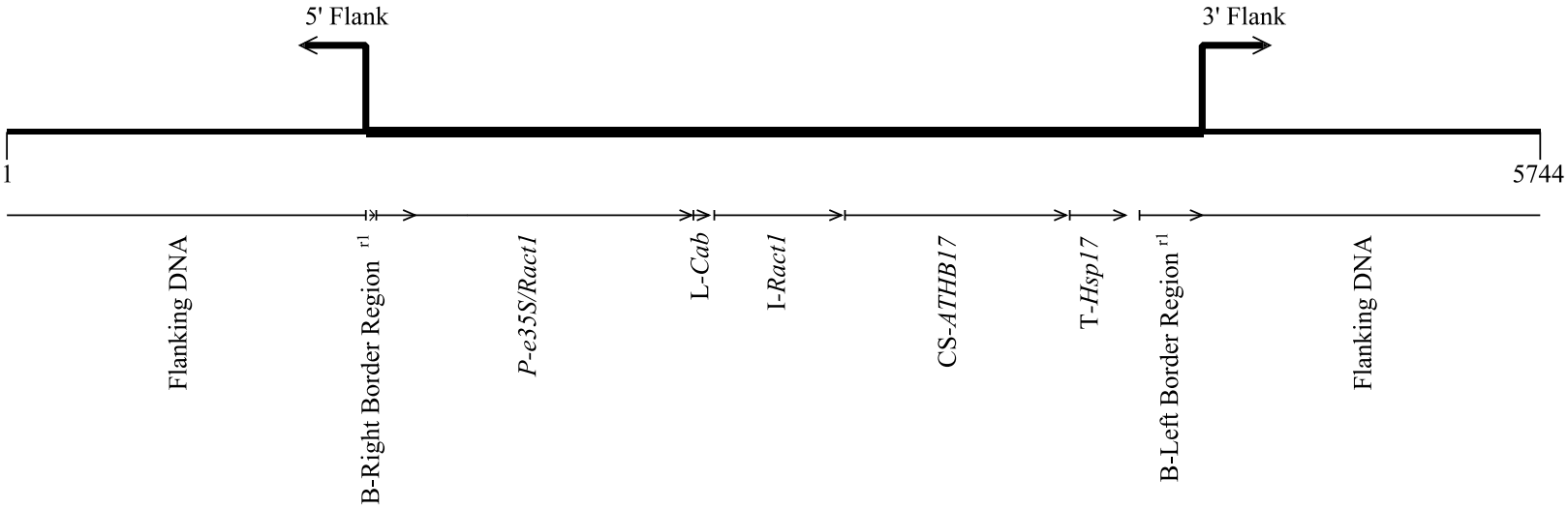


Figure 7. Schematic Representation of the Insert and Flanking Sequences in MON 87403

DNA derived from T-DNA of PV-ZMAP5714 integrated in MON 87403. Right-angled arrows indicate the ends of the integrated T-DNA and the beginning of the flanking sequence. Genetic elements within the insert are identified on the map. This schematic diagram is not drawn to scale, the exact coordinates of every element is shown in Table 3.

[r1]Superscript in Left and Right Border Regions indicate that the sequence in MON 87403 was truncated compared to the sequences in PV-ZMAP5714.

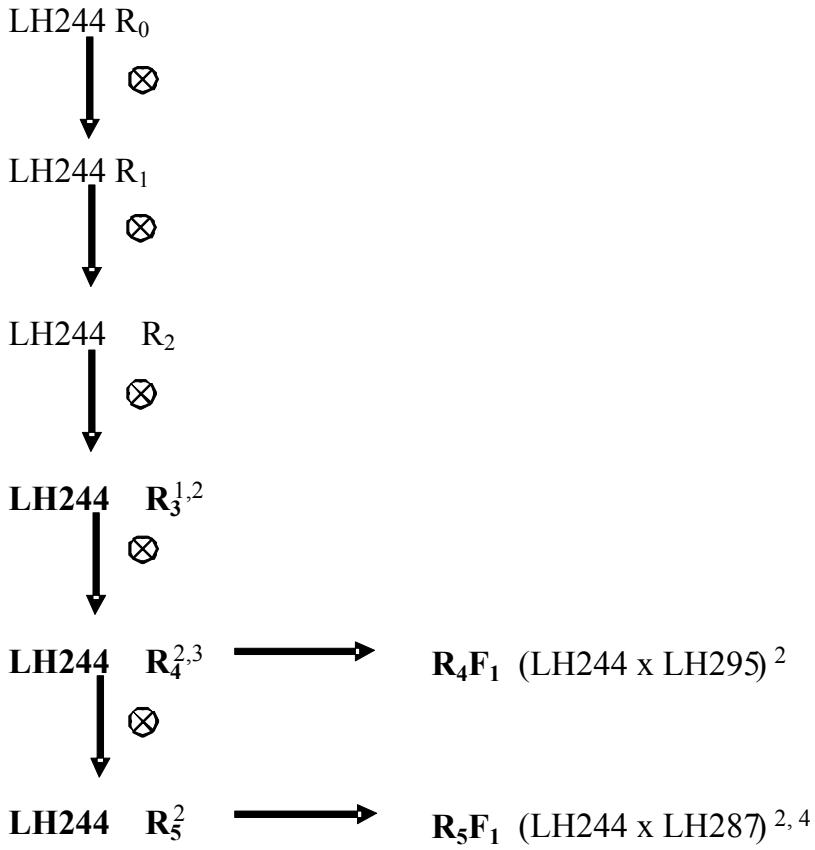


Figure 8. Breeding History of MON 87403

R_0 corresponds to the transformed plant, F# is the filial generation, ⊗ designates self-pollination.

[1] Generation used for molecular characterisation

[2] Generations used to confirm insert stability

[3] Generation used for commercial development of MON 87403

[4] Generation used for compositional analysis studies

Selection of Sequence Reads Containing Sequence of the PV-ZMAP5714

PV-ZMAP5714 was transformed into the parental variety LH244 to produce MON 87403. Consequently, any DNA inserted into MON 87403 will consist of sequences that are similar to the PV-ZMAP5714 DNA sequence. Therefore, to fully characterize the DNA from PV-ZMAP5714 inserted in MON 87403, it is sufficient to completely analyze only the sequence reads that have similarity to the transformation plasmid (Figure 4, Step 2).

Using established criteria, sequence reads similar to the transformation plasmid were selected from MON 87403 and the conventional control sequence datasets and were then used as input data for bioinformatic junction sequence analysis.

Determination of the Insert Number and Copy Number

The NGS/JSA method described above used the entire plasmid sequence as a query to determine the DNA insertion site number. Any inserted transformation plasmid sequence, regardless of origin, either T-DNA or backbone, can be identified by aligning reads to the transformation plasmid sequence while the number of inserted DNA molecules can be determined using JSA. Therefore unlike the traditional Southern blot analysis that separately hybridizes T-DNA or backbone probes, in NGS/JSA the determination of the T-DNA insert number and of the absence of backbone or unintended sequences are simply represented by the identification of sequence reads that match the entire transformation plasmid, the determination of the overall insert number in the genome followed by determination of the exact identity of any DNA insert using directed sequencing and sequence analysis.

By evaluating the number of unique junction classes, the number of DNA insertion sites and copy number can be determined (Figure 4, Step 3). If MON 87403 contains a single T-DNA insert, two junction sequence classes (JSCs), each containing portions of T-DNA sequence and flanking sequence, will be detected.

To determine the insert number in MON 87403, the selected sequence reads described above were analyzed using JSA (Kovalic *et al*., 2012). JSA uses bioinformatic analysis to find and classify partially matched reads characteristic of the ends of insertions. The number of resultant unique JSCs were determined by this analysis and are shown in Table 4.

Table 4. Unique Junction Sequence Class Results

Sample	Junction Sequence Classes Detected
MON 87403	2
LH244	0

The location and orientation of the junction sequences relative to the T-DNA insert determined for MON 87403 are illustrated in Figure 9. As shown in the figure, there are two junction

sequence classes identified in MON 87403. Junction Sequence Class A and Class B (JSC-A and JSC-B) both contain the T-DNA border sequence joined to genomic flanking sequence, indicating that they represent the sequences at the junctions of the intended T-DNA insert and genomic flanking sequence.

The presence of two, and only two, junction sequence classes (joining T-DNA border and flanking sequences) indicate this single pair of JSCs likely arises from the insertion of the single copy of PV-ZMAP5714 T-DNA containing the ATHB17 cassette at a single locus in the genome of MON 87403. JSC-A represents the junction of the T-DNA Left Border sequence to the 5' flank and JSC-B represents the junction of the T-DNA Right Border sequence to the 3' flank. Complete alignment of the JSCs to the full flank/insert sequence confirms that both of these JSCs originate from the same locus of the MON 87403 genome and are linked by contiguous, known and expected DNA that makes up the single insert.

Determination of the Inserted DNA Identity

To determine the identity of inserted DNA, all selected sequences were mapped to the transformation plasmid sequence. While thousands of sequence reads from the R_3 generation mapped to the plasmid T-DNA sequence, only 4 reads mapped to the plasmid backbone. The apparent backbone reads are associated with *rop* and *ori* sequences. These sequences are common in almost all plasmids used in molecular biology. Because we did not see the consistent presence of these 4 sequence reads across multiple generations of MON 87403, we have concluded that these reads likely resulted from contaminants in our sequencing experiments. Such contaminants could come from surfaces in the laboratory environment, bacteria in our tissue samples, or even reagents used in the study, as they are sometimes produced from recombinant bacteria. Other researchers have also noted the infrequent presence of these types of reads (Yang *et al*., 2013). Additionally, these reads are not part of a junction sequence and at a sequencing depth of 1, they are not statistically significant. As such, we conclude that MON 87403 does not contain any sequence from the transformation plasmid backbone.

Based on this comprehensive NGS/JSA study, it is concluded that MON 87403 contains one *ATHB17*-containing T-DNA inserted into a single locus, as shown in Figure 9. The identity of the DNA insert was determined by the sequencing and analysis of overlapping PCR products from this locus as described in this Section A3(d)(ii), which showed that the DNA insert only contained the *ATHB17*-containing T-DNA elements from the plasmid and the backbone sequences were not present in MON 87403.

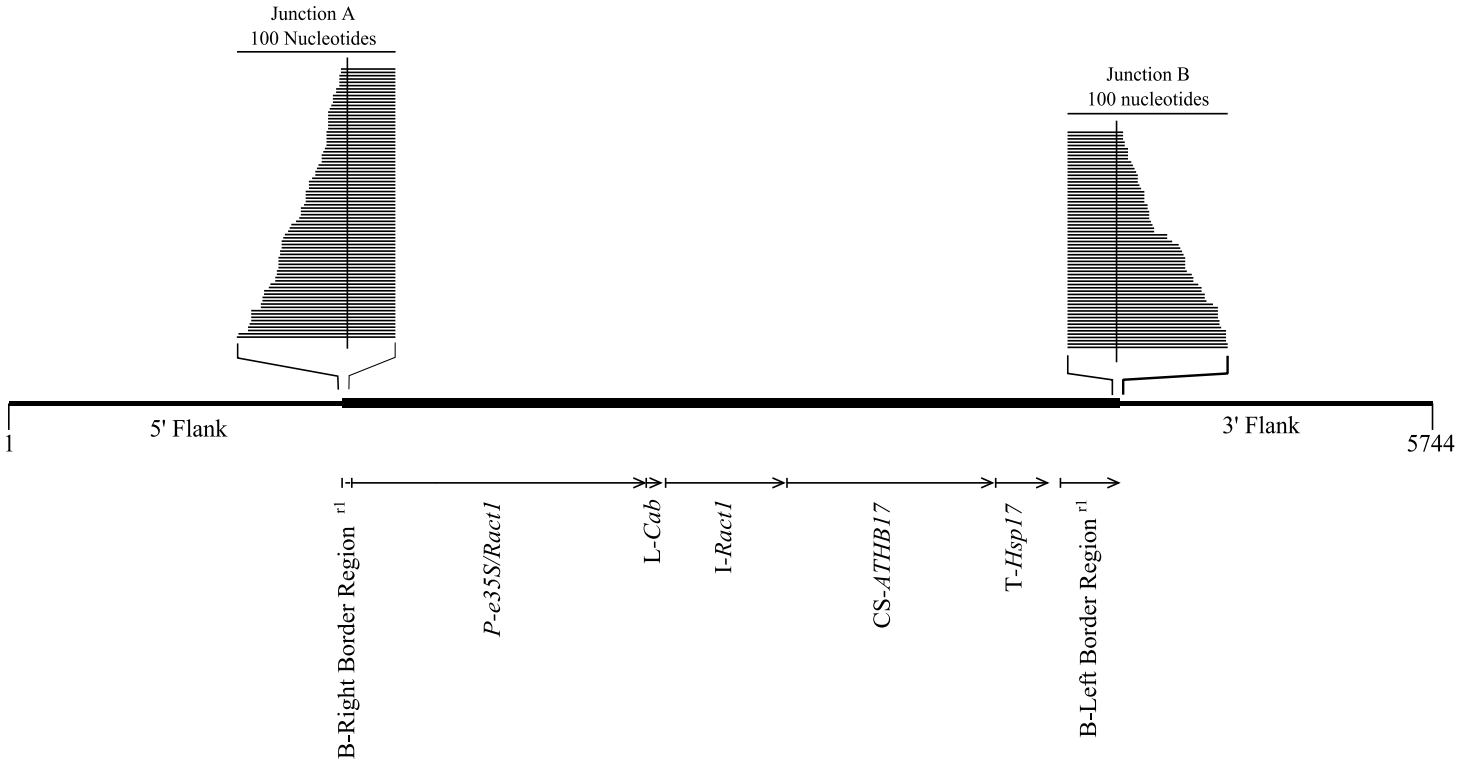


Figure 9. Junction Sequences Detected by NGS/JSA

Linear map of MON 87403 illustrating the relationship of the detected junction sequences to the insert locus. The individual junction sequences detected by JSA are illustrated as stacked bars and arbitrarily cut off in the Right or Left Border Region for representation purposes.

[r1]Superscript in Left and Right Border Regions indicate that the sequence in MON 87403 was truncated compared to the sequences in PV-ZMAP5714.

A3(d)(iii) Full DNA sequence, including junction regions

Organisation and Sequence of the Insert and Adjacent DNA in MON 87403

The organization of the elements within the DNA insert and the adjacent genomic DNA was assessed using directed DNA sequence analysis (refer to Figure 4, Step 4). PCR primers were designed to amplify three overlapping regions of the MON 87403 genomic DNA that span the entire length of the insert (Figure 10). The amplified PCR products were subjected to DNA sequencing analyses. The results of this analysis confirm that the MON 87403 insert is 3,132 bp and that each genetic element within the *ATHB17*-containing T-DNA is intact compared to the transformation plasmid PV-ZMAP5714, with the exception of the border regions. The border regions both contain small terminal deletions with the remainder of the inserted border regions being identical to the sequence in PV-ZMAP5714. The sequence and organization of the insert was also shown to be identical to the corresponding *ATHB17*-containing T-DNA of PV-ZMAP5714, confirming that a single copy of the T-DNA was inserted as intended. This analysis also shows that only *ATHB17*-containing T-DNA elements (described in Table 3) were present in the single DNA insert.

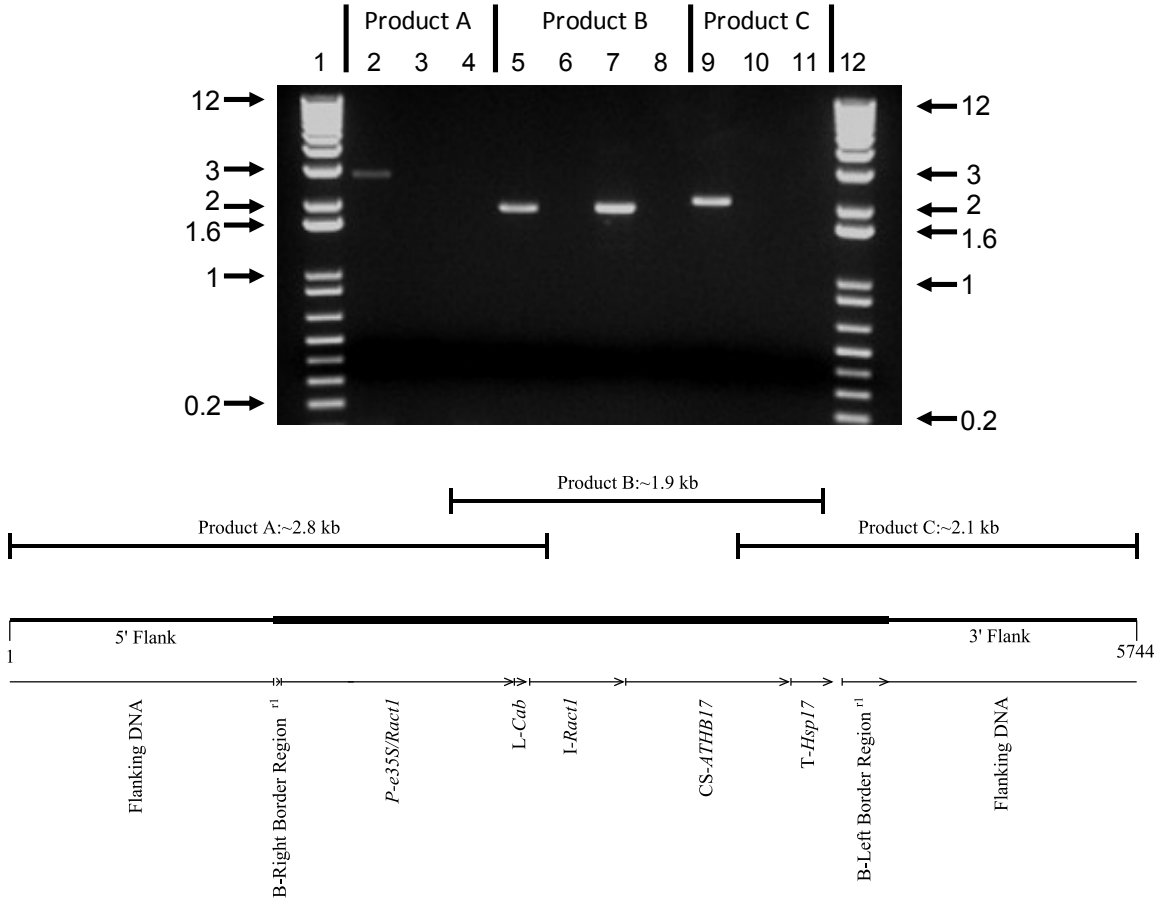


Figure 10. Overlapping PCR Analysis Across the Insert in MON 87403

PCR was performed on both conventional control genomic DNA and MON 87403 genomic DNA using three pairs of primers to generate overlapping PCR fragments from MON 87403 for sequencing analysis. To verify the PCR products, a portion of each PCR was loaded on a 1% (w/v) agarose gel and visualized by ethidium bromide staining. The expected product size for each amplicon is provided in the illustration. Lane designations are as follows:

Lane		**Lane**	
1	1 Kb Plus DNA Ladder	7	PV-ZMAP5714
2	MON 87403	8	Conventional Control
3	No Template Control	9	MON 87403
4	Conventional Control	10	No Template Control
5	MON 87403	11	Conventional Control
6	No Template Control	12	1 Kb Plus DNA Ladder
Arrows next to the agarose gel photograph denote the size of the DNA, in kilobase pairs, obtained from the 1 Kb DNA Plus Ladder (Invitrogen, Grand Island, NY) on the ethidium bromide stained gel.			

[r1] Superscript in Left and Right Border Regions indicate that the sequence in MON 87403 was truncated compared to the sequences in PV-ZMAP5714.

Sequencing of the MON 87403 Insertion Site

PCR and sequence analysis were performed on genomic DNA extracted from the conventional control to examine the insertion site in conventional maize (refer to Figure 4, Step 5). The PCR was performed with one primer specific to the genomic DNA sequence flanking the 5' end of the MON 87403 insert paired with a second primer specific to the genomic DNA sequence flanking the 3' end of the insert (Figure 11). A sequence comparison between the PCR product generated from the conventional control and the sequence generated from the 5' and 3' flanking sequences of MON 87403 indicates that 149 bases of maize genomic DNA were deleted during integration of the T-DNA. The remainder of the flanks in MON 87403 are identical to the conventional control. Such changes are common during plant transformation and these changes presumably resulted from double-stranded break repair mechanisms in the plant during the *Agrobacterium*-mediated transformation process (Salomon and Puchta, 1998).

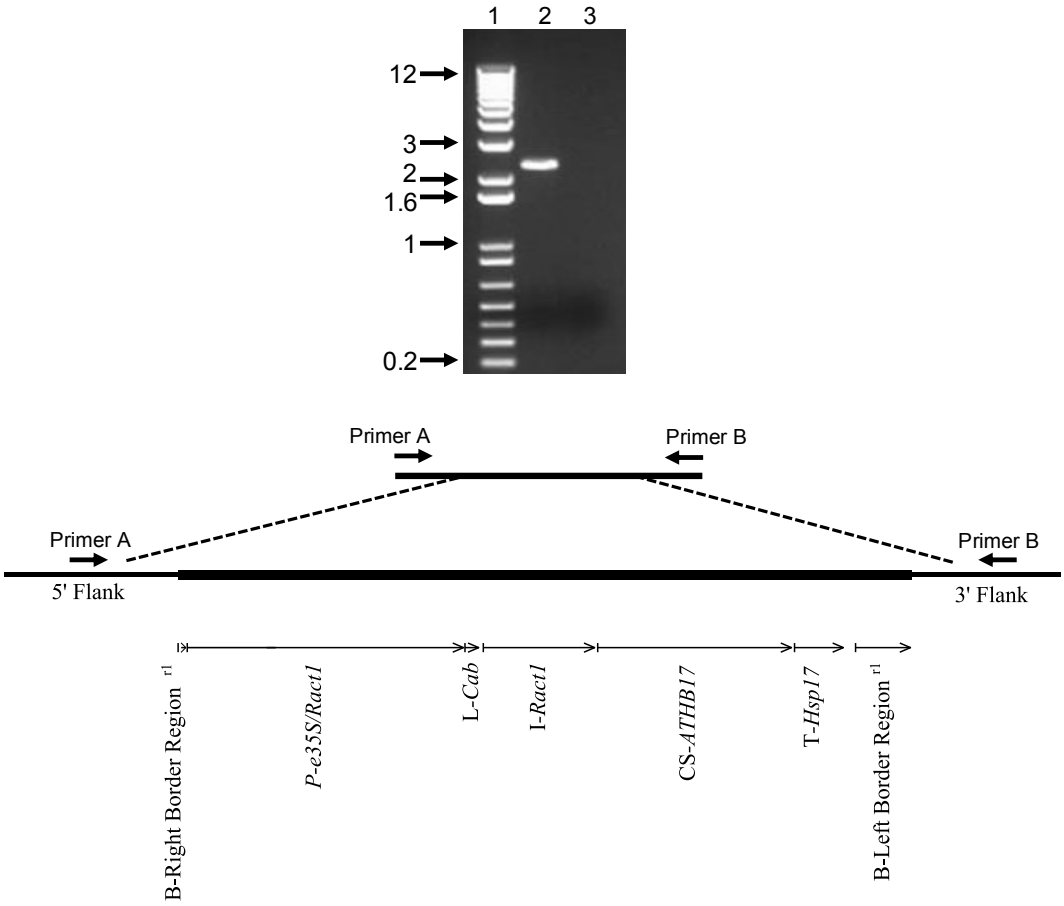


Figure 11. PCR Amplification of the MON 87403 Insertion Site

PCR analysis was performed to evaluate the insertion site. PCR was performed on conventional control DNA using Primer A, specific to the 5' flanking sequence, and Primer B, specific to the 3' flanking sequence of the insert in MON 87403. The DNA generated from the conventional control PCR was used for sequencing analysis. This illustration depicts the MON 87403 insertion site in the conventional control (upper panel) and the MON 87403 insert (lower panel). To verify the PCR products, a portion of each PCR was loaded on the gel and visualized by ethidium bromide staining. Lane designations are as follows:

Lane	
1	1 Kb Plus DNA Ladder
2	Conventional Control
3	No template DNA control

Arrows denote the size of the DNA, in kilobase pairs, obtained from the 1 kb DNA Plus Ladder (Invitrogen, Grand Island, NY) on the ethidium bromide stained gel.

r1Superscript in Left and Right Border Regions indicate that the sequence in MON 87403 was truncated compared to the sequences in PV-ZMAP5714.

For details, please also refer to [redacted]., 2014 (MSL0025909).

A3(d)(iv) Map of the organisation of the inserted DNA (each site)

PCR and DNA sequence analyses performed on MON 87403 and the conventional control determined the organisation of the genetic elements within the insert as given in Figure 10.

A3(d)(v) Identification and characterisation of unexpected ORFs

The 2009 Codex Alimentarius Commission guidelines for the safety assessment of food derived from biotechnology crops (Codex Alimentarius, 2009) includes an assessment element on the identification and evaluation of "open reading frames within the inserted DNA or created by the insertion with contiguous plant genomic DNA." These assessments examine the potential homology of any putative polypeptides or proteins that could be produced from open reading frames (ORFs) in the insert or at the plant-insert junction to known toxins or allergens. These analyses are conducted even if there is no evidence that such alternative reading frames in the insert or such ORFs at the plant-insert junction are capable of being transcribed or translated into a protein. Results from these bioinformatics analyses demonstrate that any putative polypeptides encoded by the MON 87403 event sequence are unlikely to exhibit allergenic, toxic or otherwise biologically adverse properties.

Bioinformatic analyses were performed on the MON 87403 insert and flanking genomic DNA sequences to assess the potential for allergenicity, toxicity, or biological activity of putative peptides/polypeptides encoded by all six reading frames present in the MON 87403 insert DNA as well as ORFs present in the 5' flanking-inserted T-DNA and inserted T-DNA 3' flanking sequence junctions (Figure 12). These various bioinformatic evaluations are depicted in Figure 12. The entire MON 87403 insert DNA sequence was translated in all six reading frames (three forward reading frames and three reading frames in reverse orientation) and the resulting deduced amino acid sequences were then compared to toxin, allergen and all proteins databases using bioinformatic tools. Similarly, ORFs spanning the 5' flanking sequence DNA-inserted DNA junction and inserted DNA-3' flanking sequence DNA junction were translated from stop codon to stop codon in all six reading frames (three forward reading frames and three reading frames in reverse orientation). Putative peptides/polypeptides from each reading frame of eight amino acids or greater were also subjected to bioinformatic analyses. There are no analytical data that indicate any putative peptides/polypeptides subjected to bioinformatic evaluation, other than ATHB17Δ113, are produced. Moreover, the data generated from these analyses confirm that even in the highly unlikely occurrence that a translation product other than ATHB17Δ113 was derived from frames 1 to 6 of the insert DNA, or the ORFs spanning the insert junctions; they do not share a sufficient degree of sequence similarity with other proteins to indicate they would be potentially allergenic, toxic, or have other safety implications.

Bioinformatics Assessment of Insert DNA Reading Frames

Bioinformatic analyses were performed to assess the potential of allergenicity, toxicity, or biological activity of the putative peptides/polypeptides encoded by translation of reading

frames 1 through 6 of the inserted T-DNA sequence present in MON 87403 (Figure 12). The FASTA sequence alignment tool was used to assess structural relatedness between the query sequences and any protein sequences in the AD_2014, TOX_2014, and PRT_2014[2] databases. Structural similarities shared between the sequence with each sequence in the database were examined. The extent of structural relatedness was evaluated by detailed visual inspection of the alignment, the calculated percent identity, and the *E*-score. Alignments having *E*-score less than or equal to 1e-5 (1×10^{-5}) are deemed significant because they may reflect shared structure and function among sequences. In addition to structural similarity, the sequence was screened for short polypeptide matches using a pair-wise comparison algorithm. In these analyses, eight contiguous and identical amino acids were defined as immunologically relevant, where eight represents the typical minimum sequence length likely to represent an immunological epitope (Silvanovich *et al*., 2006) and evaluated against the AD_2014 database.

The results of the FASTA and eight amino acid sliding window search comparisons showed that no relevant structural similarity to known allergens were observed for any of the putative peptide/polypeptides when compared to proteins in the allergen (AD_2014) database. Using the FASTA algorithm to search the AD_2014 database, no alignments with any of the six query sequences generated an *E*-score of less than or equal to 1e-5. Likewise, no alignment met or exceeded the Codex Alimentarius (2009) FASTA alignment threshold for potential allergenicity of 35% identity over 80 amino acids.

Potential toxicity of putative peptide/polypeptide translated from the six reading frames was assessed using the FASTA algorithm. Using the FASTA algorithm to search the TOX_2014 database no alignments with any of the six query sequences generated an *E*-score of less than or equal to 1e-5. As a result, it is unlikely that the translation of any reading frame in the MON 87403 T-DNA would yield a protein displaying toxicity.

When used to query the PRT_2014 database, a comprehensive sequence database containing 32,476,608 sequences, translations of reading frames 1 and 2 yielded alignments with *E*-scores less than or equal to 1e-5. The frame 1 translation positively identified homeobox-leucine zipper protein (GI-330250356) with a significant *E*-score, 2.7e-102, and is consistent with the known insert structure in MON 87403. The frame 2 translation positively identified a partial matrix protein sequence protein from CaMV (Cauliflower Mosaic Virus) (GI-

---

[2] The allergen, gliadin, and glutenin sequence database (AD_2014) was obtained from Food Allergy Research and Resource Program Database (FARRP_2014) (http://www.allergenonline.org) and was used for the evaluation of sequence similarities. It is referred to as the AD_2014 database and contains 1,706 sequences.

The toxin database is a subset of sequences derived from the PRT_2014 database that was selected using a keyword search then filtered to remove likely non-toxin proteins. It is referred to as the TOX_2014 database. It contains 10,419 sequences and was used for the evaluation of sequence similarities.

A comprehensive collection of protein sequences that comprises GenBank release 199 was downloaded from NCBI and was used for the evaluation of sequence similarities. It is referred to as the PRT_2014 database and contains 32,476,608 sequences.

331576) with a significant *E*-score, 7.6e-6, and is also consistent with the known insert structure in MON 87403. This alignment was not unexpected since it was consistent with the DNA source of the aligned region that was from the CaMV 35S enhancer region.

Taken together, these data demonstrate the lack of relevant similarities between known allergens or toxins for putative peptides derived from all six reading frames from the inserted DNA sequence of MON 87403. As a result, in the unlikely event that a translation product other than ATHB17Δ113 was derived from reading frames 1 to 6, these putative polypeptides are not expected to be cross-reactive allergens, toxins, or display adverse biological activity.

Insert Junction Open Reading Frame Bioinformatics Analysis

Analyses of putative polypeptides encoded by DNA spanning the 5' and 3' genomic junctions of the MON 87403 inserted DNA were performed using a bioinformatic comparison strategy (Figure 12). The purpose of the assessment is to evaluate the potential for novel open reading frames (ORFs) to have homology to known allergens, toxins, or proteins that display adverse biological activity. Sequences spanning the 5' flanking sequence DNA and inserted DNA, and the inserted DNA and 3' flanking sequence DNA (Figure 12) were translated from stop codon (TGA, TAG, TAA) to stop codon in all six reading frames. The resulting putative peptide/polypeptides from each reading frame that were eight amino acids or greater in length were compared to AD_2014, TOX_2014, and PRT_2014 databases using FASTA and to the AD_2014 database using an eight amino acid sliding window search. The FASTA sequence alignment tool was used to assess structural relatedness between the query sequences and protein sequences in the AD_2014, TOX_2014, and PRT_2014 databases. Structural similarities shared between each putative polypeptide with each sequence in the database were examined. The extent of structural relatedness was evaluated by detailed visual inspection of the alignment, the calculated percent identity and the alignment length (to ascertain if alignments exceeded Codex Alimentarius (2009) threshold of 35% identity in 80 amino acids for FASTA searches of the AD_2014 database), and the *E*-score. In addition to structural similarity, each putative polypeptide was screened for short polypeptide matches using a pair-wise comparison algorithm. In these analyses, eight contiguous and identical amino acids, as described by Codex (2009), were defined as immunologically relevant, where eight represents the typical minimum sequence length likely to represent an immunological epitope, and evaluated against the AD_2014 database.

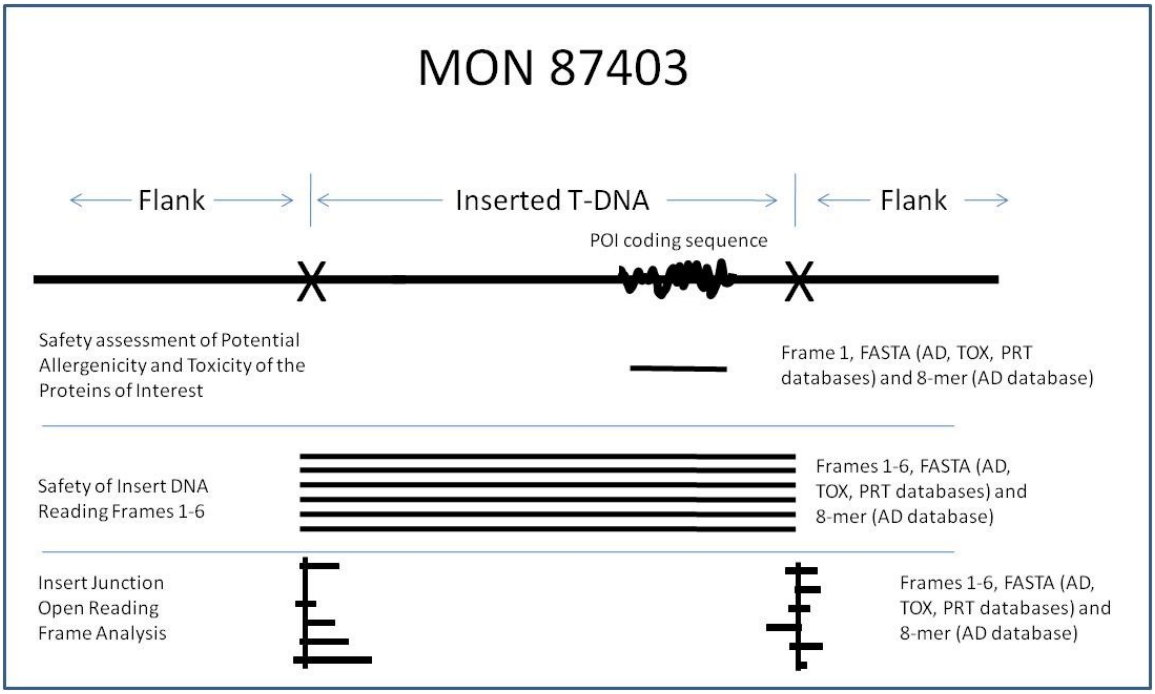
The bioinformatic analysis performed using the 10 putative peptides/popyleptides translated from junctions is theoretical as there is no reason to suspect, or evidence to indicate the presence of transcripts spanning the flank junctions. The results of bioinformatic analysis indicate that no structurally relevant sequence similarities were observed between the 10 putative flank junction derived sequences and allergens or toxins. Alignments yielding significant *E*-scores were observed between sequence 5_1 and matrix protein sequences protein from CaMV such as GI-331576. These alignments are consistent with the known insert structure in MON 87403. As a result, in the unlikely occurrence that any of the 10 putative polypeptides analyzed herein is found *in planta*, none would share significant

similarity or identity to known allergens, toxins, or other biologically active proteins that could affect human or animal health.

Bioinformatic Asessment of Allergenicity, Toxicity, and Adverse Biological Activity Potential of MON 87403 Polypeptides Putatively Encoded by the Insert and Flanking Sequences Summary and Conclusions

A conservative bioinformatic assessment of potential allergenicity, toxicity and adverse biological activity for putative polypeptides that were derived from different reading frames of the entire insert or span the 5' and 3' insert junctions was conducted for MON 87403. There are no analytical data that indicate any putative polypeptides/proteins subjected to bioinformatic evaluation are produced other than ATHB17Δ113. Moreover, the data generated from these analyses confirm that even in the highly unlikely occurrence that a translation product other than ATHB17Δ113 was derived from frames 1 to 6 of the insert DNA, or the ORFs spanning the insert junctions; they would not share a sufficient degree of sequence similarity with other proteins to indicate they would be potentially allergenic, toxic, or have other safety implications. Therefore, there is no evidence for concern regarding the putative polypeptides for MON 87403 relatedness to known toxins, allergens, or biologically active proteins.

For details, please also refer to [redacted], 2014 (MSL0025648), [redacted], 2014 (MSL0025733).



AD= AD_2014; TOX= TOX_2014 and PRT= PRT_2014 (GenBank release 199): 8-mer = the eight amino acid sliding window search.

Figure 12. Schematic Summary of MON 87403 Bioinformatic Analyses

A3(e) Family tree or breeding process

The transformation for MON 87403 was conducted with inbred maize line LH244, a patented maize line assigned to Holden's Foundation Seeds LLC in 2001 (U.S. Patent #6,252,148). LH244 is a medium season yellow dent maize line with a Stiff Stalk background that is best adapted to the central regions of the U.S. corn belt. LH244 was initiated from a single cross of LH197 × LH199 followed by a backcross to LH197. The F_2 combination ((LH197 × LH199) × LH197) was then selfed and used in the development of LH244.

Following transformation of immature LH244 embryos, a single transformed plant was selected and self-crossed to increase seed supplies. A homozygous inbred line was developed through further self-crossing and selection and was then used to produce other inbred and hybrid lines which were used for product testing, safety assessment studies, and commercial hybrid development. For breeding tree please see Figure 8.

Please also refer to section A3(f)(i)

A3(f) Evidence of the stability of the genetic changes

A3(f)(i) Pattern of inheritance of insert and number of generations monitored

Determination of Insert Stability over Multiple Generations of MON 87403

In order to demonstrate the stability of the T-DNA present in MON 87403 through multiple generations, JSA and read mapping was performed using DNA obtained from five breeding generations of MON 87403. The breeding history of MON 87403 is presented in Figure 8, and the specific generations tested are indicated in the figure legend. The MON 87403 R_3 generation was used for the molecular characterisation analyses discussed in Section A3(d)(i) and A3(d)(ii) and shown in Figure 8. To assess stability, four additional generations were evaluated by JSA and read mapping as previously described in Section A3(d)(i) , and compared to the fully characterized R_3 generation. The conventional controls used for the generational stability analysis included LH244, which included similar background genetics to the R_3, R_4 and R_5 generations and represents the original transformation line; LH244 × LH295, a hybrid with similar background genetics to the R_4F_1 hybrid; and LH244 × LH287, a hybrid with similar background genetics to the R_5F_1 hybrid. Genomic DNA isolated from each of the selected generations of MON 87403 and conventional control was used for NGS/JSA analysis.

To determine the insert number in the MON 87403 generations, the sequences selected as described in Section A3(d)(ii) were analyzed using JSA (Kovalic *et al*., 2012). Table 5 shows the number of resultant JSCs containing PV-ZMAP5714 DNA sequence determined by this analysis.

Table 5. Junction Sequence Classes Detected

Sample	Junction Sequence Classes Detected
MON 87403 (R_3)	2
MON 87403 (R_4)	2
MON 87403 (R_4F_1)	2
MON 87403 (R_5)	2
MON 87403 (R_5F_1)	2
LH244	0
LH244 × LH295	0
LH244 × LH287	0

Alignment of the JSCs from each of the assessed MON 87403 generations (R_4, R_5, R_4F_1, and R_5F_1) to the full flank/insert sequence and JSCs determined for the MON 87403 R_3 generation, confirms that the pair of JSCs originates from the same region of the MON 87403 genome and is linked by contiguous, known and expected DNA sequence. This single identical pair of JSCs is observed as a result of the insertion of PV-ZMAP5714 T-DNA at a single locus in the genome of MON 87403. The consistency of these JSC data across all

generations tested demonstrates that this single locus was stably maintained throughout the MON 87403 breeding process. Additionally, read mapping demonstrated a consistent absence of plasmid backbone sequence in all tested generations and the conventional controls.

These results demonstrate that the MON 87403 single integration locus was maintained through several generations of breeding MON 87403; thereby confirming the stability of the insert. Based on this comprehensive sequence data and bioinformatic analysis (NGS/JSA), it is concluded that MON 87403 contains a single and stable T-DNA insertion.

For details, please also refer to [redacted], 2014 (MSL0025909).

Inheritance of the Genetic Insert in MON 87403

The MON 87403 T-DNA resides at a single locus within the maize genome and therefore should be inherited according to Mendelian principles of inheritance. During development of lines containing MON 87403, genotypic segregation data were recorded to assess the inheritance and stability of the MON 87403 T-DNA using Chi-square (χ^2) analysis over several generations. The χ^2 analysis is based on comparing the observed segregation ratio to the expected segregation ratio according to Mendelian principles.

The MON 87403 breeding path for generating segregation data is described in Figure 13. The transformed R_0 plant was self-pollinated to generate R_1 seed. An individual plant homozygous for the MON 87403 T-DNA was identified in the R_1 segregating population via an End-Point TaqMan® PCR assay.

The homozygous positive R_1 plant was self-pollinated to give rise to R_2 seed. The R_2 plants were self-pollinated to produce R_3 seed. The R_3 plants were self-pollinated to produce R_4 seed. R_4 plants homozygous for the MON 87403 T-DNA were crossed via traditional breeding techniques to a Monsanto proprietary conventional recurrent parent that does not contain the *ATHB17* coding sequence to produce hemizygous R_4F_1 seed. The R_4F_1 plants were crossed with the recurrent parent to produce BC_1F_1 seed. The BC_1F_1 generation was tested for the presence of the T-DNA by End-Point TaqMan PCR to select for hemizygous MON 87403 plants. BC_1F_1 plants hemizygous for MON 87403 T-DNA were crossed with the recurrent parent to produce the BC_2F_1 plants. BC_2F_1 plants hemizygous for MON 87403 T-DNA were crossed with the recurrent parent to produce the BC_3F_1 plants.

The inheritance of the MON 87403 T-DNA was assessed in the BC_1F_1, BC_2F_1, and BC_3F_1 generations. At the BC_1F_1, BC_2F_1, and BC_3F_1 generations, the MON 87403 T-DNA was

® TaqMan is a registered trademark of Roche Molecular Systems, Inc.

predicted to segregate at a 1:1 ratio (hemizygous positive: homozygous negative) according to Mendelian inheritance principles.

A Pearson's chi-square (χ^2) analysis was used to compare the observed segregation ratios of the *ATHB17* coding sequence to the expected ratios.

The Chi-square was calculated as:

$\chi^2 = \sum [(| o - e |)^2 / e]$

where o = observed frequency of the genotype or phenotype and e = expected frequency of the genotype or phenotype. The level of statistical significance was predetermined to be 5% ($\alpha = 0.05$).

The results of the χ^2 analysis of the segregating progeny of MON 87403 are presented in Table 6. The χ^2 value in the BC_1F_1, BC_2F_1, and BC_3F_1 generations indicated no statistically significant difference between the observed and expected 1:1 segregation ratio (hemizygous positive: homozygous negative) of MON 87403 T-DNA. These results support the conclusion that the MON 87403 T-DNA resides at a single locus within the maize genome and is inherited according to Mendelian principles of inheritance. These results are also consistent with the molecular characterisation data indicating that MON 87403 contains a single intact copy of the *ATHB17* expression cassette inserted at a single locus in the maize genome.

For details, please also refer to [redacted]., 2013 (MSL0024676).

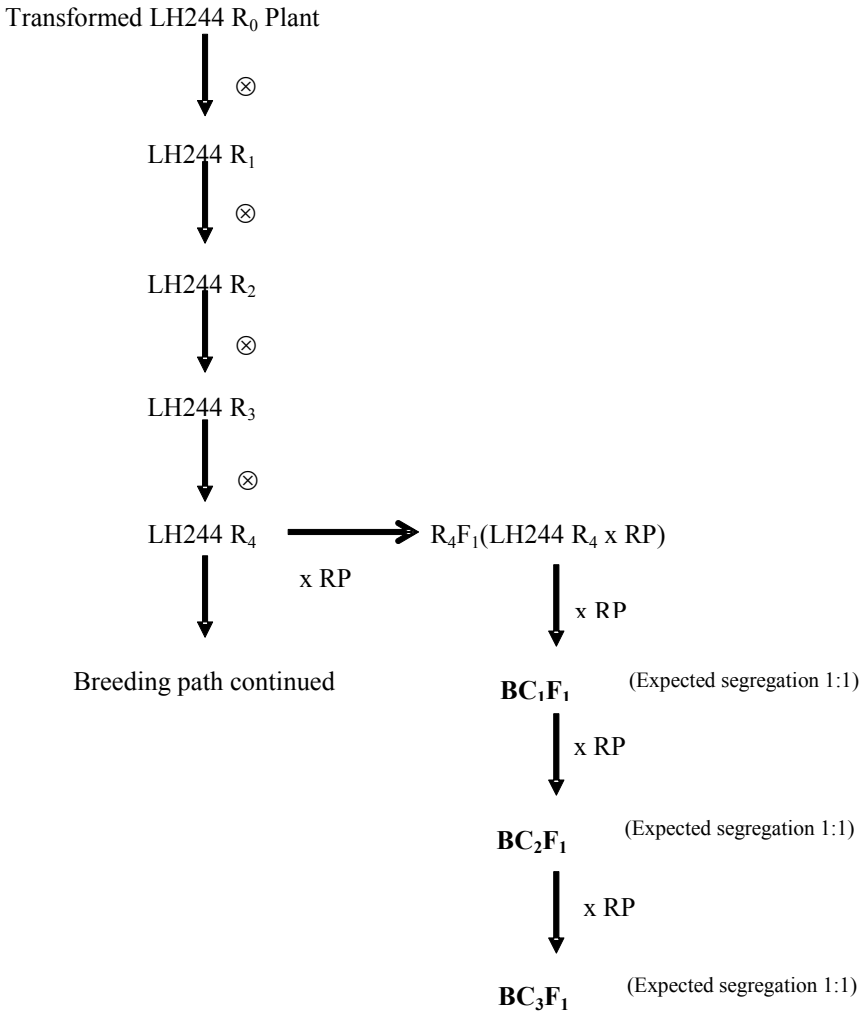


Figure 13. Breeding Path for Generating Segregation Data for MON 87403

Chi-square analysis was conducted on segregation data from BC_1F_1, BC_2F_1, and BC_3F_1 generations (bolded text).

RP: Recurring parent (HCL503).

BC: Back-Cross.

⊗: Self- Pollinated.

Table 6. Segregation of the Expression Cassette During the Development of MON 87403

Generation	Number of plants	Observed Positives	Observed Negatives	Expected Positives	Expected Negatives	χ^2	Probability
BC_1F_1	180	88	92	90.00	90.00	0.09	0.766
BC_2F_1	178	82	96	89.00	89.00	1.10	0.294
BC_3F_1	181	101	80	90.50	90.50	2.44	0.119

A3(f)(ii) Pattern of expression of phenotype over several generations

The stability of the T-DNA present in MON 87403 across multiple generations was evaluated by NGS/JSA analyses as described above. This information was used to determine the number and identity of insertion sites in multiple generations. For a single insert, two junction sequence classes are expected; each one originates from either end of the insert, both containing portions of DNA insert and flanking sequence. All integrated sequences are expected to align to the T-DNA region of the plasmid. In the case of MON 87403, two identical junction sequence classes were detected in all the generations tested, confirming that the single insert is stably inherited over multiple generations.

Segregation analysis of the T-DNA was conducted to determine the inheritance and stability of the insert in MON 87403. The results support the conclusion that the MON 87403 T-DNA resides at a single locus within the maize genome and is inherited according to Mendelian principles of inheritance. Segregation analysis corroborates the insert stability demonstrated by NGS/JSA and independently establishes the genetic behavior of the T-DNA.

For details please refer to Section A3(f)(i).

In order to assess the presence of the ATHB17Δ113 protein in MON 87403 across multiple breeding generations, western blot analysis of MON 87403 was conducted on leaf tissue collected from generations R_3, R_4, R_4F_1, R_5, and R_5F_1 of MON 87403, and on leaf tissue of the conventional control (LH244).

The presence of the ATHB17Δ113 protein was demonstrated in five breeding generations of MON 87403 using western blot analyses (Figure 14). The *E. coli*-produced ATHB17Δ113 protein reference standard (0.1 ng) was used as a reference for the positive identification of the ATHB17Δ113 protein (Figure 14, Lane 2). The presence of the ATHB17Δ113 protein in MON 87403 leaf tissue samples was determined by visual comparison of the bands detected in five breeding generations (Figure 14, Lanes 4-8) to the *E. coli*-produced ATHB17Δ113 protein reference standard (Figure 14, Lane 2). The MON 87403-produced ATHB17Δ113 protein migrated indistinguishably from that of the *E. coli*-produced protein standard analyzed on the same western blot. As expected, the ATHB17Δ113 protein was not detected in the conventional control leaf extract used as the negative control (Figure 14, Lane 9).

Additional bands corresponding to proteins other than the expected ATHB17Δ113 protein were detected in both MON 87403 and conventional leaf samples. The presence in both samples is likely the result of cross-reactivity between the primary or secondary antibodies to proteins endogenous to maize leaf. The non-specific binding of the antibodies resulted in the appearance of bands ranging from approximately 25 to 125 kDa (Figure 14, Lanes 4-9) in both MON 87403 and the conventional control samples. The presence of these non-specific bands does not affect the conclusions of the present study which establishes the generational stability of the ATHB17Δ113 protein in MON 87403.

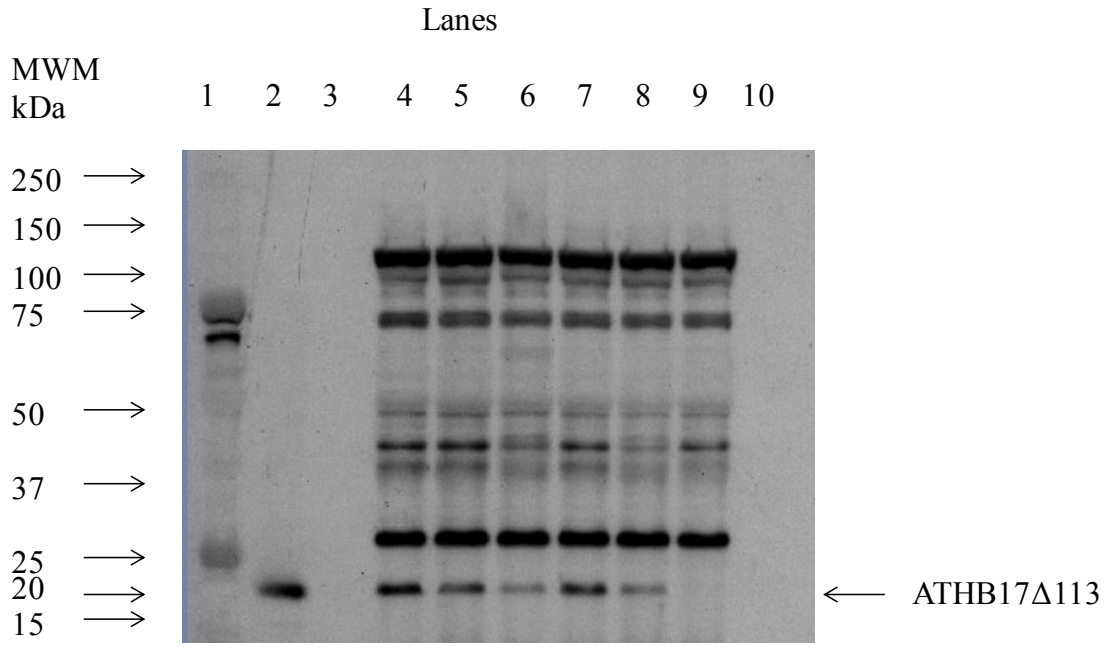


Figure 14. Presence of ATHB17Δ113 Protein in Multiple Generations of MON 87403

Extracts from five generations of MON 87403 leaf tissues, conventional control leaf tissue, *E. coli*-produced ATHB17Δ113 standard, and molecular weight markers were subjected to SDS-PAGE and electrotransferred to a nitrocellulose membrane. The membrane was incubated with rabbit anti-ATHB17Δ113 antibodies and immunoreactive bands visualized through the use of chemiluminescent reagents. The molecular weights (in kDa) of the standards are shown on the left. Lane designations are as follows:

Lane	Description	Amount Loaded on Gel
1	Precision Plus Protein Dual Color Molecular Weight Marker	10 µl
2	*E. coli*-produced ATHB17Δ113 protein (0.1 ng)	10 µl
3	Blank	N/A
4	Test Substance, R_3	10 µl
5	Test Substance, R_4	10 µl
6	Test Substance, R_4F_1	10 µl
7	Test Substance, R_5	10 µl
8	Test Substance, R_5F_1	10 µl
9	Conventional Substance, Control	10 µl
10	Blank	N/A

A4 Analytical Method for Detection

The event-specific DNA-based detection methods such as PCR can be used as the monitoring tool to determine the presence of MON 87403 in a collected sample.

B. INFORMATION RELATED TO THE SAFETY OF THE GM FOOD

B1 Equivalence Studies

B1(a) ATHB17Δ113 protein sequence identity

As described in Section A2(b)(i), the *ATHB17* gene is derived from *Arabidopsis thaliana*. Compared to the wild type sequence of the ATHB17 protein, the *ATHB17*-encoded protein expressed in MON 87403 lacks 113 amino acids due to alternative mRNA splicing, and has thus been designated ATHB17Δ113. The *ATHB17* expression cassette in MON 87403 includes the *ATHB17* coding sequence and the I-*Ract1* intron (Figure 15). Typically, pre-mRNA processing includes mRNA splicing, which consists of the removal of the introns from pre-mRNA transcripts and ligation of exons to produce mature mRNA (Simpson and Filipowicz, 1996). To determine the sequence of the *ATHB17* mature mRNA transcript in MON 87403, total RNA was extracted from leaf tissue and then PolyA$^+$ RNA was enriched. RNA transcripts of *ATHB17* were amplified by RT-PCR and then analyzed by sequencing. In MON 87403, the *ATHB17* mRNA is spliced to remove the majority of the I-*Ract1* intron as well as a portion of the *ATHB17* gene (Figure 15). In order to predict the amino acid sequence of the protein produced from the *ATHB17* mRNA, a Basic Local Alignment Search Tool (BLASTX) search was performed to compare the *ATHB17* mRNA sequence against the GenBank_Protein_Preferred protein sequence database. The top result of this search was the *Arabidopsis thaliana* ATHB17 protein sequence (GenBank ID 179876107). The first methionine (M) amino acid in the BLASTX alignment provides a predicted translation start site for the *ATHB17*-derived protein in MON 87403, and results in the protein sequence of ATHB17Δ113, which is lacking the N-terminal 113 amino acids found in wild type ATHB17 (Figure 16). To confirm the presence of the ATHB17Δ113 protein in MON 87403, western blot analysis was conducted on MON 87403 with an anti-ATHB17Δ113 polyclonal antibody. Western blots revealed the presence of an immunogenic band corresponding to the expected molecular weight of the predicted ATHB17Δ113 protein (Figure 21).

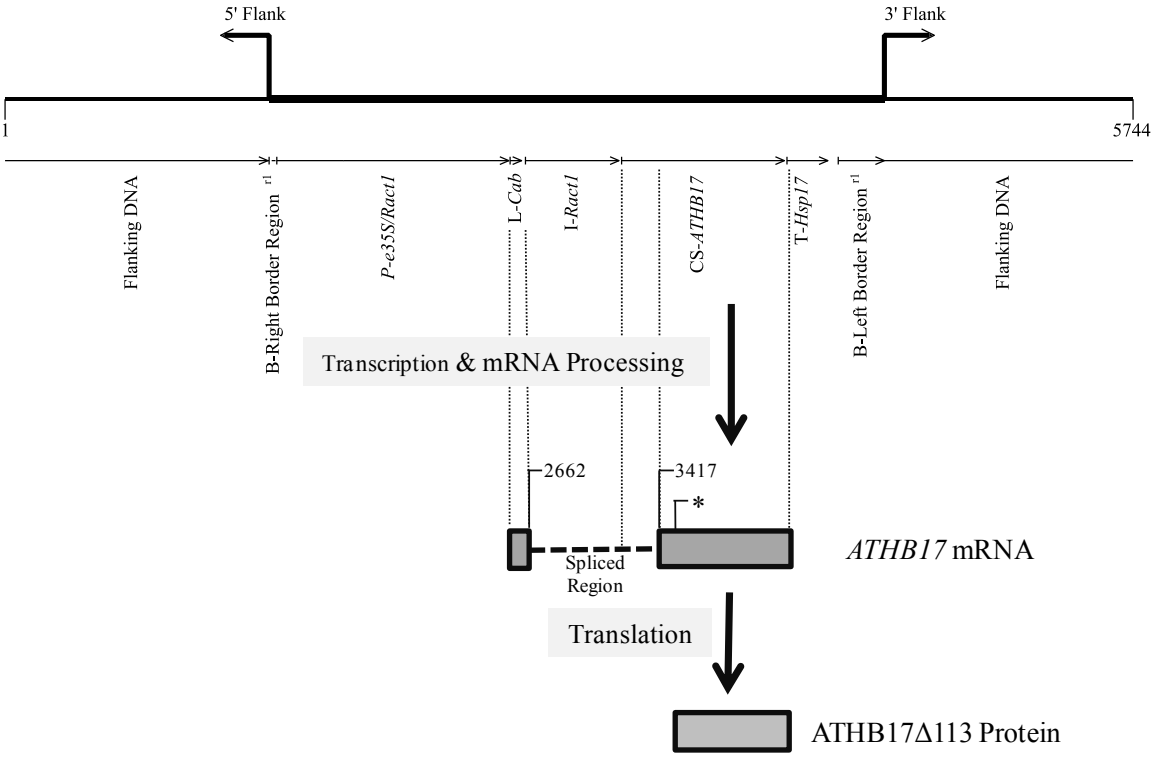


Figure 15. Insert Map of MON 87403 and Schematic Diagram of *ATHB17* Transcription, mRNA Processing, and Predicted Translation

The nucleotide sequence of the mature *ATHB17* mRNA was determined using RT-PCR followed by Sanger sequencing. The mRNA sequence that codes for the ATHB17 protein in MON 87403 was determined. This analysis allowed for the identification of the splice junction in MON 87403 *ATHB17* mRNA. The maize splicing of MON 87403 *ATHB17* mRNA results in the excision of the I-*Ract1* element (splice junction site 2662) and a portion of the CS-*ATHB17* coding region (splice junction site 3417). The resulting transcript is predicted to produce a version of the ATHB17 protein in MON 87403 without the first 113 amino acids at the N-terminus compared to the sequence of ATHB17 found in Arabidopsis (the translation initiation site is denoted by an *).

```
  1 MNRLPSSEDG DDEEFSHDDG SAPPRKKLRL TREQSRLLED SFRQNHTLNP KQKEVLAKHL
 61 MLRPRQIEVW FQNRRARSKL KQTEMECEYL KRWFGSLTEE NHRLHREVEE LRAIKVGPTT
121 VNSASSLTMC PRCERVTPAA SPSRAVVPVP AKKTFPPQER DR
```

Figure 16. Predicted Amino Acid Sequence of the Protein Produced by MON 87403 Mature ATHB17 mRNA

The consensus sequence of MON 87403 mature *ATHB17* mRNA was used in a BLASTX 2.2.23 search of the GenBank_Protein_Preferred database. The predicted amino acid sequence of MON 87403 ATHB17Δ113, beginning with the first methionine amino acid in the top BLAST result is shown above.

B1(b) ATHB17Δ113 protein identity and equivalence

For safety data generated using *E. coli*-produced protein to be applied to the plant-produced protein, the equivalence of the plant- and *E. coli*-produced proteins must be demonstrated. Due to very low ATHB17Δ113 expression, the MON 87403-produced ATHB17Δ113 protein was not isolated from grain, but instead limited amounts of the protein were enriched in leaf extract. ATHB17Δ113 from MON 87403 has characteristics of intractable proteins, which are proteins with properties that make it extremely difficult or impossible with current methods to express in heterologous systems; isolate, purify, concentrate; quantify (due to low levels); demonstrate biological activity; or prove equivalency with plant proteins (Bushey *et al*., 2014). Some of the same qualities (e.g. low stability and low quantity) that make proteins intractacable are the same qualities that indicate lack of hazards in conventional proteins (Bushey *et al*., 2014). For MON 87403-produced ATHB17Δ113, the key difficulties were its extremely low expression levels and corresponding low purity in the enriched leaf extract. In spite of these challenges, the ATHB17Δ113 protein isolated from leaf tissue of MON 87403 was characterized and equivalence of the physicochemical properties between the MON 87403-produced and *E. coli*-produced ATHB17Δ113 was established using a combination of analytical techniques applied to the enriched leaf extract, including: 1) sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis to establish equivalence of apparent molecular weight between the MON 87403-produced and *E. coli*-produced ATHB17Δ113 proteins; 2) matrix assisted laser desorption/ionization - time of flight mass spectrometry (MALDI-TOF MS) analysis of the putative ATHB17Δ113 band in the enriched leaf extract to establish protein identity, 3) liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis to establish the N-terminus, 4) western blot analysis to establish that MON 87403-produced and *E. coli*-produced ATHB17Δ113 with equivalent electrophoretic mobility are both immunoreactive to an ATHB17Δ113-specific antibody, 5) DNA binding assays to demonstrate that *E. coli*-produced protein is functional and binds to its expected target DNA sequence, 6) LC-MS/MS amino acid sequencing characterisation of MON 87403-produced ATHB17Δ113 protein indicates that it is not glycosylated. Taken together, these data provide a characterisation of the MON 87403-produced ATHB17Δ113 protein and establish its equivalence to the *E. coli*-produced ATHB17Δ113 protein. This demonstration of protein equivalence confirms that the *E. coli* produced protein is appropriate for evaluation of the safety of the MON 87403 produced ATHB17Δ113 protein.

For details, please refer to [redacted], 2014 (MSL0025829), [redacted], 2015 (MSL0026621), [redacted], 2015 (MSL0026645).

B1(b)(i) Molecular weight and purity analysis of the MON 87403 ATHB17Δ113 protein

To assess the molecular weight (MW) and purity of MON 87403-produced ATHB17Δ113 protein in the enriched leaf extract, samples of the enriched extract were subjected to SDS-PAGE. *E. coli*-produced ATHB17Δ113 protein was also loaded, as a reference. Following electrophoresis, the gel was stained with silver stain and analyzed by densitometry.

The apparent MW of the protein band representing MON 87403-produced ATHB17Δ113 protein, which was confirmed by MALDI-TOF MS analysis (see Section B1(b)(ii)) and Western blotting (see Section B1(b)(iv)), was calculated to be 22.4 kDa based on SDS-PAGE mobility (Figure 17). Because the experimentally determined apparent MW of the MON 87403-produced ATHB17Δ113 protein was within the 95% prediction interval experimentally derived from the *E. coli*-produced ATHB17Δ113 protein (Table 7), the MON 87403-produced and *E. coli*-produced ATHB17Δ113 proteins were determined to have equivalent apparent molecular weights.

The purity of the MON 87403-produced ATHB17Δ113 protein was calculated based on the six loads (Figure 17, Lanes 3-8) on the gel. The amount loaded in each lane was calculated based upon the concentration of MON 87403-produced ATHB17Δ113 protein determined by the Protein Concentration Analysis described in Section B1(b)(ii). The mean purity was determined to be 3%.

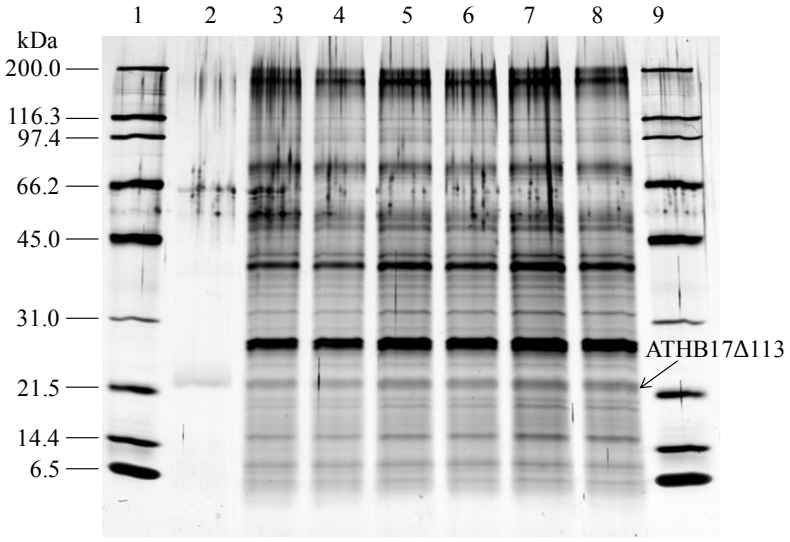


Figure 17. Molecular Weight and Purity Analysis of the MON 87403-produced ATHB17Δ113 Protein

Aliquots of the MON 87403-produced ATHB17Δ113 and the *E. coli*-produced ATHB17Δ113 proteins were analyzed by SDS-PAGE and silver staining. The molecular weights (kDa) of the standards are shown on the left. Lane 10 was cropped from the image. Lane designations are as follows:

Lane	Sample	Amount (ng)
1	Broad Range MW Standards	90
2	*E. coli*-produced ATHB17Δ113	4
3	MON 87403-produced ATHB17Δ113	8
4	MON 87403-produced ATHB17Δ113	8
5	MON 87403-produced ATHB17Δ113	12
6	MON 87403-produced ATHB17Δ113	12
7	MON 87403-produced ATHB17Δ113	16
8	MON 87403-produced ATHB17Δ113	16
9	Broad Range MW Standards	90
10	Blank	-

Table 7. Molecular Weight Comparison Between the MON 87403- and *E. coli*-produced ATHB17Δ113 Proteins Based on SDS-PAGE

Apparent MW of MON 87403-produced ATHB17Δ113 Protein (kDa)	Apparent MW of *E. coli*-produced ATHB17Δ113 Protein[1] (kDa)	Acceptance Limits[2] (kDa)
22.4	22.2	22.2 – 23.5

[1] The molecular weight of the *E. coli*-produced ATHB17Δ113 protein as reported on its Certificate of Analysis.
[2] Data obtained for the *E. coli*-produced protein were used to generate the prediction interval.

B1(b)(ii) Identification of MON 87403-produced ATHB17Δ113 protein by MALDI-TOF Tryptic Mass Map analysis

The identity of the MON 87403-produced ATHB17Δ113 protein was confirmed by MALDI-TOF MS analysis of peptide fragments produced by trypsin digestion of the putative protein band identified as ATHB17Δ113 (Figure 17) after excision from an SDS-PAGE gel of the enriched leaf extract. There were 8 unique peptides identified that corresponded to the expected masses (Table 8). The identified masses were used to assemble a peptide coverage map of the ATHB17Δ113 protein (Figure 18). The experimentally determined coverage of the ATHB17Δ113 protein was 50% (81 out of 162 amino acids) and mapped peptides spanned almost the entire length of the protein (from amino acid position 4 to amino acid position 159 on the coverage map).

Table 8. Summary of the Tryptic Masses Identified for the MON 87403-produced ATHB17Δ113 Protein Using MALDI-TOF MS

Experimental Mass[1]	Calculated Mass[2]	Diff.[3]	Fragment[4]	Sequence[5]
873.46	873.43	0.03	154-160	TFPPQER
878.47	878.45	0.02	37-43	LLEDSFR
1001.56	1001.53	0.03	153-160	KTFPPQER
1218.65	1218.61	0.04	66-74	QIEVWFQNR
1374.67	1374.63	0.04	93-103	WFGS…ENHR
1378.68	1378.68	0.00	33-43	EQSR…DSFR
2224.17	2223.99	0.18	116-135	VGPT…RCER
2358.02	2357.95	0.07	4-25	LPSS…APPR

[1] Only experimental masses that matched calculated masses are listed in the table.
[2] The calculated mass is the relative molecular mass calculated from the matched peptide sequence.
[3] The calculated difference between the experimental mass and the calculated mass.
[4] Fragment numbering is based on the predicted N-terminus of the ATHB17Δ113 protein.
[5] For peptide matches greater than nine amino acids in length the first 4 residues and last 4 residues are show separated by dots (…).

```
001 MNRLPSSEDG DDEEFSHDDG SAPPRKKLRL TREQSRLLED SFRQNHTLNP
051 KQKEVLAKHL MLRPRQIEVW FQNRRARSKL KQTEMECEYL KRWFGSLTEE
101 NHRLHREVEE LRAIKVGPTT VNSASSLTMC PRCERVTPAA SPSRAVVPVP
151 AKKTFPPQER DR
```

Figure 18. MALDI-TOF MS Coverage Map of the MON 87403-produced ATHB17Δ113 Protein

The amino acid sequence of the MON 87403-produced ATHB17Δ113 protein was deduced from the *ATHB17* gene present in MON 87403 and RT-PCR analysis of MON 87403 *ATHB17* RNA. Boxed regions correspond to peptides that were identified from the MON 87403-produced ATHB17Δ113 protein sample using MALDI-TOF MS. In total, 50% (81 out of 162 amino acids) of the deduced protein sequence spanning from amino acid 4 to amino acid 159 on the coverage map.

Additional characterisation of the MON 87403-produced ATHB17Δ113 protein byMass Spectrometry

The MON 87403-produced ATHB17Δ113 protein in the enriched leaf extract was further characterized by amino acid sequencing using a Thermo Orbitrap Fusion LC-MS/MS instrument. The identity of the MON 87403-produced ATHB17Δ113 protein was confirmed by LC-MS/MS sequencing analysis of peptide fragments produced by trypsin digestion of the enriched MON 87403-produced ATHB17Δ113 protein sample.

The MS/MS dataset contained 49 unique peptides that corresponded to the expected peptides for the trypsin digested ATHB17Δ113 protein (Table 9). The characterized peptides were used to assemble a coverage map of the ATHB17Δ113 protein (Figure 19). The experimentally determined coverage of the ATHB17Δ113 protein was 99% (160 out of 162 amino acids).

In addition, the alignment of the experimentally determined N-terminal sequence of the MON 87403-produced ATHB17Δ113 protein with the expected sequence and the sequence determined for the *E. coli*-produced ATHB17Δ113 protein demonstrates that both proteins have the same expected N-terminal sequence (Figure 20). Therefore, sequencing by LC-MS/MS confirms the identity of the ATHB17Δ113 protein enriched from the leaf of MON 87403 and establishes that the predicted N-terminus is intact.

Table 9. Summary of the Tryptic Peptides Identified for the MON 87403-produced ATHB17Δ113 Protein[1]

Experimental Mass (Da)[2]	Calculated Mass (Da)[3]	Diff. (Da)[4]	Fragment[5]	Sequence[6]	Modification
2759.1384	2759.1365	0.0019	1-25	MNRL…APPR	
2801.1498	2801.1471	0.0027	1-25	MNRL…APPR	Acetyl (Protein N-term)
2817.1448	2817.142	0.0028	1-25	MNRL…APPR	Acetyl (Protein N-term); Oxidation (M)
2775.1353	2775.1314	0.0038	1-25	MNRL…APPR	Oxidation (M)
2929.2461	2929.242	0.004	1-26	MNRL…PPRK	Acetyl (Protein N-term)
2945.239	2945.237	0.002	1-26	MNRL…PPRK	Acetyl (Protein N-term); Oxidation (M)
2903.2292	2903.2264	0.0028	1-26	MNRL…PPRK	Oxidation (M)
2628.0977	2628.096	0.0017	2-25	NRLP…APPR	
2670.1089	2670.1066	0.0023	2-25	NRLP…APPR	Acetyl (Protein N-term)
785.5232	785.5235	-0.0004	27-32	KLRLTR	
1157.6624	1157.6629	-0.0005	28-36	LRLT…EQSR	
1748.9185	1748.9169	0.0016	30-43	LTRE…DSFR	
1378.6834	1378.6841	-0.0006	33-43	EQSR…DSFR	
2311.1654	2311.1669	-0.0014	33-51	EQSR….LNPK	
878.4502	878.4498	0.0005	37-43	LLEDSFR	
1810.9337	1810.9326	0.0012	37-51	LLED…LNPK	
937.5283	937.528	0.0004	59-65	HLMLRPR	Oxidation (M)
2138.1325	2138.132	0.0005	59-74	HLML…FQNR	Oxidation (M)
1218.6148	1218.6145	0.0003	66-74	QIEV…FQNR	
1802.8441	1802.843	0.0011	78-91	SKLK…EYLK	Oxidation (M)
1786.8488	1786.8481	0.0007	78-91	SKLK…EYLK	
1587.7157	1587.716	-0.0003	80-91	LKQT…EYLK	Oxidation (M)
1743.8169	1743.8171	-0.0002	80-92	LKQT…YLKR	Oxidation (M)
1502.6355	1502.6381	-0.0026	82-92	QTEM…YLKR	Oxidation (M)
1530.7332	1530.7328	0.0005	92-103	RWFG…ENHR	
2536.2585	2536.2571	0.0014	93-112	WFGS…EELR	
1179.6362	1179.636	0.0002	104-112	LHRE…EELR	
1491.8523	1491.8521	0.0001	104-115	LHRE…RAIK	
773.3917	773.3919	-0.0002	107-112	EVEELR	
1085.6087	1085.608	0.0007	107-112	EVEE…RAIK	
2106.0465	2106.0449	0.0016	113-132	AIKV…MCPR	Oxidation (M)
2090.0524	2090.05	0.0024	113-132	AIKV…MCPR	
2552.2055	2552.2033	0.0022	113-135	AIKV…RCER	Oxidation (M)
2536.2101	2536.2084	0.0018	113-135	AIKV…RCER	
1793.8319	1793.8288	0.0031	116-132	VGPT…MCPR	Oxidation (M)
1777.8401	1777.8339	0.0062	116-132	VGPT…MCPR	
2239.989	2239.9872	0.0019	116-135	VGPT…RCER	Oxidation (M)
3106.4551	3106.4482	0.0069	116-144	VGPT…SPSR	Oxidation (M)
1330.6316	1330.6299	0.0016	133-144	CERV…SPSR	
2092.1093	2092.1099	-0.0006	133-152	CERV…PVPAK	
1645.9526	1645.9515	0.0011	136-152	VTPA…PVPAK	
1774.0481	1774.0465	0.0016	136-153	VTPA…PAKK	
779.492	779.4905	0.0014	145-152	AVVPVPAK	
907.5856	907.5855	0.0001	145-153	AVVP…PAKK	
1763.0093	1763.0094	0.0001	145-160	AVVP…PQER	
1001.5296	1001.5294	0.0001	153-160	KTFPPQER	
1272.6567	1272.6575	-0.0008	153-162	KTFP…ERDR	
873.4349	873.4345	0.0004	154-160	TFPPQER	
1144.5628	1144.5625	0.0003	154-162	TFPP…ERDR	

[1] Peptides corresponding to the putative N-glycosylation site are in red.
[2] Only experimental masses that matched calculated masses are listed in the table.
[3] The calculated mass is the relative molecular mass calculated from the matched peptide sequence.
[4] The calculated difference between the experimental mass and the calculated mass.
[5] Fragment numbering is based on the predicted N-terminus of the ATHB17Δ113 protein.
[6] For peptide matches greater than nine amino acids in length the first 4 residues and last 4 residues are show separated by dots (…).

001 MNRLPSSEDG DDEEFSHDDG SAPPRKKLRL TREQSRLLED SFRQNHTLNP
051 KQKEVLAKHL MLRPRQIEVW FQNRRARSKL KQTEMECEYL KRWFGSLTEE
101 NHRLHREVEE LRAIKVGPTT VNSASSLTMC PRCERVTPAA SPSRAVVPVP
151 AKKTFPPQER DR

Figure 19. Sequence Coverage Map of the MON 87403-produced ATHB17Δ113 Protein

Boxed regions correspond to peptides that were identified from the MON 87403-produced ATHB17Δ113 protein sample by LC-MS/MS analysis. In total, 99% (160 of 162) of the expected protein sequence was identified. The expected amino acid sequence of the MON 87403-produced ATHB17Δ113 protein was deduced from the *ATHB17* gene present in MON 87403 and RT-PCR analysis of MON 87403 *ATHB17* RNA transcript.

B1(b)(iii) N-terminal sequence analysis

Because of the low levels of ATHB17Δ113 protein in MON 87403, determination of the N-terminal sequence was not possible using the typical Edman degradation method. Instead, the N-terminal sequence was determined by LC-MS/MS (See Section B1(b)(ii)) and the first 15 amino acids were compared to the N-terminal sequence of the *E. coli*-produced ATHB17Δ113 protein (determined by Edman degradation) and the expected sequence. The N-terminal sequence of MON 87403-produced ATHB17Δ113 protein was as expected with the exception that a portion of the protein population lacked the N-terminal methionine residue (Figure 20). Removal of the N-terminal methionine by methionine aminopeptidase is a common co-translational and/or post-translational modification in plants (Meinnel and Giglione, 2008). The N-terminal sequence information, therefore, confirms the identity of the ATHB17Δ113 protein isolated from MON 87403 and the intactness of its N-terminus.

Figure 20. N-terminal Sequence Evaluation of the MON 87403-produced ATHB17Δ113 Protein

The experimental sequence obtained from the MON 87403-produced ATHB17Δ113 protein was compared to the expected sequence and the sequence obtained from the *E. coli*-produced ATHB17Δ113 protein (lot# 11380003). The single letter IUPAC-IUB amino acid code is M, methonine; N, asparagine; R, arginine; L, leucine; P, proline; S, serine I, E; glutamic acid; D, asparagine; G, glycine; F, phenylalanine;. X indicates that the residue was not identified.

B1(b)(iv) Results of western blot analysis of the MON 87403-produced ATHB17Δ113 and immunoreactivity comparison to *E. coli*-produced ATHB17Δ113

Western blot analysis of the enriched leaf extract from MON 87403 was conducted using rabbit anti-ATHB17Δ113 polyclonal antibodies to 1) to further confirm the identity of MON 87403-produced ATHB17Δ113, 2) determine the concentration of the MON 87403-produced ATHB17Δ113 in the enriched sample from the leaf tissue of MON 87403; and 3) assess the immunoreactivity of the MON 87403-produced and the *E. coli*-produced ATHB17Δ113 proteins.

The results demonstrated that immunoreactive proteins with the same electrophoretic mobility were present in lanes loaded with the MON 87403-produced (Figure 21, Lanes 8-11) or the *E. coli*-produced (Figure 21, Lanes 2-7) ATHB17Δ113 protein. As expected, the signal intensity increased with increasing load amounts of the MON 87403-produced and *E. coli*-produced ATHB17Δ113 proteins, supporting identification of MON 87403-produced ATHB17Δ113 protein. Quantitative western blot analysis was conducted to determine the concentration of MON 87403-produced ATHB17Δ113 in the enriched sample. The concentration of the MON 87403-produced ATHB17Δ113 protein in the enriched sample from MON 87403 leaf tissue was determined to be 0.008 mg/ml (Table 10). The intractability of the MON 87403-produced ATHB17Δ113 protein limited the technical methods available for quantification of the ATHB17Δ113 protein in the sample. The very

low total protein concentration and limited quantity of the enriched MON 87403-produced ATHB17Δ113 protein sample prevented accurate determination of the total protein concentration in the sample. Therefore, the total amount of protein loaded on the western blot for the MON 87403-produced ATHB17Δ113 was not defined and a direct quantitative comparison of the immunoreactive signal from the *E. coli*- and MON 87403-produced proteins was not feasible. However, immunoreactivity of both the MON 87403-produced and the *E. coli*-produced ATHB17Δ113 proteins with the same anti-ATHB17Δ113 antibody supports that these two proteins are equivalent.

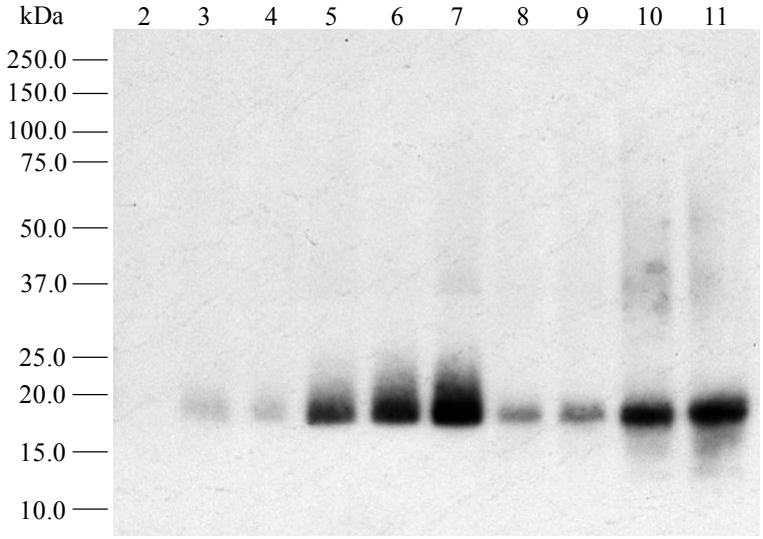


Figure 21. Western Blot Analysis of the MON 87403-produced and *E. coli*-produced ATHB17Δ113 Protein

Aliquots of the MON 87403-produced and the *E. coli*-produced ATHB17Δ113 proteins were subjected to SDS-PAGE and electrotransferred to a nitrocellulose membrane. Proteins were detected using anti-ATHB17Δ113 antibodies as the primary antibodies. Immunoreactive bands were visualized using HRP-conjugated secondary antibodies and an ECL system. The molecular weights (kilodalton, kDa) of the standards are shown on the left. The 10 minute exposure is shown. Lanes 1 and 12 were cropped from the image. Lane designations are as follows.

Lane	Sample	Amount (ng)	Dilution
1	Precision Plus Protein™ Standards	-	-
2	*E. coli*-Produced ATHB17Δ113	0.4	-
3	*E. coli*-Produced ATHB17Δ113	0.7	-
4	*E. coli*-Produced ATHB17Δ113	1.1	-
5	*E. coli*-Produced ATHB17Δ113	1.4	-
6	*E. coli*-Produced ATHB17Δ113	1.8	-
7	*E. coli*-Produced ATHB17Δ113	2.1	-
8	MON 87403-produced ATHB17Δ113	-	1:80
9	MON 87403-produced ATHB17Δ113	-	1:80
10	MON 87403-produced ATHB17Δ113	-	1:40
11	MON 87403-produced ATHB17Δ113	-	1:40
12	Precision Plus Protein™ Standards	-	-

Table 10. Concentration of the MON 87403-produced ATHB17Δ113 Protein in an Enriched Sample from the Leaf Tissue of MON 87403

Concentration (mg/ml)[1]	SD[2]
0.008	0.0017

[1] Value refers to mean calculated based on n=4. The result of the concentration determination was round to three decimal places.
[2] Standard Deviation

B1(b)(v) Functional activity of the MON 87403 ATHB17Δ113 protein

Due to the extremely low level of the ATHB17Δ113 protein in the seed and leaf of MON 87403 (<0.00028 µg/g dry weight in seed and 0.014 µg/g dry weight in leaf) the final MON 87403-produced ATHB17Δ113 protein used in the initial characterisation/equivalence study was limited in both purity and quantity, precluding determination of the DNA binding activity of the MON 87403-produced ATHB17Δ113 protein. The functional equivalence of the MON 87403-produced and *E.coli*-produced ATHB17Δ113 proteins was evaluated indirectly by measuring the DNA binding activity of the *E.coli*-produced ATHB17Δ113 protein and evaluating the results in the context of the known activity of other Arabidopsis HD-Zip II proteins.

Arabidopsis HD-ZipII proteins can bind to the DNA sequence CAAT(C/G)ATTG *in vitro* (Sessa *et al*., 1993; Rice *et al*., 2014) and when Arabidopsis HD-ZipII proteins, including ATHB17, are overexpressed in transgenic Arabidopsis plants, they can repress transcription from promoters containing the DNA sequence CAAT(C/G)ATTG (Ohgishi *et al*., 2001; Sawa *et al*., 2002; Hymus *et al*., 2013). The DNA binding activity of the *E. coli*-produced ATHB17Δ113 protein used in safety assessments was evaluated by measuring the ability of ATHB17Δ113 to specifically bind to the immobilized target DNA sequence CAAT(C/G)ATTG in a plate-based ELISA-like assay. In this assay, results are reported as the amount of ATHBΔ113 bound per well of a microplate and expressed as the percent of ATHB17Δ113 bound relative to the amount of reference ATHB17Δ113 protein bound per well of a microplate (or relative DNA binding activity).

The relative DNA binding activity of the *E. coli*-produced ATHB17Δ113 protein was determined to be 85%, 75%, and 80% during three separate characterisations (Table 11). Thus, *E. coli*-produced ATHB17Δ113 binds to the target DNA sequence CAAT(C/G)ATTG, which confirms the functional activity of the protein. The activity measured for *E. coli*-produced ATHB17Δ113 protein is consistent with what would be expected for ATHB17Δ113 in MON 87403 based on extensive characterisation of Arabidopsis HD-Zip II proteins in the published literature, providing an indirect demonstration of functional equivalence between the two proteins. This further supports that the *E. coli*-produced ATHB17Δ113 protein can serve as an appropriate surrogate for the MON 87403-produced ATHB17Δ113 protein in protein safety assessments.

Table 11. DNA binding activity of *E.coli*-produced ATHB17Δ1131

Sample	ATHB17Δ113 per Well (ng)	Bound ATHB17Δ113 per Well (ng)[1]	Relative DNA Binding Activity[2]
Characterisation	5.0	4.2	85%
Re-Characterisation No. 1	3.5	2.6	75%
Re-Characterisation No. 2	5.0	4.0	80%

[1]Results obtained from Characterisation, Re-characterisation No. 1 and Re-characterisation No. 2 data records.
[2]Mean relative DNA binding activity determined from n=3.

B1(b)(vi) Glycosylation analysis of the MON 87403 ATHB17Δ113 protein

The ATHB17Δ113 protein sequence contains a consensus potential N-glycosylation sequence (NHT, starting at amino acid position 45), however it lacks the N-terminal signal sequence required for transport to the endoplasmic reticulum, the site of N-glycosylation (Pattison and Amtmann, 2009; Vitale and Denecke, 1999). Thus, this suggests that the MON 87403-produced ATHB17Δ113 protein is predicted not to be glycosylated. This is similar to the erythropoietin protein, which contains potential N-glycosylation sites but is not glycosylated (Elliott *et al*., 2004). The LC-MS/MS amino acid sequencing characterisation of MON 87403-produced ATHB17Δ113 protein (see Section B1(b)(ii)) indicate that it is not glycosylated. Two peptides that correspond to the sequence region that contains the putative N-linked glycosylation site of MON 87403-produced ATHB17Δ113 were observed and sequenced (Table 8). These peptides cover residues 33 to 51 and 37 to 51 of the expected ATHB17Δ113 sequence. The difference between the experimentally observed masses for both peptides and the calculated masses based upon unmodified amino acid sequences was less than 0.002 Da. This result supports the conclusion that there are no sugar moieties or other post-translational modifications on the putative N-linked glycosylation site of the MON 87403-produced ATHB17Δ113 protein. Except for a few known proteins, *E.coli* produced proteins are not glycosylated. This result further justifies *E.coli*-produced ATHB17Δ113 protein as an appropriate surrogate for evaluation of the safety of the MON 87403-produced ATHB17Δ113 protein.

B1(b)(vii) MON 87403-produced ATHB17Δ113 protein identity and equivalence conclusion

A combination of analytical techniques was used to characterize the MON 87403-produced ATHB17Δ113 protein enriched from leaf of MON 87403. The concentration of the MON 87403-produced ATHB17Δ113 was 0.008 mg/ml in the enriched leaf fraction. The apparent MW and purity of the MON 87403-produced ATHB17Δ113 were 22.4 kDa and 3%, respectively. Identity of the MON 87403-produced ATHB17Δ113 was confirmed by both MALDI-TOF tryptic mass fingerprint analysis and western blot analysis using immunoreactive band recognition with anti-ATHB17Δ113 antibodies. LC-MS/MS analysis of MON 87403-produced ATHB17Δ113 identified the N-terminus of the protein and confirmed that the sequence of the protein matches the expected sequence.

The equivalence of the MON 87403-produced and *E. coli*-produced ATHB17Δ113 proteins was directly evaluated by comparing their apparent MW and immunoreactivity with anti-ATHB17Δ113 antibodies. MON 87403-produced and *E. coli*-produced ATHB17Δ113 both reacted with anti-ATHB17Δ113 antibodies, and the MON 87403-produced ATHB17Δ113 had equivalent apparent MW to *E. coli*-produced ATHB17Δ113, as assessed by SDS-PAGE. In addition, the *E. coli*-produced ATHB17Δ113 protein exhibits the DNA binding activity that is expected based on the characterisation of Arabidopsis HD-Zip II transcription factors. Due to the low concentration and purity of the MON 87403-produced ATHB17Δ113 protein in the enriched leaf extract, equivalence of immunoreactivity and functional activity were assessed qualitatively. Quantitative differences in immunoreactivity and functional activity would be indicative of potential differences in either sequence or post-translational modification status between the two proteins. However, sequencing of MON 87403-produced ATHB17Δ113 protein in the enriched leaf extract by LC-MS/MS of tryptic peptides confirms that it is identical in sequence to the *E. coli*-produced protein and indicates that it has no natural modifications except those common to the N-terminal end of proteins. The LC-MS/MS amino acid sequencing characterisation of MON 87403-produced ATHB17Δ113 protein also indicates that it is not glycosylated. Taken together, the results of these studies demonstrate that the MON 87403-produced ATHB17Δ113 protein and the *E. coli*-produced ATHB17Δ113 protein are equivalent. This demonstration of protein equivalence confirms that the *E. coli*-produced protein is appropriate for evaluation of the safety of the MON 87403-produced ATHB17Δ113 protein.

For details, please refer to [redacted], 2014 (MSL0025829), [redacted], 2015 (MSL0026621), and [redacted], 2015 (MSL0026645).

B2 Antibiotic Resistance Marker Genes

MON 87403 does not contain genes that encode resistance to antibiotic markers. Molecular characterisation data presented in Section A demonstrate the absence of antibiotic resistance marker gene in MON 87403.

B2(a) Clinical importance of antibiotic that GM is resistant to (if any)

Not applicable.

B2(b) Presence in food of antibiotic resistance protein (if any)

Not applicable.

B2(c) Safety of antibiotic protein

Not applicable.

B2(d) If GM organism is micro-organism, is it viable in final food?

Not applicable.

B3 Characterisation of Novel Proteins or Other Novel Substances

B3(a) Biochemical function and phenotypic effects of novel substances

The *Arabidopsis thaliana* HB17 (ATHB17) protein is a member of the Homeodomain-Leucine zipper (HD-Zip) family of transcription factors (Ariel *et al*., 2007). The HD-Zip protein family is characterized by the presence of a leucine zipper (LZ) domain adjacent to the C-terminus of a homeodomain (HD) (Mukherjee *et al*., 2009; Ruberti *et al*., 1991). These two domains, HD and LZ, are individually present in transcription factors found across eukaryotic organisms, however such combination of HD and LZ domains in a single transcription factor is unique to plants; HD-Zip genes have been identified in representative members of all plant lineages, but not in animals or fungi (Ariel *et al*., 2007). The family of HD-Zip proteins is further segregated into four distinct subfamilies designated as classes I, II, III, and IV, based upon the amino acid sequences of conservative structural domains and motifs that convey their DNA specificity and physiological function (Ariel *et al*., 2007). HD-Zip proteins have been shown to bind to specific target DNA sequences and form homo- or hetero- protein dimers with members of their own subfamily, but not with members of other subfamilies of HD-Zip proteins (Sessa *et al*., 1993). Many HD-Zip proteins function as repressors of gene expression, and have been shown in some cases to down regulate the transcription of HD-Zip genes (Henriksson *et al*., 2005; Ohgishi *et al*., 2001; Rice *et al*., 2014; Sessa *et al*., 1993; Sorin *et al*., 2009; Steindler *et al*., 1999).

The ATHB17 protein is part of the HD-Zip II subfamily of proteins (Ariel *et al*., 2007). The activity of HD-Zip II proteins affects diverse processes throughout plant growth and development, including in shade avoidance responses, development of photosynthetic capacity, and reproduction (Ciarbelli *et al*., 2008; Hymus *et al*., 2013; Sawa *et al*., 2002; Sorin *et al*., 2009; Steindler *et al*., 1999). In maize, 18 HD-Zip II genes have been identified through a systematic bioinformatic analysis, and characterisation of 13 of these 18 proteins has shown that they each can function as transcriptional repressors (Rice *et al*., 2014; Zhao *et al*., 2011). ATHB17, like the maize HD-Zip II proteins, can also function as a transcriptional repressor (Rice *et al*., 2014).

As described in Section B1(a), expression of the *ATHB17* gene in MON 87403 results in production of the ATHB17Δ113 protein, which consists of a single polypeptide chain of 162 amino acids, with an apparent molecular weight of ~22 kDa. The ATHB17Δ113 protein, a truncated version of the ATHB17 protein, retains several distinctive structural domains of ATHB17 that place it in the HD-Zip II protein family. These characteristics, highly conserved in HD-Zip II members, include a HD that recognizes the pseudo-palindromic DNA sequence CAAT(C/G)ATTG, and an LZ domain responsible for protein dimerization (Ariel *et al*., 2007).

The truncation of the first 113 amino acids of ATHB17 results of the loss of its unique N-terminal domain and a large part of the repression domain (Ciarbelli *et al*., 2008; Rice *et al*., 2014). Experiments in Arabidopsis protoplasts have shown that this truncation abolishes cytoplasmic localization of the protein while nuclear localization, the expected site of action

for a transcriptional repressor protein, remains apparent (Rice *et al*., 2014). Although ATHB17Δ113 does not function as a transcriptional repressor, the protein retains expected DNA binding properties, as well as an ability to form homo- and hetero- protein dimers with maize HD-Zip II proteins (Rice *et al*., 2014).

Therefore, the expression of the ATHB17Δ113 protein in MON 87403 likely modulates HD-Zip II regulated pathways in the maize ear through dominant-negative interactions with endogenous HD-Zip II proteins. This leads to increased ear biomass and thus, increased sink size at the early reproductive stages compared to control plants providing an increased yield opportunity over a range of environmental conditions where MON 87403 is expected to be grown.

B3(b) Identification of novel substances (e.g. metabolites), levels and site

ATHB17Δ113 protein levels in various tissues of MON 87403 relevant to the risk assessment were determined by a validated enzyme-linked immunosorbent assay (ELISA). Tissues of MON 87403 were collected from four replicate plots planted in a randomized complete block field design during the 2012 growing season from the following five field sites in the United States: Jackson County, Arkansas (site code ARNE); Story County, Iowa (site code IALL); Jefferson County, Iowa (site code IARL); Pawnee County, Kansas (site code KSLA) and Lehigh County, Pennsylvania (site code PAGR). The field sites were representative of maize-producing regions suitable for commercial production. OSL1, OSR1, forage, and grain tissue samples were collected from each replicated plot at all field sites.

The ELISA results obtained for each sample were averaged across the five sites and are summarized in Table 12. The ATHB17Δ113 protein levels in MON 87403 across all samples analyzed from all sites ranged from <LOD to 0.017 µg/g dw. The mean ATHB17Δ113 protein level among all tissue types was highest in OSL1 at 0.014 µg/g dw. The mean ATHB17Δ113 protein level was lowest in grain at <LOD µg/g dw, which corresponds to < 0.00028 µg/g dw, an extremely low expression level in the maize product.

For details, please refer to [redacted], 2015, MSL0026598.

Table 12. Summary of ATHB17Δ113 Protein Levels in Tissues from MON 87403 Grown in 2012 United States Field Trials

Tissue Type[1]	Development Stage[2]	Mean(SD) Range (µg/g dw)[3]	LOQ/LOD[4] (µg/g dw)
Leaf	V3-V4	0.014 (0.0020) 0.0096 – 0.017	0.00109/0.00049
Root	V3-V4	0.0023 (0.0016) 0.00083 – 0.0058	0.00078/0.00065
Forage	R5	0.0018 (0.00064) 0.0011 – 0.0035	0.00078/0.00063
Grain	R6	<LOD (N/A) N/A – N/A	0.00156/0.00028

[1]OSL= over season leaf
OSR= over season root
[2]The crop development stage each tissue was collected.
[3]Protein levels are expressed as the arithmetic mean and standard deviation (SD) as microgram (µg) of protein per gram (g) of tissue on a dry weight basis (dw). The means, SD, and ranges (minimum and maximum values) were calculated for each tissue across all sites. (n=20, except OSR1 where n=16 due to four samples resulting in inconclusive levels.
[4]LOQ=limit of quantitation; LOD=limit of detection.

B3(c) Site of expression of all novel substances and levels

Please refer to Section B3(b).

B3(d) Post-translational modifications to the novel protein(s)

Not applicable.

B3(e) Evidence of silencing, if silencing is the method of modification

Not applicable.

B3(f) History of human consumption of novel substances or similarity to substances previously consumed in food

As described in Section A2(a), the gene encoding ATHB17 is from *Arabidopsis thaliana,* a common plant species that is not intentionally consumed but is phylogenetically related to commonly consumed *Brassica* species.

Because ATHB17 is from Arabidopsis, a non-food plant, the homology of ATHB17Δ113 to proteins present in species widely used for food has been evaluated. The publically available databases Phytozome[3] and BRAD[4] were queried to identify orthologs to ATHB17. Eleven amino acid sequences from several different food crop plant species, including soybean, rice, corn, tomato, potato, orange, papaya, grape, sorghum, napa cabbage (*Brassica rapa* subsp. *pekinensis*), and cabbage (*Brassica oleracea* species) were retrieved. The degree of amino acid sequence identity between ATHB17Δ113 and the retrieved sequences was assessed using a BLAST analysis with ATHB17Δ113 as the query. The results are summarized in Table 13 and Table 14. The highest sequence identity was observed when the ATHB17Δ113 sequence was aligned with sequences from the Brassica species *Brassica rapa* (a species including common food crops such as turnip and napa cabbage) and *Brassica oleracea* (a species including common food crops such as cabbage and Brussels sprouts), which show a high degree of identity (80 - 83% identical) across the entire ATHB17Δ113 amino acid sequence (Table 13). A high degree of identity (up to 77%) with the core of the ATHB17Δ113 amino acid sequence which encompasses the functional HD and LZ domains, is observed for proteins from a number of other species (Table 14). Thus ATHB17Δ113 shares significant sequence identity and structural similarity with proteins that are present in frequently consumed food plants, including plants whose food products are commonly cooked before consumption and those whose products are frequently consumed without cooking. This similiarity establishes that there is a history of safe consumption of this class of proteins by humans and other animals.

3 http://www.phytozome.net
4 http://brassicadb.org/brad/

Table 13. Amino Acid Sequence Identity from Sequence Alignment between MON 87403-produced ATHB17Δ113 and Other Proteins Present in Plants

Crop Species	Common Name	Sequence ID[1]	Sequence Identity[2]	Amino Acid Range[3]
Brassica rapa	Napa cabbage	Bra024888	83%	1-161
Brassic oleracea	Cabbage	Bol014821	80%	1-161
Carica papaya	Papaya	evm.model.supercontig_65.141	67%	1-158

[1]Sequences were retrieved from the Phytozome v9.1 website (http://www.phytozome.net) or, for *Brassica oleracea*, the Brassica Database (BRAD) website (http://brassicadb.org/brad/), using the ATHB17 amino acid sequence (ID: AT2G01430.1) as the query.
[2]Sequence identity is as reported in a BLAST alignment report, using ATHB17Δ113 as the query sequence.
[3]Amino acid range refers to the aligned span of amino acids in ATHB17Δ113, which is 162 amino acids in length.

Table 14. Amino Acid Sequence Identity from Partial Sequence Alignment between the HD and LZ of MON 87403-produced ATHB17Δ113 and Other Proteins Present in Plants

Crop Species	Common Name	Sequence ID[1]	Sequence Identity[2]	Amino Acid Range[3]
Vitis vinifera	Grape	GSVIVT01011754001	77%	11-137
Solanum lycopersicum	Tomato	Solyc05g008050.1.1	75%	22-137
Citrus sinensis	Sweet orange	orange1.1g026325m	73%	22-137
Solanum tuberosum	Potato	PGSC0003DMP400053056	75%	11-140
Glycine max	Soybean	Glyma20g01770.1	72%	1-140
Sorghum bicolor	Sorghum	Sobic.003G235900.1	67%	6-136
Oryza sativa	Rice	LOC_Os01g45570.1	69%	10-142
Zea mays	Corn	Zmhdz18_GRMZM2G126239_T01	61%	7-158

[1]Sequences were retrieved from the Phytozome v9.1 website (http://www.phytozome.net) or, for *Brassica oleracea*, the Brassica Database (BRAD) website (http://brassicadb.org/brad/), using the ATHB17 amino acid sequence (ID: AT2G01430.1) as the query.
[2]Sequence identity is as reported in a BLAST alignment report, using ATHB17Δ113 as the query sequence.
[3]Amino acid range refers to the aligned span of amino acids in ATHB17Δ113, which is 162 amino acids in length.

B4 Assessment of Potential Toxicity

The assessment of the potential toxicity of an introduced protein takes into account several aspects of its biochemical characteristics (Delaney *et al*., 2008). A protein introduced into maize is not likely to be associated with toxicity if: 1) the protein lacks any structural similarity to known toxins or other biologically-active proteins that could cause adverse effects in humans or animals; 2) the protein is structurally and functionally related to proteins with a history of safe consumption; and 3) the protein is readily inactivated or degraded in response to common food processing conditions (*e.g*., heating) and/or digestive enzymes. The lack of any effects in an acute oral mammalian toxicity study performed at dose levels substantially greater than anticipated human exposure levels can provide further confirmation that an introduced protein is unlikely to pose a significant risk to human or animal health.

B4(a) Bioinformatic comparison (aa) of novel protein(s) to toxins

The assessment of the potential for protein toxicity includes bioinformatic analysis of the amino acid sequence of the introduced protein. The goal of the bioinformatic analysis is to ensure that the introduced protein does not share homology to known toxins or anti-nutritional proteins associated with adverse health effects.

Potential structural similarities shared between ATHB17Δ113 protein with sequences in a protein database were evaluated using the FASTA sequence alignment tool. The FASTA program directly compares amino acid sequences (i.e., primary, linear protein structure) and the alignment data may be used to infer shared higher order structural similarities between two sequences (i.e., secondary and tertiary protein structures). Proteins that share a high degree of similarity throughout the entire sequence are often homologous. Homologous proteins often have common secondary structures, common three-dimensional configuration, and, consequently, may share similar functions (Caetano-Anollés *et al*., 2009; Illergård *et al*., 2009).

FASTA bioinformatic alignment searches using the ATHB17Δ113 amino acid sequence were performed with a toxin database to identify possible homology with proteins that may be harmful to human and animal health. The toxin database, TOX_2013, is a subset of sequences derived from the PRT_2013 database, that was selected using a keyword search and filtered to remove likely non-toxin proteins. The TOX_2013 database contains 8,881 sequences.

An *E*-score acceptance criteria of 1×10^{-5} or less for any alignment was used to identify proteins from the TOX_2013 database with potential for significant shared structural similarity and function with ATHB17Δ113 protein. The *E*-score is a statistical measure of the likelihood that the observed similarity score could have occurred by chance in a search. A larger *E*-score indicates a lower degree of similarity between the query sequence and the sequence from the database. Typically, alignments between two sequences require an *E*-score of 1×10^{-5} or less to be considered to have sufficient sequence similarity to infer

homology. The results of the search comparisons showed that no relevant alignments were observed against proteins in the TOX_2013 database.

The results of the bioinformatic analyses demonstrated that no structurally relevant similarity exists between the ATHB17Δ113 protein and any sequence in the TOX_2013 database, as no alignments displaying an *E*-score < 1×10^5 were observed.

For details, please refer to [redacted], 2013 (MSL0025242).

Similarity of ATHB17Δ113 to Consumed Proteins

As described in Section B3(f), ATHB17Δ113 shares significant sequence identity and structural similarity with proteins present in frequently consumed food plants, including plants whose food products are commonly cooked before consumption and those whose products are frequently consumed without cooking. This similiarity establishes that there is a history of safe consumption of this class of proteins by humans and other animals.

B4(b) Degradation and heat stability of the ATHB17Δ113 protein

B4(b)(i) Stability to proteolysis in gastrointestinal model system

Proteins newly-produced in biotechnology-derived crops are evaluated for their safety for human and animal consumption. Proteins are an essential dietary component for humans and animals, and most are rapidly degraded to the component amino acids for nutritional purposes (Hammond and Jez, 2011). Although the vast majority of ingested proteins are non-allergenic, a small set of proteins or their fragments have been associated with a variety of gastrointestinal and systemic manifestations of immune-mediated allergy. One characteristic of several food allergens is their ability to withstand proteolytic digestion to amino acids by enzymes present in the gastrointestinal tract (Astwood *et al*., 1996; Moreno *et al*., 2005; Vassilopoulou *et al*., 2006; Vieths *et al*., 1999), although like most aspects of biology, exceptions to the correlation can be identified (Fu *et al*., 2002). The enzymatic degradation of an ingested protein by exposure to gastric pepsin and/or intestinal pancreatic proteases (e.g., pancreatin) makes it highly unlikely that either the intact protein or protein fragment(s) will reach the absorptive epithelial cells of the small intestine where antigen processing cells reside (Moreno *et al*., 2005). Therefore, the susceptibility of ATHB17Δ113 to the presence of pepsin, using an assay protocol that has been standardized based on results obtained from an international, multi-laboratory ring study (Thomas *et al*., 2004). The susceptibility of proteins in the presence of pancreatin has also been used as a separate test system to assess the digestibility of food components (Okunuki *et al*., 2002; Yagami *et al*., 2000).

Degradation of MON 87403 ATHB17Δ113 in the Presence of Pepsin

Digestibility of the *E. coli*-produced ATHB17Δ113 protein in simulated gastric fluid (SGF), in the presence of pepsin, was evaluated over time by analyzing digestion mixtures incubated for targeted time intervals following a standardized protocol validated in an international, multi-laboratory ring study (Thomas *et al*., 2004). The study showed that the results of *in*

vitro pepsin digestion assays using this protocol were reproducibile and consistent for determining the digestive stability of a protein. This standardized *in vitro* pepsin digestion protocol utilized a physiologically relevant acidic buffer appropriate for pepsin activity. The susceptibility of *E.coli*-produced ATHB17Δ113 to pepsin degradation was assessed by visual analysis of a Brilliant Blue G Colloidal stained SDS-PAGE gel and by visual analysis of a western blot probed with an anti-ATHB17Δ113 polyclonal antibody. Both visualization methods were run concurrently with separate SDS-PAGE and western blot analyses to estimate the limit of detection (LOD) of the ATHB17Δ113 protein for each method.

For SDS-PAGE analysis of susceptibility of ATHB17Δ113 to pepsin degradation, the gel was loaded with 1 µg of a protein sample (based on pre-pepsin protein concentrations) containing the *E. coli*-produced ATHB17Δ113 protein for each time point (Figure 22, Panel A). Visual examination of the SDS-PAGE data showed that full-length ATHB17Δ113 protein was degraded below the LOD within 0.5 min of incubation with pepsin (Figure 22, Panel A, Lane 5). The LOD of the ATHB17Δ113 protein was observed at approximately 25 ng total protein loading (Figure 22, Panel B). The LOD was used to calculate the maximum relative amount of full-length ATHB17Δ113 protein that could remain in the absence of visual detection after pepsin treatment, which corresponded to approximately 2.5% of the ATHB17Δ113 protein loaded. Therefore, based on the LOD, more than 97.5% (100% - 2.5% = 97.5%) of the full-length ATHB17Δ113 protein was degraded within 0.5 min of exposure to pepsin. A peptide fragment of ~4-5 kDa was observed in the 0.5 minute time points of pepsin exposure, but was not observed after 2 minutes of incubation with pepsin (Figure 22, Panel A, Lanes 5-6).

No change in the ATHB17Δ113 protein band intensity was observed in the absence of pepsin in the SGF 0 minute No Pepsin Control and SGF 60 minute No Pepsin Control (Figure 22, Panel A, Lanes 3 and 12). This indicates that the digestion of the ATHB17Δ113 protein was due to pepsin proteolytic activity and not due to instability of the protein in the test system.

The SGF 0 minute No Test Protein Control and SGF 60 minute No Test Protein Control (Figure 22, Panel A, Lanes 2 and 13) demonstrated that the pepsin is stable throughout the experimental phase of this assessment.

For details, please refer to [redacted], 2014 (MSL0025516).

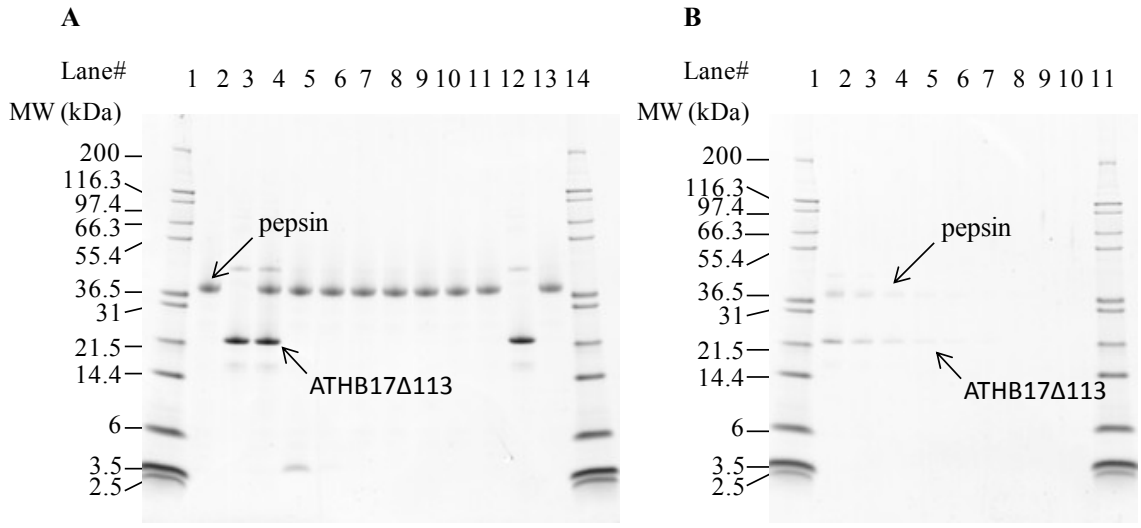


Figure 22. Colloidal Blue Stained SDS-PAGE Gel Showing the Degradation of Purified *E. coli*-produced ATHB17Δ113 Protein in Simulated Gastric Fluid

Colloidal Brilliant Blue G stained SDS-PAGE gels were used to assess the degradation of ATHB17Δ113 in SGF. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded. In each gel, ATHB17Δ113 protein migrated to approximately 22 kDa and pepsin to approximately 38 kDa. Blank lanes were cropped from the images.

A: ATHB17Δ113 protein degradation in SGF. Based on initial protein concentrations, 1 µg of test protein was loaded in each lane containing ATHB17Δ113 protein.

B: LOD determination. Indicated amounts of the ATHB17Δ113 protein from the SGF T0 sample were loaded to estimate the LOD of the ATHB17Δ113 protein.

Lane	Sample	Incubation Time (min)
1	Mark 12 MWM	-
2	SGF 0 min No Test Protein Control	0
3	SGF 0 min No Pepsin Control	0
4	SGF T0	0
5	SGF T1	0.5
6	SGF T2	2
7	SGF T3	5
8	SGF T4	10
9	SGF T5	20
10	SGF T6	30
11	SGF T7	60
12	SGF 60 min No Pepsin Control	60
13	SGF 60 min No Test Protein Control	60
14	Mark 12 MWM	-

Lane	Sample	Amount (ng)
1	Mark 12 MWM	-
2	SGF T0	202
3	SGF T0	101
4	SGF T0	50
5	SGF T0	25
6	SGF T0	13
7	SGF T0	6.3
8	SGF T0	3.2
9	SGF T0	1.6
10	SGF T0	0.8
11	Mark 12 MWM	-
		-
		-
		-

For western blot analysis of ATHB17Δ113 pepsin susceptibility, the ATHB17Δ113 protein was loaded with approximately 20 ng per lane of total protein (based on initial protein concentrations) for each of the time points. Visual examination of the western blot data showed that the full-length ATHB17Δ113 protein was degraded below the LOD within 0.5 minutes of incubation in SGF (Figure 23, Panel A, Lane 5). The LOD of the ATHB17Δ113 protein by western blot analysis was determined to be approximately 1.3 ng (Figure 23, Panel B). The LOD was used to calculate the maximum relative amount of full-length ATHB17Δ113 protein that could remain visually undetected after pepsin exposure, which corresponded to approximately 6.5% of the total protein loaded. Based on the western blot LOD for the ATHB17Δ113 protein, more than 93.5% (100% - 6.5% = 93.5%) of the ATHB17Δ113 protein was degraded within 0.5 minutes. No peptide fragments were detected at any time point in the western blot analysis.

No change in the ATHB17Δ113 protein band intensity was observed in the absence of pepsin in the SGF 0 minute No Pepsin Control and SGF 60 minute No Pepsin Control (Figure 23, Panel A, Lanes 3 and 12) thus, the ATHB17Δ113 protein was stable in the SGF without pepsin.

No immunoreactive bands were observed in SGF 0 minute No Protein Control and SGF 60 minute No Protein Control (Figure 23, Panel A, Lanes 2 and 13). This result indicates that there was no non-specific interaction between the SGF and the ATHB17Δ113-specific antibody under these experimental conditions.

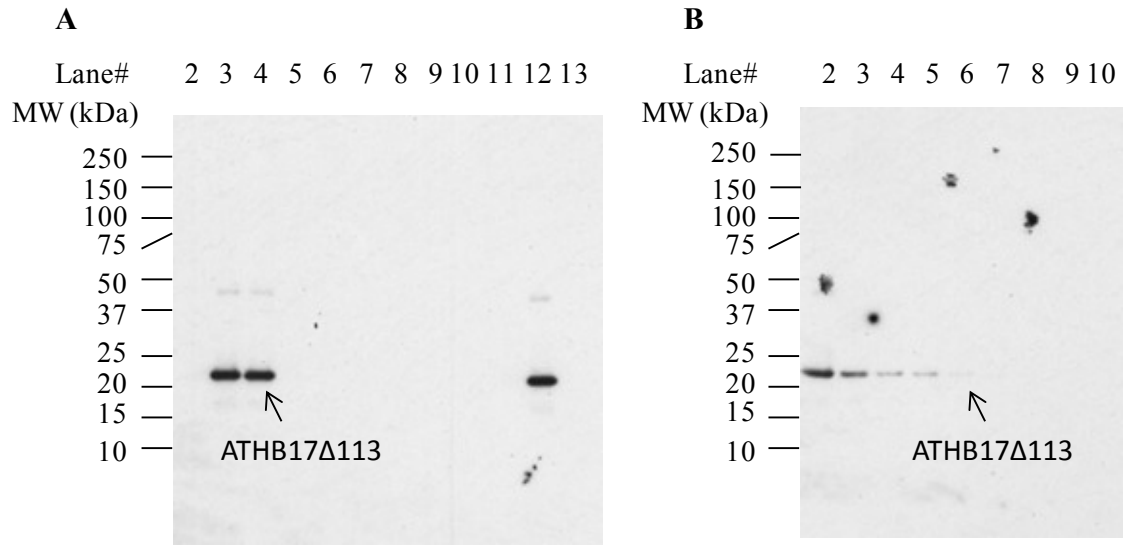


Figure 23. Western Blot Analysis of Purified *E. coli*-produced ATHB17Δ113 Protein in Simulated Gastric Fluid

Western blots probed with an anti-ATHB17Δ113 antibody were used to assess the degradation of ATHB17Δ113 in SGF. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded (cropped from images). Blank lanes were cropped from the images. A 2 min exposure is shown.
A: ATHB17Δ113 protein degradation in SGF. Based on initial protein concentrations, 20 ng of test protein was loaded in each lane containing ATHB17Δ113 protein.
B: LOD determination. Indicated amounts of the ATHB17Δ113 protein from the SGF T0 sample were loaded to estimate the LOD of the ATHB17Δ113 protein.

Lane	Sample	Incubation Time (min)
1	Precision Plus MWM	-
2	SGF 0 min No Test Protein Control	0
3	SGF 0 min No Pepsin Control	0
4	SGF T0	0
5	SGF T1	0.5
6	SGF T2	2
7	SGF T3	5
8	SGF T4	10
9	SGF T5	20
10	SGF T6	30
11	SGF T7	60
12	SGF 60 min No Pepsin Control	60
13	SGF 60 min No Test Protein Control	60
14	Precision Plus MWM	-

Lane	Sample	Amount (ng)
1	Precision Plus MWM	-
2	SGF T0	20.2
3	SGF T0	10.1
4	SGF T0	5.1
5	SGF T0	2.5
6	SGF T0	1.3
7	SGF T0	0.6
8	SGF T0	0.3
9	SGF T0	0.2
10	SGF T0	0.1
11	Precision Plus MWM	-

Degradation of MON 87403 ATHB17Δ113 in the Presence of Pancreatin

The degradation of the *E. coli*-produced ATHB17Δ113 protein in simulated intestinal fluid (SIF) containing pancreatin was assessed by western blot analysis (Figure 24). The western blot used to assess the ATHB17Δ113 protein degradation in the presence of pancreatin (Figure 24, Panel A) was run concurrently with the western blot used to estimate the LOD (Figure 24, Panel B) of the ATHB17Δ113 protein. The observed LOD of the ATHB17Δ113 protein by this western blot analysis was approximately 1.3 ng of protein per lane. The LOD was used to calculate the maximum relative amount of ATHB17Δ113 protein that could remain visually undetected after degradation, which corresponded to approximately 6.5% of the total protein loaded.

The gel used to assess degradation of the ATHB17Δ113 protein by western blot was loaded with 20 ng total protein (based on pre-degradation concentrations) for each of the incubation time points. Western blot analysis demonstrated that a band corresponding to the ATHB17Δ113 protein was degraded to a level below the LOD within 5 minutes of incubation in SIF (Figure 24, Panel A, Lane 5), the first time point assessed. Therefore, based on the LOD, more than 93.5% (100% - 6.5% = 93.5%) of the ATHB17Δ113 protein was digested within 5 minutes. No peptide fragments were detected at any time point in the western blot analysis.

No change in the ATHB17Δ113 band intensity was observed in the absence of pancreatin in the SIF 0 hour No Pancreatin Control and SIF 24 hour No Pancreatin Control (Figure 24, Panel A, Lanes 3 and 13). This indicates that the ATHB17Δ113 was stable in the SIF without pancreatin at 37 ± 2ºC over the course of the experiment.

No immunoreactive bands were observed in SIF 0 hour No Test Protein Control and SIF 24 hour No Test Protein Control (Figure 24, Panel A, Lanes 2 and 14), demonstrating the absence of non-specific antibody interactions within the SIF test system.

For details, please refer to [redacted], 2014 (MSL0025516).

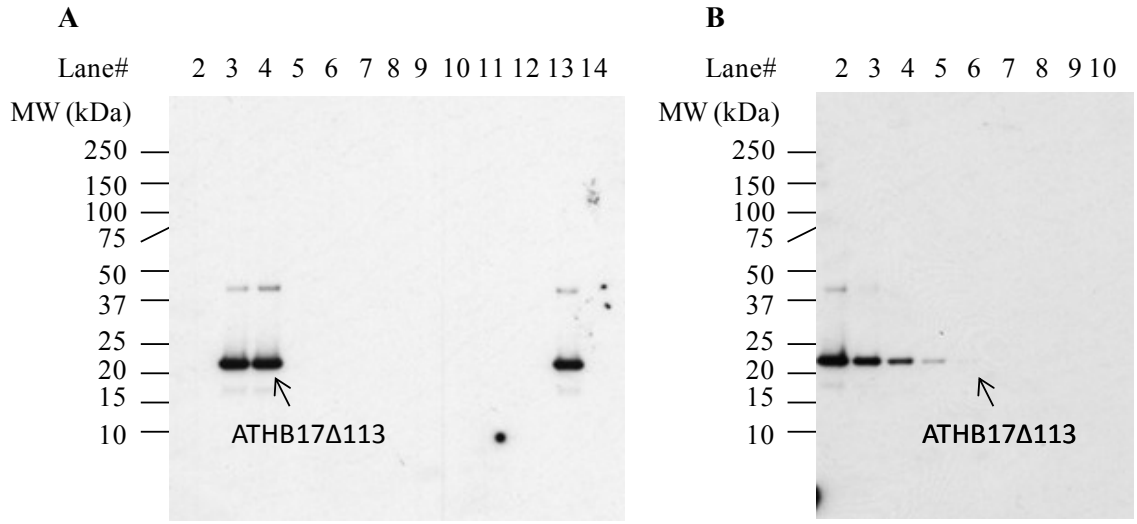


Figure 24. Western Blot Analysis of Purified *E. coli*-produced ATHB17Δ113 Protein in Simulated Intestinal Fluid

Western blots probed with an anti-ATHB17Δ113 antibody were used to assess degradation of ATHB17Δ113 in SIF. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded (cropped from images). Blank lanes were cropped from the images. A 5 minute exposure is shown.

A: ATHB17Δ113 protein degradation in SIF. Based on initial protein concentrations, 20 ng of test protein was loaded in each lane containing ATHB17Δ113 protein.

B: LOD determination. Indicated amounts of the test protein from the SIF T0 sample were loaded to estimate the LOD of the ATHB17Δ113 protein.

Lane	Sample	Incubation Time
1	Precision Plus MWM	-
2	SIF 0 min No Test Protein Control	0
3	SIF 0 min No Pancreatin Control	0
4	SIF T0	0
5	SIF T1	5 min
6	SIF T2	15 min
7	SIF T3	30 min
8	SIF T4	1 hr
9	SIF T5	2 hr
10	SIF T6	4 hr
11	SIF T7	8 hr
12	SIF T8	24 hr
13	SIF 24 hr No Pancreatin Control	24 hr
14	SIF 24 hr No Test Protein Control	24 hr
15	Precision Plus MWM	-

Lane	Sample	Amount (ng)
1	Precision Plus MWM	-
2	SIF T0	20.0
3	SIF T0	10.0
4	SIF T0	5.0
5	SIF T0	2.5
6	SIF T0	1.3
7	SIF T0	0.6
8	SIF T0	0.3
9	SIF T0	0.2
10	SIF T0	0.1
11	Precision Plus MWM	-

B4(b)(ii) Heat stability of the ATHB17Δ113 protein

Temperature can have a profound effect on the structure and function of proteins. Heat treatment is widely used in the preparation of foods derived from maize grain (Hammond and Jez, 2011). The ATHB17Δ113 protein is undetectable in MON 87403 grain and, therefore there is immeasurably low exposure to the protein in foods derived from MON 87403 grain. Consequently, it is highly improbable that there is a meaningful allergenicity or toxicity risk posed by food uses of MON 87403. However, an assessment of the effect of heating, as a surrogate for the conditions encountered during the preparation of foods from MON 87403 grain was conducted. It is reasonable that such processing will have an effect on the functional activity and structure of ATHB17Δ113 when consumed in different food products, thus reducing any potential safety concerns posed by the protein.

The effect of heat treatment on the activity of ATHB17Δ113 was evaluated using the *E. coli*-produced ATHB17Δ113 protein. Heat-treated samples and an unheated control sample of the *E. coli*-produced ATHB17Δ113 protein were analyzed: 1) using a functional assay to assess the impact of temperature on the DNA binding activity of the ATHB17Δ113 protein; and 2) using SDS-PAGE to assess the impact of temperature on protein integrity.

Aliquots of the ATHB17Δ113 protein were heated to 25, 37, 55, 75, and 95°C for 15 and 30 minutes, while a separate aliquot of the ATHB17Δ113 protein was maintained on ice for the duration of the heat treatments to serve as a temperature control. The effect of heat treatment on the relative activity of the ATHB17Δ113 protein was evaluated using a DNA binding activity assay. The effect of heat treatment on the integrity of the the ATHB17Δ113 protein was evaluated using SDS-PAGE analysis of the heated and temperature controlled *E.coli*-produced ATHB17Δ113 protein samples.

The effects of heating on the relative DNA binding activity of the ATHB17Δ113 protein are presented in Table 15 and Table 16. The relative activity was unaffected at 25, 37 and 55ºC for the 15 and 30 minute incubation periods. The relative activity was reduced significantly at 75 and 95°C for both incubation periods, indicating the ATHB17Δ113 protein is partially labile when heated for 15 or 30 minutes at temperatures ≥75°C. These results suggest that temperatures typical of food uses of maize grain would reduce the activity of the ATHB17Δ113 protein.

Analysis by SDS-PAGE stained with Brilliant Blue G Colloidal demonstrated that the reference standard and control samples contained major bands with an apparent molecular weight of ~22 kDa, corresponding to the ATHB17Δ113 protein (Figure 25 and Figure 26). No apparent decrease in the intensity of this band was observed in heat treated ATHB17Δ113 protein at any temperatures at 15 or 30 minutes (Figure 25, Lanes 3-7 and Figure 26, Lanes 3-7). Heat treatment of the ATHB17Δ113 protein at 95 ºC for 15 or 30 min resulted in the appearance of lower molecular weight species, which may be due to slight degradation of the protein when exposed to high temperatures (Figure 25, Lane 7 and Figure 26, Lane 7).

These data demonstrate that *E. coli*-produced ATHB17Δ113 protein behaves with a predictable tendency toward loss of functional activity at elevated temperatures. Heat treatment is widely used in the preparation of foods containing components derived from maize grain. Therefore, it is reasonable to conclude that MON 87403 ATHB17Δ113 activity, if present at all, would be decreased in heat-treated food products. The retention of some intact and functionally active ATHB17Δ113 protein after heat treatment would not represent a safety concern because (1) there is no bioinformatic similarity indicating structural relatedness to known toxins or allergens, (2) an acute mouse toxicity study indicated no toxicity of the intact protein, and (3) the protein is rapidly degraded by the digestive enzymes pepsin and pancreatin.

For details, please refer to [redacted], 2014 (MSL0025364).

Table 15. Activity of *E. coli*-produced ATHB17Δ113 Protein after 15 Minutes at Elevated Temperatures

Temperature	Relative DNA Binding Activity[1]	Relative Activity (% of control substance) [2,3]
0 ºC (control substance)	0.63 ± 0.05	100 %
25 ºC	0.70 ± 0.05	111 %
37 ºC	0.80 ± 0.10	127 %
55 ºC	0.69 ± 0.12	110 %
75 ºC	0.48 ± 0.01	76 %
95 ºC	0.37 ± 0.05	59 %

[1]Relative DNA binding activity = [assay-determined protein amount/starting protein amount]. Means and standard deviations were determined from n=3.
[2]ATHB17Δ113 protein activity of the unheated control was assigned 100 % active.
[3]Relative Activity = [relative DNA binding activity of sample/relative DNA binding activity of the unheated control] x 100

Table 16. Activity of *E. coli*-produced ATHB17Δ113 Protein after 30 Minutes at Elevated Temperatures

Temperature	Relative DNA Binding Activity[1]	Relative Activity (% of control substance) [2,3]
0 ºC (control substance)	0.63 ± 0.05	100 %
25 ºC	0.78 ± 0.08	124 %
37 ºC	0.80 ± 0.09	127 %
55 ºC	0.68 ± 0.03	108 %
75 ºC	0.39 ± 0.02	62 %
95 ºC	0.24 ± 0.03	38 %

[1]Relative DNA binding activity = [assay-determined protein amount/starting protein amount]. Means and standard deviations were determined from n=3.
[2]ATHB17Δ113 protein activity of the unheated control was assigned 100 % active.
[3]Relative Activity = [relative DNA binding activity of sample/relative DNA binding activity of the unheated control] x 100

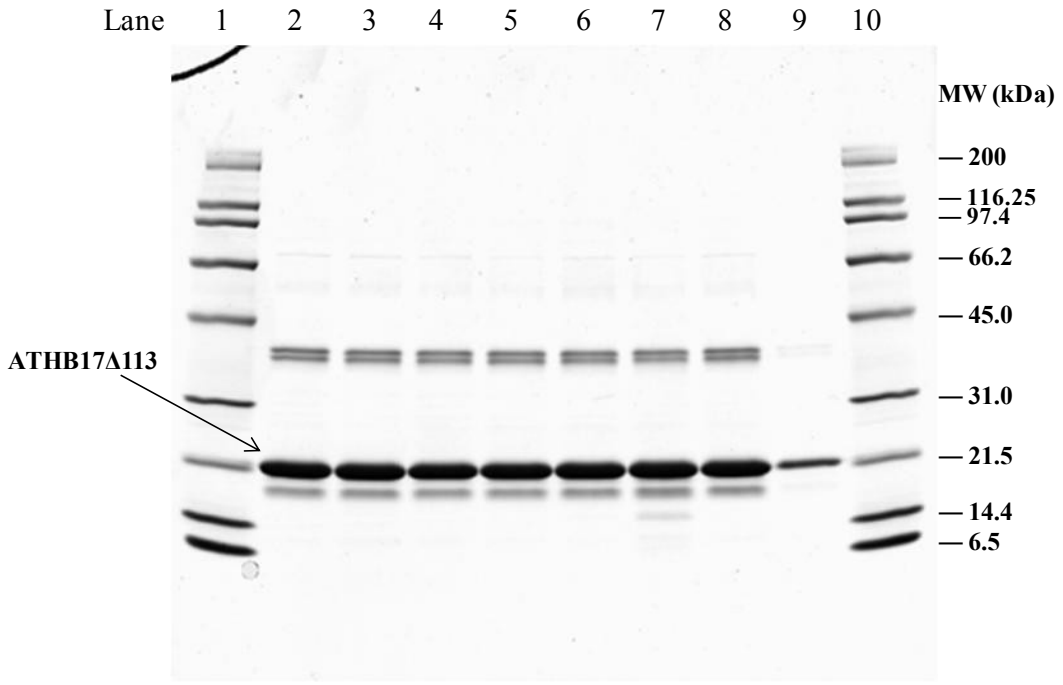


Figure 25. SDS-PAGE of *E. coli*-produced ATHB17Δ113 Protein Following Heat Treatment for 15 Minutes

Heat-treated samples of ATHB17Δ113 were subjected to SDS-PAGE under denaturing and reducing conditions. Gel was stained with Brilliant Blue G Colloidal. Approximate molecular weights (kDa) are shown on the right and correspond to molecular weight standards in lanes 1 and 10. Lane designations are as follows:

Lane	Description	Total Amount
1	Broad Range Molecular Weight Standards	4.5 µg
2	ATHB17Δ113 Protein Control Substance	2.1 µg
3	ATHB17Δ113 Protein 25 ºC for 15 min	2.1 µg
4	ATHB17Δ113 Protein 37 ºC for 15 min	2.1 µg
5	ATHB17Δ113 Protein 55 ºC for 15 min	2.1 µg
6	ATHB17Δ113 Protein 75 ºC for 15 min	2.1 µg
7	ATHB17Δ113 Protein 95 ºC for 15 min	2.1 µg
8	ATHB17Δ113 Protein Reference 100 % Equivalence	2.1 µg
9	ATHB17Δ113 Protein Reference 10 % Equivalence	0.2 µg
10	Broad Range Molecular Weight Standards	4.5 µg

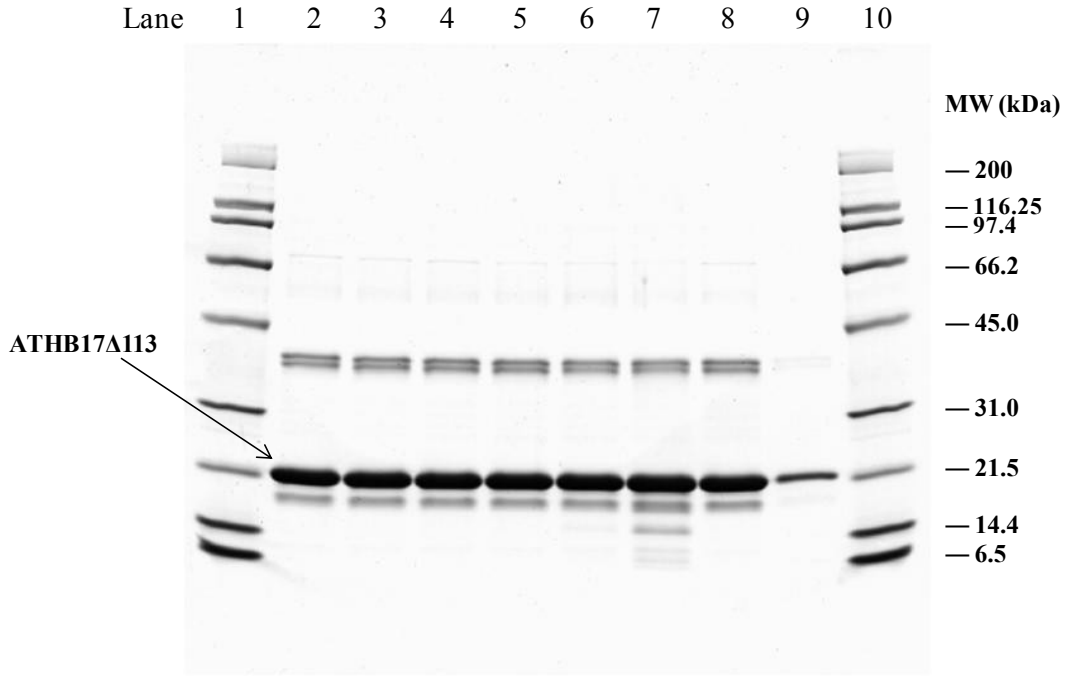


Figure 26. SDS-PAGE of *E. coli*-produced MON 87403 ATHB17Δ113 Protein Following Heat Treatment for 30 Minutes

Heat-treated samples of ATHB17Δ113 were subjected to SDS-PAGE denaturing and reducing conditions. Gel was stained with Brilliant Blue G Colloidal. Approximate molecular weights (kDa) are shown on the right and correspond to molecular weight standards in lanes 1 and 10. Lane designations are as follows:

Lane	Description	Total Amount
1	Broad Range Molecular Weight Standards	4.5 µg
2	ATHB17Δ113 Protein Control Substance	2.1 µg
3	ATHB17Δ113 Protein 25 ºC for 30 minutes	2.1 µg
4	ATHB17Δ113 Protein 37 ºC for 30 minutes	2.1 µg
5	ATHB17Δ113 Protein 55 ºC for 30 minutes	2.1 µg
6	ATHB17Δ113 Protein 75 ºC for 30 minutes	2.1 µg
7	ATHB17Δ113 Protein 95 ºC for 30 minutes	2.1 µg
8	ATHB17Δ113 Protein Reference 100 % Equivalence	2.1 µg
9	ATHB17Δ113 Protein Reference 10 % Equivalence	0.2 µg
10	Broad Range Molecular Weight Standards	4.5 µg

B4(c) Acute oral toxicity study with the ATHB17Δ113 protein

Toxicology studies are generally not considered necessary for proteins that have a history of safe use or are closely related to proteins with a history of safe use. Additional factors taken into account as part of the safety assessment include the protein's biological function, the level of exposure, and whether the protein shows similarity to proteins that are harmful to mammals (Codex Alimentarius, 2009; Delaney *et al*., 2008; Hammond *et al*., 2013). Based on the undetectable level of expression of the ATHB17Δ113 protein expressed in MON 87403 grain, and on the safety assessments conducted (Sections B3(d), and B4(a)-B4(b)), toxicology studies are not considered necessary. However, an acute toxicology study using ATHB17Δ113 was conducted to provide further confirmation of the safety of this protein.

The *E. coli*-produced ATHB17Δ113 protein was administered twice daily by oral gavage to 10 male and 10 female CD-1 mice for a total daily dose of 1335 mg/kg body wt/day. Control mice were administered a comparable dose of bovine serum albumin. Following dosing, all mice were given detailed clinical observations once daily (twice on day of dosing) for signs of mortality or toxicity. Food consumption was measured on days 0, 7, and 14. Body weights were measured prior to dosing and on study days 0, 7, and 14. All animals were sacrificed on day 14 and subjected to a gross necropsy. There were no treatment-related effects of ATHB17Δ113 on survival, clinical observations, body weight gain, food consumption or gross pathology. The No Observable Adverse Effect Level (NOAEL) for ATHB17Δ113 was therefore concluded to be 1335 mg/kg body weight/day, the highest dose tested, in males and females.

Most known protein toxins act through acute mechanisms to exert toxicity (Hammond and Fuchs, 1998; Pariza and Johnson, 2001; Sjoblad *et al*., 1992). The primary exceptions to this rule consist of certain anti-nutritional proteins such as lectins and protease inhibitors, which manifest toxicity in a short term (few weeks) feeding study (Liener, 1994). The amino acid sequence of the ATHB17Δ113 protein is not similar to any of these anti-nutritional proteins or to any other known protein toxins (Section B4(a)). Thus, an acute study was considered sufficient to evaluate the toxicity of the ATHB17Δ113 protein.

For details, please refer to [redacted], 2013 (CRO-2013-0121)

B5 Assessment of Potential Allergenicity

Assessment of potential allergenicity of an introduced protein is conducted following the guidelines adopted by the Codex Alimentarius Commission (Codex Alimentarius, 2009). A protein introduced into maize is not likely to be associated with allergenicity if: 1) the protein is from a nonallergenic source, 2) the protein lacks structural similarity to known allergens based on the amino acid sequence, and 3) the protein is susceptible to pepsin digestion. Other factors that are evaluated include the level of exposure to the introduced protein and the effects of heat treatment, which, as a surrogate for food processing conditions, can provide additional information regarding the stability of the protein in food.

B5(a) Source of introduced protein

As described in Section A2(a), the gene encoding ATHB17Δ113 is from *Arabidopsis thaliana,* a plant species that is not intentionally consumed but is phylogenetically related to commonly consumed Brassica species. Arabidopsis is not considered an allergenic source organism, and no Arabidopsis proteins have been reported in a peer-reviewed database of known allergens (FARRP, 2013). One case of occupational asthma has been reported in a laboratory worker due to exposure to Arabidopsis pollen (Yates *et al*., 2008).

B5(b) Bioinformatic comparison (aa) of novel protein(s) to allergens

The Codex guidelines for the evaluation of the allergenicity potential of introduced proteins (Codex Alimentarius, 2009) are based on the comparison of amino acid sequences between introduced proteins and allergens, where allergenic cross-reactivity may exist if the introduced protein is found to have at least 35% amino acid identity with an allergen over any segment of at least 80 amino acids. The Codex guideline also suggests that a sliding window search with a scientifically justified peptide size be used to identify immunologically relevant peptides in otherwise unrelated proteins. Therefore, the extent of sequence similarities between the ATHB17Δ113 protein sequence and known allergens, gliadins, and glutenins was assessed using the FASTA sequence alignment tool along with an eight-amino acid sliding window search (Codex Alimentarius, 2009; Thomas *et al*., 2005). The data generated from these analyses confirm that the ATHB17Δ113 protein does not share amino acid sequence similarities with known allergens, gliadins, or glutenins.

The FASTA program directly compares amino acid sequences (*i.e*., primary, linear protein structure). This alignment data may be used to infer shared higher order structural similarities between two sequences (*i.e*., secondary and tertiary protein structures). Proteins that share a high degree of similarity throughout the entire sequence are often homologous. By definition, homologous proteins have common secondary structures, and three-dimensional configuration, and, consequently, may share similar functions. The allergen, gliadin, and glutenin sequence database (AD_2013) was obtained from Food Allergy Research and Resource Program Database (FARRP, 2013) and was used for the evaluation of sequence similarities shared between the ATHB17Δ113 protein and all proteins in the database. The AD_2013 database contains 1,630 sequences. When used to align the

sequence of the introduced protein to each protein in the database, the FASTA algorithm produces an *E*-score (expectation score) for each alignment. The *E*-score is a statistical measure of the likelihood that the observed similarity score could have occurred by chance in a search. A larger *E*-score indicates a low degree of similarity between the query sequence and the sequence from the database. Typically, alignments between two sequences which have an *E*-score of less than or equal to 1×10^{-5} are considered to have meaningful homology. Results indicate that the ATHB17Δ113 protein sequence does not share meaningful similarity with sequences in the allergen database. No alignment met or exceeded the threshold of 35% identity over 80 amino acids recommended by Codex Alimentarius (2009) or had an *E*-score of less than or equal to 1×10^{-5}.

A second bioinformatic tool, an eight-amino acid sliding window search, was used to specifically identify short linear polypeptide matches to known allergens. It is possible that proteins structurally unrelated to allergens, gliadins, and glutenins may contain smaller immunologically meaningful epitopes. An amino acid sequence may have allergenic potential if it has an exact sequence identity of at least eight linearly contiguous amino acids with a potential allergen epitope (Hileman *et al*., 2002; Metcalfe *et al*., 1996). Using a sliding window of less than eight amino acids can produce matches containing considerable uncertainty depending on the length of the query sequence (Silvanovich *et al*., 2006), and is not useful to the allergy assessment process (Thomas *et al*., 2005). No eight contiguous amino acid identities were detected when the ATHB17Δ113 protein sequence was compared to the proteins in the AD_2013 sequence database.

The bioinformatic results demonstrated there were no biologically relevant sequence similarities to allergens when the ATHB17Δ113 protein sequence was used as a query for a FASTA search of the AD_2013 database. Furthermore, no short (eight amino acid) polypeptide matches were shared between the ATHB17Δ113 protein sequence and proteins in the allergen database. These data show that ATHB17Δ113 protein sequence lacks both structurally and immunologically relevant similarities to known allergens, gliadins, and glutenins.

For details, please refer to [redacted], 2013 (MSL0025242).

B5(c) Structural properties, including digestion by pepsin, heat treatment

B5(c)(i) Digestibility of proteins

Proteins are an essential dietary component for humans and animals, and most are rapidly degraded to the component amino acids for nutritional purposes and, therefore, are not toxic when ingested (Hammond and Jez, 2011). The enzymatic degradation of an ingested protein by exposure to gastric pepsin and/or intestinal pancreatic proteases makes it highly unlikely that either the intact protein or protein fragment(s) will retain any functionality as a toxin in food or feed. Furthermore, ATHB17Δ113 is undetectable in maize grain and, therefore, is a highly unlikely risk of allergenicity due to the extremely low exposure anticipated for food and feed uses of MON 87403. Nonetheless, the degradation of

ATHB17Δ113 was evaluated by incubation with gastric pepsin and intestinal pancreatic enzymes, and the results show that ATHB17Δ113 was readily digested by both proteolytic conditions (Section B4(b)).

B5(c)(2) Heat stability of the ATHB17Δ113 protein

Although the exposure to ATHB17Δ113 is highly unlikely due to undetectable levels of this protein, the effect of heat treatment on the activity of ATHB17Δ113 protein was evaluated using a DNA binding assay, and SDS-PAGE was used to assess the impact of temperature on protein integrity. The results show that *E. coli*-produced ATHB17Δ113 protein behaves with a predictable tendency toward loss of functional activity at elevated temperatures (Section B4(c)). This tendency, combined with the undetectably low ATHB17Δ113 protein levels in grain (Section B3(c)), indicates that exposure to functionally active ATHB17Δ113 protein from the consumption of MON 87403 or foods derived from MON 87403 will be negligible.

B5(d) Specific serum screening if protein from allergenic source

Not applicable.

B5(e) Protein as a proportion of total protein

The ATHB17Δ113 protein has very low expression levels detectable in most of the plant tissues assayed, and was undetectable in grain (Table 12). Since maize enters the human diet only through products derived from maize grain, the lack of detection of ATHB17Δ113 in MON 87403 grain makes exposure to the protein immeasurably low. However, in order to be able to provide a conservative estimate of the percent of ATHB17Δ113 protein in the total grain protein of MON 87403, the LOD of the ELISA method used to detect ATHB17Δ113 in grain, 0.00028 µg/g dw (Table 12), was used as a theoretical expression value for the purposes of this calculation. This value was not converted from dry weight to fresh weight as the protein was not detectable in grain, therefore the dry weight LOD provides a more conservative value. Total protein as a percentage of dry weight in harvested grain from MON 87403 is 10.13% (or 101300 µg/g, Table 17). The theoretical percentage of ATHB17Δ113 protein in total protein of MON 87403 grain is calculated as follows:

(0.00028 µg/g ÷ 101300 µg/g) × 100% ≈ 0.0000003% of total maize protein

This value represents the maximum percentage of ATHB17Δ113 protein in the total protein of MON 87403 grain because the actual concentration of ATHB17Δ113 protein in MON 87403 grain was below the LOD. Thus, the MON 87403-produced ATHB17Δ113 protein represents an extremely small portion of the total MON 87403 grain protein.

B6 Toxicity of Novel Herbicide Metabolites in GM Herbicide-Tolerant Plants

Not Applicable

B7 Compositional Assessment

Safety assessments of biotechnology-derived crops follow the comparative safety assessment process (Codex Alimentarius, 2009) in which the composition of grain and/or other raw agricultural commodities of the biotechnology-derived crop are compared to the appropriate conventional control that has a history of safe use. Compositional assessments are performed using the principles and analytes outlined in the OECD consensus document for maize composition (OECD, 2002a).

A recent review of compositional assessments conducted according to OECD guidelines that encompassed a total of seven biotechnology-derived crop varieties, nine countries and eleven growing seasons concluded that incorporation of biotechnology-derived agronomic traits has had little impact on crop composition compared to natural variation. Most compositional variation is attributable to growing region, agronomic practices, and genetic background (Harrigan *et al*., 2010). Numerous scientific publications have further documented the extensive variability in the concentrations of crop nutrients and anti-nutrients and secondary metabolites that reflect the influence of environmental and genetic factors as well as extensive conventional breeding efforts to improve nutrition, agronomics, and yield (Harrigan *et al*., 2010; Harrigan *et al*., 2009; Ridley *et al*., 2011; Zhou *et al*., 2011).

Compositional equivalence between biotechnology-derived and conventional crops supports an "equal or increased assurance of the safety of foods derived from genetically modified plants" (OECD, 2002b). OECD consensus documents on compositional considerations for new crop varieties emphasize quantitative measurements of essential nutrients and known anti-nutrients. These quantitative measurements effectively discern any compositional changes that imply potential nutritional or safety (e.g., anti-nutritional) concerns. Levels of the components in grain and/or other raw agricultural commodities of the biotechnology-derived crop are compared to: 1) corresponding levels in a conventional comparator, a genetically similar conventional line, grown concurrently, under similar field conditions, and 2) natural ranges generated from an evaluation of commercial reference hybrids grown concurrently and from data published in the scientific literature. The comparison to data published in the literature places any potential differences between the assessed crop and its comparator in the context of the well-documented variation in the concentrations of crop nutrients, anti-nutrients, and secondary metabolites.

This section provides analyses of concentrations of key nutrients, anti-nutrients, and secondary metabolites of MON 87403 compared to that of a conventional control maize hybrid grown and harvested under similar conditions. In addition, conventional commercial reference maize hybrids were included in the composition analyses to provide additional information on the range of natural variability for each component. The production of materials for the compositional analyses used sufficient variety of field trial sites, robust field designs (randomized complete block design with four blocks), and sensitive analytical methods that allow accurate assessments of compositional characteristics over a range of environmental conditions under which MON 87403 is expected to be grown.

The information provided in this section also addresses the relevant factors in Codex Plant Guidelines, Section 4, paragraphs 44 and 45 for compositional analyses (Codex Alimentarius, 2009).

Compositional Equivalence of MON 87403 Grain and Forage to Conventional Maize

Grain and forage samples were collected from MON 87403, a conventional control maize, and a total of 17 different reference hybrids grown in the U.S. during a 2012 field production. The reference hybrids were included in the composition analyses to provide data on the natural variability for each component. The field production was conducted at eight sites. The field sites were planted in a randomized complete block design with four blocks per site. MON 87403, the conventional control, and reference hybrids were grown in areas of the U.S. that were typical for maize production and under normal agronomic field conditions for their respective geographic regions.

The evaluation of MON 87403 followed considerations relevant to the compositional quality of maize as defined by the OECD consensus document (OECD, 2002a). Grain samples were analyzed for levels of nutrients including proximates, carbohydrates by calculation, fiber, amino acids, fatty acids, minerals, and vitamins. The anti-nutrients analyzed in grain included phytic acid and raffinose. Secondary metabolites analyzed in grain included furfural, ferulic acid, and p-coumaric acid. Forage samples were analyzed for levels of proximates, carbohydrates by calculation, fiber, and minerals. In total, 78 different components were assayed (nine in forage and 69 in grain).

Of those 78 components, 14 fatty acids (caprylic, capric, lauric, myristic, myristoleic, pentadecanoic, pentadecenoic, palmitoleic, heptadecanoic, heptadecenoic, gamma linolenic, eicosadienoic, eicosatrienoic, and arachidonic acids), sodium, and furfural had more than 50% of observations below the assay limit of quantitation (LOQ) and were excluded from statistical analysis. Moisture for grain and forage was measured for conversion of components to dry weight, but was not statistically analyzed. Therefore, 60 components were statistically analyzed.

The statistical comparison of MON 87403 and the conventional control was based on compositional data combined across all field sites. Statistically significant differences were identified at the 5% level ($\alpha = 0.05$). The compositional data from the reference hybrids were combined across all field sites to calculate a 99% tolerance interval for each component to estimate the natural variability of each component in maize.

A statistically significant difference between MON 87403 and the conventional control does not necessarily imply biological relevance from a food and feed safety perspective. Therefore, statistically significant differences observed are typically evaluated in the context of natural variability to determine whether a detected difference indicates a biologically relevant compositional change. However, in this study, no significant differences between MON 87403 and the conventional control were observed in any of the measured components.

Measurements of grain nutrients, including protein, amino acids (18 components), total fat, fatty acids (22 components), carbohydrates by calculation, fiber (3 components), ash, minerals (9 components), vitamins (7 components), anti-nutrients (phytic acid and raffinose), and secondary metabolites (furfural, ferulic acid, and p-coumaric acid) demonstrated that values in MON 87403 were not statistically significantly different from those in the conventional control (Table 17 to Table 23). Forage component levels, including ash, protein, total fat, carbohydrates by calculation, fiber (ADF and NDF), and minerals (calcium and phosphorus) were also not statistically significantly different between MON 87403 and the conventional control (Table 23). International Life Sciences Institute Crop Composition Database (ILSI-CCDB) and published literature values for all analytes are provided in Table 24.

The compositional analysis provided a comprehensive comparative assessment of the levels of key nutrients, anti-nutrients, and secondary metabolites in maize grain and forage of MON 87403 and the conventional control. The lack of any statistically significant differences between MON 87403 and the conventional control demonstrated that MON 87403 was not a major contributor to variation in nutrient, anti-nutrient, or secondary metabolite component levels in maize grain or forage and confirmed the compositional equivalence of MON 87403 to the conventional control in levels of these components.

For details, please refer to [redacted]., 2015 (MSL0025787).

Table 17. Summary of Maize Grain Protein and Amino Acids for MON 87403, Conventional Control, and Conventional References

					Difference (Test minus Control)		
Component (% dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Mean (S.E.)	p-Value	% Relative[6]
Protein	10.13 (0.34) 8.41 - 11.90	10.15 (0.34) 8.49 - 12.50	(8.58 - 12.63) 7.72, 12.67	4.01	-0.016 (0.14)	0.911	-0.16
Alanine	0.77 (0.030) 0.63 - 0.91	0.77 (0.030) 0.62 - 0.98	(0.63 - 0.99) 0.55, 1.01	0.36	0.00036 (0.012)	0.975	0.05
Arginine	0.48 (0.0093) 0.41 - 0.53	0.48 (0.0093) 0.41 - 0.54	(0.43 - 0.61) 0.41, 0.59	0.13	-0.00007 (0.0051)	0.988	-0.01
Aspartic Acid	0.63 (0.018) 0.54 - 0.72	0.63 (0.018) 0.52 - 0.74	(0.56 - 0.79) 0.52, 0.78	0.21	0.00058 (0.0080)	0.943	0.09
Cystine/Cysteine	0.21 (0.0041) 0.18 - 0.23	0.21 (0.0041) 0.18 - 0.24	(0.17 - 0.26) 0.16, 0.25	0.06	0.0017 (0.0032)	0.589	0.83
Glutamic Acid	1.85 (0.077) 1.48 - 2.26	1.86 (0.077) 1.43 - 2.34	(1.50 - 2.47) 1.26, 2.52	0.91	-0.0056 (0.033)	0.870	-0.30

Table 17. Summary of Maize Grain Protein and Amino Acids for MON 87403, Conventional Control, and Conventional References (continued)

					Difference (Test minus Control)		
Component (% dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Mean (S.E.)	p-Value	% Relative[6]
Glycine	0.37 (0.0065) 0.32 - 0.41	0.37 (0.0065) 0.33 - 0.42	(0.32 - 0.44) 0.32, 0.44	0.08	0.00016 (0.0036)	0.966	0.04
Histidine	0.28 (0.0068) 0.24 - 0.33	0.28 (0.0068) 0.24 - 0.33	(0.22 - 0.35) 0.20, 0.35	0.09	0.0016 (0.0034)	0.648	0.55
Isoleucine	0.36 (0.014) 0.29 - 0.43	0.36 (0.014) 0.28 - 0.46	(0.28 - 0.44) 0.26, 0.46	0.17	-0.00044 (0.0061)	0.944	-0.12
Leucine	1.27 (0.059) 0.98 - 1.58	1.27 (0.059) 0.93 - 1.64	(0.98 - 1.65) 0.81, 1.73	0.71	0.0015 (0.026)	0.956	0.12
Lysine	0.27 (0.0048) 0.24 - 0.30	0.27 (0.0048) 0.23 - 0.31	(0.23 - 0.33) 0.24, 0.31	0.07	0.0013 (0.0027)	0.625	0.50
Methionine	0.21 (0.0063) 0.17 - 0.23	0.20 (0.0063) 0.17 - 0.25	(0.16 - 0.26) 0.15, 0.26	0.08	0.00071 (0.0028)	0.803	0.35

Table 17. Summary of Maize Grain Protein and Amino Acids for MON 87403, Conventional Control, and Conventional References (continued)

					Difference (Test minus Control)		
Component (% dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Mean (S.E.)	p-Value	% Relative[6]
Phenylalanine	0.52 (0.022) 0.41 - 0.63	0.52 (0.022) 0.40 - 0.64	(0.42 - 0.65) 0.36, 0.68	0.24	0.0024 (0.0093)	0.801	0.47
Proline	0.95 (0.030) 0.80 - 1.10	0.95 (0.030) 0.76 - 1.18	(0.73 - 1.12) 0.64, 1.17	0.43	0.0032 (0.015)	0.830	0.34
Serine	0.44 (0.015) 0.35 - 0.51	0.45 (0.015) 0.37 - 0.55	(0.37 - 0.58) 0.34, 0.57	0.18	-0.0068 (0.0072)	0.356	-1.51
Threonine	0.36 (0.0098) 0.30 - 0.40	0.35 (0.0098) 0.30 - 0.41	(0.31 - 0.44) 0.28, 0.43	0.11	0.0025 (0.0043)	0.570	0.69
Tryptophan	0.077 (0.0015) 0.063 - 0.086	0.077 (0.0015) 0.069 - 0.084	(0.067 - 0.096) 0.064, 0.093	0.01	-0.00019 (0.00085)	0.827	-0.25
Tyrosine	0.41 (0.016) 0.32 - 0.48	0.41 (0.016) 0.32 - 0.52	(0.34 - 0.55) 0.30, 0.53	0.20	0.00043 (0.0081)	0.959	0.10

Table 17. Summary of Maize Grain Protein and Amino Acids for MON 87403, Conventional Control, and Conventional References (continued)

					Difference (Test minus Control)		
Component (% dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Mean (S.E.)	p-Value	% Relative[6]
Valine	0.46 (0.013)	0.46 (0.013)	(0.38 - 0.57)	0.17	-0.00092 (0.0061)	0.880	-0.20
	0.38 - 0.52	0.39 - 0.56	0.35, 0.57				

[1]dwt = dry weight.

[2]Mean (S.E.) = least-square mean (standard error).

[3]Range is the minimum and maximum raw values for the conventional reference maize hybrids.

[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference maize hybrids.
Negative limits set to zero.

[5]Maximum value minus minimum value for the control maize hybrid.

[6]The relative magnitude of the difference in mean values between MON 87403 and the control, expressed as a percent of the control.

Table 18. Summary of Maize Grain Total Fat and Fatty Acids for MON 87403, Conventional Control, and Conventional References

Component	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (Test minus Control) Mean (S.E.)	Difference (Test minus Control) p-Value
Total Fat (% dwt)[1]	3.56 (0.072) 3.10 - 3.90	3.54 (0.072) 3.11 - 4.09	(2.49 - 4.70) 1.93, 5.49	0.98	0.026 (0.060)	0.682
16:0 Palmitic[6]	12.77 (0.14) 12.27 - 13.50	12.68 (0.14) 10.59 - 13.59	(9.90 - 13.12) 7.90, 14.94	2.99	0.097 (0.072)	0.219
18:0 Stearic	2.04 (0.022) 1.87 - 2.14	2.06 (0.022) 1.82 - 2.28	(1.46 - 2.42) 1.05, 2.72	0.46	-0.025 (0.019)	0.206
18:1 Oleic	21.84 (0.30) 20.77 - 23.46	21.77 (0.30) 20.50 - 28.08	(21.13 - 34.04) 14.77, 38.25	7.59	0.070 (0.19)	0.716
18:2 Linoleic	61.23 (0.30) 59.60 - 62.78	61.39 (0.30) 57.12 - 63.36	(51.19 - 62.88) 45.88, 70.60	6.25	-0.15 (0.16)	0.372
18:3 Linolenic	1.31 (0.014) 1.22 - 1.40	1.30 (0.014) 1.16 - 1.40	(0.78 - 1.48) 0.69, 1.67	0.24	0.012 (0.010)	0.267

Table 18. Summary of Maize Grain Total Fat and Fatty Acids for MON 87403, Conventional Control, and Conventional References (continued)

Component	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (Test minus Control) Mean (S.E.)	Difference (Test minus Control) p-Value
20:0 Arachidic	0.45 (0.012) 0.41 - 0.52	0.45 (0.012) 0.40 - 0.54	(0.30 - 0.52) 0.26, 0.55	0.14	-0.0017 (0.0034)	0.634
20:1 Eicosenoic	0.20 (0.0030) 0.18 - 0.22	0.20 (0.0030) 0.18 - 0.23	(0.19 - 0.30) 0.14, 0.33	0.05	-0.0020 (0.0017)	0.243
22:0 Behenic	0.16 (0.0081) 0.062 - 0.21	0.16 (0.0081) 0.069 - 0.20	(0.055 - 0.23) 0, 0.26	0.13	0.0029 (0.0045)	0.544

[1]dwt = dry weight.

[2]Mean (S.E.) = least-square mean (standard error).

[3]Range is the minimum and maximum raw values for the conventional reference maize hybrids.

[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference maize hybrids. Negative limits set to zero.

[5]Maximum value minus minimum value for the control maize hybrid.

[6]Expressed as % total fatty acid. Prefix numbers refer to number of carbon atoms and number of carbon-carbon double bonds in the fatty acid molecule; 16:0 means sixteen carbon atoms and zero double bonds. Numbers are not included in text discussion for reasons of clarity. The following fatty acids with more than 50% of observations below the assay LOQ were excluded from statistical analysis: caprylic acid, capric acid, lauric acid, myristic acid, myristoleic acid, pentadecanoic acid, pentadecenoic acid, palmitoleic acid, heptadecanoic acid, heptadecenoic acid, gamma linolenic acid, eicosadienoic acid, eicosatrienoic acid, and arachidonic acid.

Table 19. Summary of Maize Grain Carbohydrates by Calculation and Fiber for MON 87403, Conventional Control, and Conventional References

					Difference (Test minus Control)	
Component (% dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Mean (S.E.)	p-Value
Carbohydrates by	84.98 (0.34)	84.98 (0.34)	(82.55 - 86.97)	3.84	0.0021 (0.12)	0.986
Calculation	83.09 - 86.73	82.77 - 86.61	81.80, 87.71			
Acid Detergent Fiber	3.63 (0.068)	3.62 (0.068)	(2.52 - 4.42)	1.49	0.0087 (0.062)	0.893
	3.16 - 4.07	2.91 - 4.40	2.36, 4.43			
Neutral Detergent Fiber	9.47 (0.19)	9.41 (0.19)	(6.86 - 12.18)	3.48	0.063 (0.16)	0.696
	8.11 - 10.67	7.53 - 11.01	5.32, 12.85			
Total Dietary Fiber	13.04 (0.15)	12.95 (0.15)	(9.83 - 17.30)	4.19	0.098 (0.22)	0.657
	11.53 - 14.70	11.50 - 15.69	10.05, 15.51			

[1]dwt = dry weight.

[2]Mean (S.E.) = least-square mean (standard error).

[3]Range is the minimum and maximum raw values for the conventional reference maize hybrids.

[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference maize hybrids. Negative limits set to zero.

[5]Maximum value minus minimum value for the control maize hybrid.

Table 20. Summary of Maize Grain Ash and Minerals for MON 87403, Conventional Control, and Conventional References

Component	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (Test minus Control) Mean (S.E.)	Difference (Test minus Control) p-Value
Ash (% dwt)[1]	1.31 (0.026) 1.09 - 1.52	1.33 (0.026) 1.10 - 1.51	(1.02 - 1.53) 1.08, 1.60	0.42	-0.017 (0.019)	0.371
Calcium (% dwt)	0.0037 (0.00021) 0.0026 - 0.0049	0.0037 (0.00021) 0.0025 - 0.0049	(0.0011 - 0.0058) 0.0011, 0.0059	0.00	0.00006 (0.00006)	0.321
Copper (mg/kg dwt)	1.57 (0.11) 1.13 - 3.35	1.55 (0.11) 1.04 - 2.91	(0.92 - 6.11) 0.29, 3.17	1.87	0.023 (0.12)	0.852
Iron (mg/kg dwt)	19.33 (0.82) 15.60 - 24.15	19.60 (0.82) 15.81 - 23.90	(14.66 - 25.54) 10.87, 27.03	8.09	-0.27 (0.21)	0.240
Magnesium (% dwt)	0.12 (0.0033) 0.10 - 0.14	0.12 (0.0033) 0.096 - 0.14	(0.093 - 0.14) 0.092, 0.15	0.04	-0.00010 (0.0013)	0.939
Manganese (mg/kg dwt)	6.17 (0.31) 4.36 - 7.85	6.14 (0.31) 4.55 - 8.55	(4.02 - 9.46) 2.59, 10.23	3.99	0.029 (0.11)	0.787

Table 20. Summary of Maize Grain Ash and Minerals for MON 87403, Conventional Control, and Conventional References (continued)

Component	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (Test minus Control) Mean (S.E.)	Difference (Test minus Control) p-Value
Phosphorus (% dwt)	0.30 (0.0068) 0.25 - 0.33	0.30 (0.0068) 0.24 - 0.34	(0.24 - 0.36) 0.24, 0.38	0.10	-0.0017 (0.0041)	0.683
Potassium (% dwt)	0.33 (0.0076) 0.30 - 0.40	0.33 (0.0076) 0.30 - 0.40	(0.27 - 0.42) 0.23, 0.42	0.10	-0.0023 (0.0044)	0.611
Zinc (mg/kg dwt)	20.52 (0.83) 16.30 - 25.20	20.98 (0.83) 16.49 - 27.14	(15.56 - 30.10) 9.09, 32.95	10.65	-0.46 (0.30)	0.165

[1]dwt = dry weight.

[2]Mean (S.E.) = least-square mean (standard error).

[3]Range is the minimum and maximum raw values for the conventional reference maize hybrids.

[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference maize hybrids. Negative limits set to zero.

[5]Maximum value minus minimum value for the control maize hybrid.

Table 21. Summary of Maize Grain Vitamins for MON 87403, Conventional Control, and Conventional References

Component (mg/kg dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (Test minus Control) Mean (S.E.)	p-Value
Vitamin A (β-Carotene)	1.16 (0.039) 0.89 - 1.37	1.14 (0.039) 0.88 - 1.37	(0.48 - 2.85) 0, 3.10	0.49	0.025 (0.023)	0.319
Vitamin B1 (Thiamin)	3.48 (0.13) 2.72 - 4.19	3.52 (0.13) 2.82 - 4.23	(2.54 - 4.99) 1.73, 5.12	1.41	-0.046 (0.042)	0.277
Vitamin B2 (Riboflavin)	1.83 (0.059) 1.48 - 2.43	1.71 (0.059) 1.22 - 2.22	(1.35 - 2.35) 1.25, 2.22	1.00	0.12 (0.053)	0.057
Vitamin B3 (Niacin)	16.47 (0.78) 12.74 - 21.74	16.59 (0.78) 12.61 - 26.88	(12.77 - 30.15) 7.36, 30.18	14.27	-0.12 (0.46)	0.788
Vitamin B6 (Pyridoxine)	7.11 (0.17) 5.50 - 10.63	6.89 (0.17) 5.85 - 9.36	(4.66 - 8.80) 4.51, 8.98	3.51	0.22 (0.19)	0.255
Vitamin B9 (Folic Acid)	0.39 (0.016) 0.27 - 0.50	0.39 (0.016) 0.28 - 0.47	(0.22 - 0.77) 0.038, 0.69	0.20	0.00065 (0.0086)	0.940

Table 21. Summary of Maize Grain Vitamins for MON 87403, Conventional Control, and Conventional References (continued)

					Difference (Test minus Control)	
Component (mg/kg dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Mean (S.E.)	p-Value
Vitamin E (α-Tocopherol)	17.87 (0.65)	18.33 (0.65)	(8.68 - 25.90)	8.20	-0.47 (0.24)	0.095
	14.70 - 20.76	15.03 - 23.23	2.50, 27.12			

[1]dwt = dry weight.

[2]Mean (S.E.) = least-square mean (standard error).

[3]Range is the minimum and maximum raw values for the conventional reference maize hybrids.

[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference maize hybrids. Negative limits set to zero.

[5]Maximum value minus minimum value for the control maize hybrid.

Table 22. Summary of Maize Grain Anti-nutrients and Secondary Metabolites for MON 87403, Conventional Control, and Conventional References

Component	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (Test minus Control) Mean (S.E.)	Difference (Test minus Control) p-Value
Anti-nutrient (% dwt[1])						
Phytic Acid	0.89 (0.017) 0.70 - 1.04	0.87 (0.017) 0.67 - 1.07	(0.68 - 1.15) 0.68, 1.18	0.40	0.023 (0.016)	0.164
Raffinose	0.22 (0.012) 0.14 - 0.30	0.23 (0.012) 0.17 - 0.31	(0.062 - 0.35) 0.00088, 0.40	0.14	-0.0030 (0.0041)	0.487
Secondary Metabolite (µg/g dwt)						
Ferulic Acid	2262.60 (61.27) 1833.71 - 2587.97	2213.54 (61.27) 1344.15 - 2694.48	(1381.65 - 2990.97) 827.07, 3473.40	1350.33	49.06 (48.06)	0.341
p-Coumaric Acid	216.73 (5.25) 160.59 - 242.25	212.58 (5.25) 173.81 - 260.72	(103.35 - 383.41) 6.62, 433.65	86.91	4.15 (5.09)	0.441

[1]dwt = dry weight.

[2]Mean (S.E.) = least-square mean (standard error).

[3]Range is the minimum and maximum raw values for the conventional reference maize hybrids.

[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference maize hybrids. Negative limits set to zero.

[5]Maximum value minus minimum value for the control maize hybrid.

Table 23. Summary of Maize Forage Proximates, Fiber, and Minerals for MON 87403, Conventional Control, and Conventional References

Component (% dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (Test minus Control) Mean (S.E.)	Difference (Test minus Control) p-Value
Ash	4.00 (0.35) 2.17 - 5.87	4.14 (0.35) 2.14 - 6.24	(2.05 - 7.98) 0.67, 7.56	4.10	-0.14 (0.11)	0.199
Carbohydrates by Calculation	86.98 (0.43) 84.01 - 90.21	86.76 (0.43) 83.73 - 89.74	(82.09 - 90.15) 81.04, 92.37	6.01	0.22 (0.28)	0.449
Protein	7.08 (0.27) 5.60 - 8.92	7.19 (0.27) 4.42 - 9.34	(4.14 - 10.27) 3.56, 10.69	4.92	-0.12 (0.19)	0.548
Total Fat	1.96 (0.12) 0.92 - 2.89	1.98 (0.12) 0.53 - 2.94	(0.62 - 3.18) 0.81, 3.33	2.41	-0.023 (0.12)	0.851
Acid Detergent Fiber	23.01 (0.89) 16.18 - 33.12	22.44 (0.89) 17.33 - 30.17	(16.01 - 37.25) 17.89, 28.94	12.85	0.57 (0.83)	0.513
Neutral Detergent Fiber	36.83 (1.06) 28.67 - 44.62	37.23 (1.06) 29.77 - 56.33	(27.09 - 54.66) 30.85, 44.85	26.57	-0.40 (1.33)	0.768

Table 23. Summary of Maize Forage Proximates, Fiber, and Minerals for MON 87403, Conventional Control, and Conventional References (continued)

					Difference (Test minus Control)	
Component (% dwt)[1]	MON 87403 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Mean (S.E.)	p-Value
Calcium	0.22 (0.015) 0.12 - 0.34	0.23 (0.015) 0.16 - 0.33	(0.12 - 0.40) 0.10, 0.36	0.17	-0.013 (0.0083)	0.150
Phosphorus	0.16 (0.0058) 0.12 - 0.22	0.16 (0.0058) 0.13 - 0.22	(0.10 - 0.30) 0.045, 0.30	0.09	0.0011 (0.0049)	0.828

[1]dwt = dry weight.

[2]Mean (S.E.) = least-square mean (standard error).

[3]Range is the minimum and maximum raw values for the conventional reference maize hybrids.

[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference maize hybrids. Negative limits set to zero.

[5]Maximum value minus minimum value for the control maize hybrid.

Table 24. Literature and ILSI Ranges for Components in Maize Forage and Grain

Grain Tissue Components[1]	Literature Range[2]	ILSI Range[3]
Grain Nutrients		
Proximates (% dwt)		
Ash	1.17 – 2.01[a]; 1.27 – 1.63[b]	0.616 – 6.282
Carbohydrates by calculation	81.31 – 87.06[a]; 82.10 – 85.98[b]	77.4 – 89.5
Fat, total	2.95 – 4.40[a]; 3.18 – 4.23[b]	1.742 – 5.900
Protein	8.27 – 13.33[a]; 9.17 – 12.19[b]	6.15 – 17.26
Fiber (% dwt)		
Acid detergent fiber	1.82 – 4.48[a]; 1.83 – 3.39[b]	1.82 – 11.34
Neutral detergent fiber	6.51 –12.28[a]; 6.08 – 10.36[b]	5.59 – 22.64
Total dietary fiber	10.65 – 16.26[a]; 10.57 – 14.56[b]	9.01 – 35.31
Amino Acids (% dwt)		
Alanine	0.60 – 1.04[a]; 0.68 – 0.96[b]	0.44 - 1.39
Arginine	0.34 – 0.52[a]; 0.34 – 0.50[b]	0.12 - 0.64
Aspartic acid	0.52 – 0.78[a]; 0.59 – 0.76[b]	0.33 – 1.21
Cystine	0.19 – 0.26[a]; 0.20 – 0.26[b]	0.13 – 0.51
Glutamic acid	1.54 – 2.67[a]; 1.71 – 2.44[b]	0.97 – 3.54
Glycine	0.33 – 0.43[a]; 0.33 – 0.42[b]	0.18 – 0.54
Histidine	0.25 – 0.37[a]; 0.27 – 0.34[b]	0.14 – 0.43
Isoleucine	0.30 – 0.48[a]; 0.32 – 0.44[b]	0.18 – 0.69
Leucine	1.02 – 1.87[a]; 1.13 – 1.65[b]	0.64 – 2.49
Lysine	0.26 – 0.33[a]; 0.28 – 0.31[b]	0.17 – 0.67
Methionine	0.17 – 0.26[a]; 0.16 – 0.30[b]	0.12 – 0.47
Phenylalanine	0.43 – 0.72[a]; 0.45 – 0.63[b]	0.24 – 0.93
Proline	0.74 – 1.21[a]; 0.78 – 1.11[b]	0.46 – 1.63
Serine	0.39 – 0.67[a]; 0.43 – 0.60[b]	0.24 – 0.77
Threonine	0.29 – 0.45[a]; 0.31 – 0.39[b]	0.22 – 0.67
Tryptophan	0.047 – 0.085[a]; 0.042 – 0.070[b]	0.027 – 0.215
Tyrosine	0.13 – 0.43[a]; 0.12 – 0.41[b]	0.10 – 0.64
Valine	0.42 – 0.62[a]; 0.45 – 0.58[b]	0.27 – 0.86
Fatty Acids (% Total FA)		
16:0 Palmitic	8.80 – 13.33[a]; 9.84 – 12.33[b]	7.94 – 20.71
18:0 Stearic	1.36 – 2.14 [a]; 1.30 – 2.10[b]	1.02 – 3.40
18:1 Oleic	19.50 – 33.71[a]; 19.59 – 29.13[b]	17.4 – 40.2
18:2 Linoleic	49.31 – 64.70[a]; 56.51 – 65.65[b]	36.2 – 66.5
18:3 Linolenic	0.89 – 1.56[a]; 1.03 – 1.38[b]	0.57 – 2.25
20:0 Arachidic	0.30 – 0.49[a]; 0.30 – 0.41[b]	0.279 – 0.965
20:1 Eicosenoic	0.17 – 0.29[a]; 0.17 – 0.27[b]	0.170 – 1.917
22:0 Behenic	0.069 – 0.28[a]; 0.059 – 0.18[b]	0.110 – 0.349

Table 22. Literature and ILSI Database Ranges for Components in Maize Forage and Grain (continued)

Grain Tissue Components[1]	Literature Range[2]	ILSI Range[3]
Minerals		
Calcium (% dwt)	0.0030 – 0.0083[c]	0.00127 – 0.02084
Copper (mg/kg dwt)	0.85 – 3.54[c]	0.73 – 18.50
Iron (mg/kg dwt)	10.58 – 30.65	10.42 – 49.07
Magnesium (% dwt)	0.085 – 0.15[c]	0.0594 – 0.194
Manganese (mg/kg dwt)	3.67 – 9.39[c]	1.69 – 14.30
Phosphorous (% dwt)	0.25 – 0.38[c]	0.147 – 0.533
Potassium (% dwt)	0.29 – 0.47[c]	0.181 – 0.603
Zinc (mg/kg dwt)	16.67 – 31.38[c]	6.5 – 37.2
Vitamins (mg/kg dwt)		
Folic acid	0.19 – 0.35[a]; 0.23 – 0.42[b]	0.147 – 1.464
Vitamin A [β–Carotene]	0.122-4.740	0.19 – 46.81
Vitamin B_1 [Thiamine]	2.33 – 4.17[a]; 2.71 – 4.33[b]	1.26 – 40.00
Vitamin B_2 [Riboflavin]	0.94 – 2.42[a]; 1.64 – 2.81[b]	0.50 – 2.36
Vitamin B_3 [Niacin]	15.07 – 32.38[a]; 13.64 – 42.06[b]	10.37 – 46.94
Vitamin B_6 [Pyridoxine]	4.93 – 7.53[a]; 4.97 – 8.27[b]	3.68 – 11.32
Vitamin E [α–Tocopherol]	5.96 – 18.44[a]; 2.84 – 15.53[b]	1.537 – 68.672
Grain Anti–Nutrients (% dwt)		
Phytic acid	0.69 – 1.09[a]; 0.60 – 0.94[b]	0.111 – 1.570
Raffinose	0.079 – 0.22[a]; 0.061 – 0.15[b]	0.020 – 0.320
Grain Secondary Metabolites (µg/g dwt)		
Ferulic acid	1205.75 – 2873.05[a]; 1011.40 – 2539.86[b]	291.9 – 3885.8
p–Coumaric acid	94.77 – 327.39[a]; 66.48 – 259.68[b]	53.4 – 576.2

Forage Tissue Components[1]	Literature Range[2]	ILSI Range[3]
Forage Nutrients		
Proximates (% dwt)		
Ash	2.67 – 8.01[a]; 4.59 – 6.90[b]	1.527 – 9.638
Carbohydrates by calculation	81.88 – 89.26[a]; 84.11 – 87.54[b]	76.4 – 92.1
Fat, total	1.28 – 3.62[a]; 0.20 – 1.76[b]	0.296 – 4.570
Protein	5.80 – 10.24[a]; 5.56 – 9.14[b]	3.14 – 11.57
Fiber (% dwt)		
Acid detergent fiber	19.11 – 30.49[a]; 20.73 – 33.39[b]	16.13 – 47.39
Neutral detergent fiber	27.73 – 49.62[a]; 31.81 – 50.61[b]	20.29 – 63.71
Minerals (% dwt)		
Calcium	0.12 – 0.33[a]; 0.21 – 0.41[b]	0.07139 – 0.57679
Phosphorous	0.090 – 0.26[a]; 0.13 – 0.21[b]	0.09362 – 0.37041

[1]dwt=dry weight; FA = fatty acids.
[2]Literature range references: [a]US and [b]Chile (Harrigan *et al*., 2009), [c]France (Ridley *et al*., 2011), [d](Safawo *et al*., 2010).
[3]ILSI range is from ILSI Crop Composition Database, 2011 [Accessed 9 January 2013] (ILSI, 2011).

Compositional Assessment of MON 87403 Conclusion

Compositional analysis was conducted on grain and forage of MON 87403 grown at eight sites in a 2012 field production in the United States that are representative of typical agricultural regions for maize production. The compositional analysis, based on the OECD consensus document for maize, included measurement of nutrients, anti-nutrients and secondary metabolites in conventional reference hybrids to provide data on the natural variability of each compositional component analyzed.

Of the 60 components statistically assessed for MON 87403, none of the components showed a significant difference between MON 87403 and the conventional control. These results support the overall conclusion that MON 87403 was not a major contributor to variation in component levels in maize grain and forage and confirmed the compositional equivalence of MON 87403 to the conventional control in levels of these components.

B7(b) Levels of other GM-influenced constituents

Not applicable.

B7(c) Levels of naturally-occurring allergenic proteins

Not applicable.

C. NUTRITIONAL IMPACT

C1 Data on Nutritional Impact of Compositional Changes

Not Applicable.

C2 Data from an Animal Feeding Study, if Available

The data and information presented in this submission demonstrate that the food and feed derived from MON 87403 are as safe and nutritious as those derived from commercially-available, conventional maize for which there is an established history of safe consumption. Therefore, animal feeding studies do not add value to the safety of MON 87403.

PART 3 STATUTORY DECLARATION – AUSTRALIA

I, Nina McCormick, declare that the information provided in this application fully sets out the matters required and that the same are true to the best of my knowledge and belief, and that no information has been withheld that might prejudice this application.

Signature: _____

Declared before me ……

This ……12th…….. day of ……June…….. 2015.

PART 4 REFERENCES

AAFCO. 2011. Official names and definitions of feed ingredients. Pages 437-438 in 2011 Official Publication. Association of American Feed Control Officials Incorporated, Charleston, West Virginia.

Ariel, F.D., P.A. Manavella, C.A. Dezar and R.L. Chan. 2007. The true story of the HD-Zip family. Trends in Plant Science 12: 419-426.

Astwood, J.D., J.N. Leach and R.L. Fuchs. 1996. Stability of food allergens to digestion in vitro. Nature Biotechnology 14: 1269-1273.

Barker, R.F., K.B. Idler, D.V. Thompson and J.D. Kemp. 1983. Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. Plant Molecular Biology 2: 335-350.

Barry, G.F., G.M. Kishore, S.R. Padgette and W.C. Stallings. 2001. Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases. Patent 6,248,876, U.S. Patent Office, Washington, D.C.

Bevan, M., W.M. Barnes and M.-D. Chilton. 1983. Structure and transcription of the nopaline synthase gene region of T-DNA. Nucleic Acids Research 11: 369-385.

Borrás, L., G.A. Slafer and M.E. Otegui. 2004. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: A quantitative reappraisal. Field Crops Research 86: 131-146.

Borrás, L. and M.E. Westgate. 2006. Predicting maize kernel sink capacity early in development. Field Crops Research 95: 223-233.

Bruns, H.A. and H.K. Abbas. 2003. Effects of plant populations on maize hybrids in the sub-tropical Mid South USA. Maydica 48: 21-27.

Bushey, D.F., G.A. Bannon, B.F. Delaney, G. Graser, M. Hefford, X. Jiang, T.C. Lee, K.M. Madduri, M. Pariza, L.S. Privalle, R. Ranjan, G. Saab-Rincon, B.W. Schafer, J.J. Thelen, J.X. Zhang and M.S. Harper. 2014. Characteristics and safety assessment of intractable proteins in genetically modified crops. Regulatory Toxicology and Pharmacology 69: 154-170.

Caetano-Anollés, G., M. Wang, D. Caetano-Anollés and J.E. Mittenthal. 2009. The origin, evolution and structure of the protein world. Biochemical Journal 417: 621-637.

Ciarbelli, A.R., A. Ciolfi, S. Salvucci, V. Ruzza, M. Possenti, M. Carabelli, A. Fruscalzo, G. Sessa, G. Morelli and I. Ruberti. 2008. The Arabidopsis homeodomain-leucine zipper II gene family: Diversity and redundancy. Plant Molecular Biology 68: 465-478.

Codex Alimentarius. 2009. Foods derived from modern biotechnology. Second Edition. Codex Alimentarius Commission, Joint FAO/WHO Food Standards Programme, Food and Agriculture Organization of the United Nations, Rome, Italy.

Delaney, B., J.D. Astwood, H. Cunny, R.E. Conn, C. Herouet-Guicheney, S. MacIntosh, L.S. Meyer, L. Privalle, Y. Gao, J. Mattsson and M. Levine. 2008. Evaluation of protein safety in the context of agricultural biotechnology. Food and Chemical Toxicology 46: S71-S97.

Depicker, A., S. Stachel, P. Dhaese, P. Zambryski and H.M. Goodman. 1982. Nopaline synthase: Transcript mapping and DNA sequence. Journal of Molecular and Applied Genetics 1: 561-573.

DuBose, A.J., S.T. Lichtenstein, N. Narisu, L.L. Bonnycastle, A.J. Swift, P.S. Chines and F.S. Collins. 2013. Use of microarray hybrid capture and next-generation sequencing to identify the anatomy of a transgene. Nucleic Acids Research 41: e70.

Earle, F.R. and J.J. Curtis. 1946. Composition of the component parts of the maize kernel. Cereal Chemistry 23: 504-511.

Edgerton, M.D. 2009. Increasing crop productivity to meet global needs for feed, food, and fuel. Plant Physiology 149: 7-13.

Elliott, S., D. Chang, E. Delorme, T. Eris and T. Lorenzini. 2004. Structural requirements for additional N-linked carbohydrate on recombinant human erythropoietin. Journal of Biological Chemistry 279: 16854-16862.

Facciola, S. 1998. Brassicaceae. Pages 52-61 in Cornucopia II: A Source Book of Edible Plants. Kampong Publications, Vista, California.

FARRP. 2013. Allergen database, version 13. University of Nebraska, Food Allergy Research and Resource Program, Lincoln, Nebraska. http://www.allergenonline.org/.

Fisher, K.S. and A.F.E. Palmer. 1983. Maize. Pages 155-248 in Potential Productivity of Field Crops Under Different Environments. International Rice Research Institute, Manila, Philippines.

Flannery, M.L., F.J. Mitchell, S. Coyne, T.A. Kavanagh, J.I. Burke, N. Salamin, P. Dowding and T.R. Hodkinson. 2006. Plastid genome characterisation in *Brassica* and Brassicaceae using a new set of nine SSRs. Theoretical and Applied Genetics 113: 1221-1231.

Fling, M.E., J. Kopf and C. Richards. 1985. Nucleotide sequence of the transposon Tn*7* gene encoding an aminoglycoside-modifying enzyme, 3"(9)-*O*-nucleotidyltransferase. Nucleic Acids Research 13: 7095-7106.

Fraley, R.T., S.G. Rogers, R.B. Horsch, P.R. Sanders, J.S. Flick, S.P. Adams, M.L. Bittner, L.A. Brand, C.L. Fink, J.S. Fry, G.R. Galluppi, S.B. Goldberg, N.L. Hoffmann and S.C. Woo. 1983. Expression of bacterial genes in plant cells. Proceedings of the National Academy of Sciences of the United States of America 80: 4803-4807.

Fu, T.-J., U.R. Abbott and C. Hatzos. 2002. Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid - A comparative study. Journal of Agricultural and Food Chemistry 50: 7154-7160.

Giza, P.E. and R.C.C. Huang. 1989. A self-inducing runaway-replication plasmid expression system utilizing the Rop protein. Gene 78: 73-84.

Goodman, M.M. 1988. The history and evolution of maize. Critical Reviews in Plant Sciences 7: 197-220.

Hammond, B., J. Kough, C. Herouet-Guicheney and J.M. Jez. 2013. Toxicological evaluation of proteins introduced into food crops. Critical Reviews in Toxicology 43: 25-42.

Hammond, B.G. and R.L. Fuchs. 1998. Safety evaluation for new varieties of food crops developed through biotechnology. Pages 61-79 in Biotechnology and Safety Assessment. Second Edition. J.A. Thomas (ed.). Taylor & Francis, Philadelphia, Pennsylvania.

Hammond, B.G. and J.M. Jez. 2011. Impact of food processing on the safety assessment for proteins introduced into biotechnology-derived soybean and corn crops. Food and Chemical Toxicology 49: 711-721.

Harrigan, G.G., D. Lundry, S. Drury, K. Berman, S.G. Riordan, M.A. Nemeth, W.P. Ridley and K.C. Glenn. 2010. Natural variation in crop composition and the impact of transgenesis. Nature Biotechnology 28: 402-404.

Harrigan, G.G., W.P. Ridley, K.D. Miller, R. Sorbet, S.G. Riordan, M.A. Nemeth, W. Reeves and T.A. Pester. 2009. The forage and grain of MON 87460, a drought-tolerant corn hybrid, are compositionally equivalent to that of conventional corn. Journal of Agricultural and Food Chemistry 57: 9754-9763.

Henriksson, E., A.S.B. Olsson, H. Johannesson, H. Johansson, J. Hanson, P. Engström and E. Söderman. 2005. Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships. Plant Physiology 139: 509-518.

Herrmann, K.M. 1995. The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. The Plant Cell 7: 907-919.

Hileman, R.E., A. Silvanovich, R.E. Goodman, E.A. Rice, G. Holleschak, J.D. Astwood and S.L. Hefle. 2002. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. International Archives of Allergy and Immunology 128: 280-291.

Huang, S., L.A. Gilbertson, T.H. Adams, K.P. Malloy, E.K. Reisenbigler, D.H. Birr, M.W. Snyder, Q. Zhang and M.H. Luethy. 2004. Generation of marker-free transgenic maize by regular two-border *Agrobacterium* transformation vectors. Transgenic Research 13: 451-461.

Hymus, G.J., S. Cai, E.A. Kohl, H.E. Holtan, C.M. Marion, S. Tiwari, D.R. Maszle, M.R. Lundgren, M.C. Hong, N. Channa, P. Loida, R. Thompson, J.P. Taylor, E. Rice, P.P. Repetti, O.J. Ratcliffe, T.L. Reuber and R.A. Creelman. 2013. Application of HB17, an *Arabidopsis* class II homeodomain-leucine zipper transcription factor, to regulate chloroplast number and photosynthetic capacity. Journal of Experimental Botany 64: 4479-4490.

Illergård, K., D.H. Ardell and A. Elofsson. 2009. Structure is three to ten times more conserved than sequence - A study of structural response in protein cores. Proteins 77: 499-508.

ILSI. 2011. Crop Composition Database, Version 4.2. International Life Sciences Institute, Washington, D.C. http://www.cropcomposition.org/.

Jones, R.J., B.M.N. Schreiber and J.A. Roessler. 1996. Kernel sink capacity in maize: Genotypic and maternal regulation. Crop Science 36: 301-306.

Kay, R., A. Chan, M. Daly and J. McPherson. 1987. Duplication of CaMV 35*S* promoter sequences creates a strong enhancer for plant genes. Science 236: 1299-1302.

Kiesselbach, T.A. 1980. Development and structure of vegetative parts. Pages 10-37 in The Structure and Reproduction of Corn. University of Nebraska Press, Lincoln, Nebraska.

Klee, H.J., Y.M. Muskopf and C.S. Gasser. 1987. Cloning of an *Arabidopsis thaliana* gene encoding 5-enolpyruvylshikimate-3-phosphate synthase: Sequence analysis and manipulation to obtain glyphosate-tolerant plants. Molecular and General Genetics 210: 437-442.

Kovalic, D., C. Garnaat, L. Guo, Y. Yan, J. Groat, A. Silvanovich, L. Ralston, M. Huang, Q. Tian, A. Christian, N. Cheikh, J. Hjelle, S. Padgette and G. Bannon. 2012. The use of next generation sequencing and junction sequence analysis bioinformatics to achieve molecular characterization of crops improved through modern biotechnology. The Plant Genome 5: 149-163.

Lamppa, G.K., G. Morelli and N.-H. Chua. 1985. Structure and developmental regulation of a wheat gene encoding the major chlorophyll a/b-binding polypeptide. Molecular and Cellular Biology 5: 1370-1378.

Leath, M.N. and L.D. Hill. 1987. Economics of production, marketing, and utilization. Pages 210-219 in Corn: Chemistry and Technology. S.A. Watson and P.E. Ramstad (eds.). American Association of Cereal Chemists, St. Paul, Minnesota.

Lee, E.A. and M. Tollenaar. 2007. Physiological basis of successful breeding strategies for maize grain yield. Crop Science 47: S202-S215.

Leibman, M., J.J. Shryock, M.J. Clements, M.A. Hall, P.J. Loida, A.L. McClerren, Z.P. McKiness, J.R. Phillips, E.A. Rice and S.B. Stark. 2014. Comparative analysis of maize (*Zea mays*) crop performance: Natural variation, incremental improvements and economic impacts. Plant Biotechnology Journal 12: 941-950.

Liener, I.E. 1994. Implications of antinutritional components in soybean foods. Critical Reviews in Food Science and Nutrition 34: 31-67.

May, J.B. 1987. Wet milling: Process and products. Pages 377-397 in Corn: Chemistry and Technology. S.A. Watson and P.E. Ramstad (eds.). American Association of Cereal Chemists, St. Paul, Minnesota.

McElroy, D., W. Zhang, J. Cao and R. Wu. 1990. Isolation of an efficient actin promoter for use in rice transformation. The Plant Cell 2: 163-171.

McElwain, E.F. and S. Spiker. 1989. A wheat cDNA clone which is homologous to the 17 kd heat-shock protein gene family of soybean. Nucleic Acids Research 17: 1764.

Meinke, D.W., J.M. Cherry, C. Dean, S.D. Rounsley and M. Koornneef. 1998. *Arabidopsis thaliana*: A model plant for genome analysis. Science 282: 662-682.

Meinnel, T. and C. Giglione. 2008. Tools for analyzing and predicting N-terminal protein modifications. Proteomics 8: 626-649.

Metcalfe, D.D., J.D. Astwood, R. Townsend, H.A. Sampson, S.L. Taylor and R.L. Fuchs. 1996. Assessment of the allergenic potential of foods derived from genetically engineered crop plants. Critical Reviews in Food Science and Nutrition 36: S165-S186.

Moreno, F.J., F.A. Mellon, M.S.J. Wickham, A.R. Bottrill and E.N.C. Mills. 2005. Stability of the major allergen Brazil nut 2S albumin (Ber e 1) to physciologically relevant *in vitro* gastrointestinal digestion. FEBS Journal 272: 341-352.

Morris, M.L. 1998. Overview of the world maize economy. Pages 13-34 in Maize Seed Industries in Devoloping Countries. M.L. Morris (ed.). Lynne Rienner Publishers, Inc., Boulder, Colorado.

Mukherjee, K., L. Brocchieri and T.R. Bürglin. 2009. A comprehensive classification and evolutionary analysis of plant homeobox genes. Molecular Biology and Evolution 26: 2775-2794.

NCGA. 2014. World of corn 2014. National Corn Growers Association, Chesterfield, Missouri.

OECD-FAO. 2008. Agricultural outlook 2008-2017. Organisation for Economic Co-operation and Development, Food Agriculture Organization of the United Nations, Paris, France.

OECD. 2002a. Consensus document on compositional considerations for new varieties of maize (*Zea mays*): Key food and feed nutrients, antinutrients and secondary plant metabolites. ENV/JM/MONO(2002)25. Organisation for Economic Co-operation and Development, Paris, France.

OECD. 2002b. Report of the OECD workshop on the toxicological and nutritional testing of novel foods. SG/ICGB(1998)1/FINAL. Organisation for Economic Co-operation and Development, Paris, France.

Ohgishi, M., A. Oka, G. Morelli, I. Ruberti and T. Aoyama. 2001. Negative autoregulation of the *Arabidopsis* homeobox gene *ATHB-2*. The Plant Journal 25: 389-398.

Okunuki, H., R. Techima, T. Shigeta, J. Sakushima, H. Akiyama, Y. Goda, M. Toyoda and J. Sawada. 2002. Increased digestibility of two products in genetically modified food (CP4 EPSPS and Cry1Ab) after preheating. Journal of the Food Hygienic Society of Japan 43: 68-73.

Padgette, S.R., D.B. Re, G.F. Barry, D.E. Eichholtz, X. Delannay, R.L. Fuchs, G.M. Kishore and R.T. Fraley. 1996. New weed control opportunities: Development of soybeans with a Roundup Ready™ gene. Pages 53-84 in Herbicide-Resistant Crops: Agricultural, Environmental, Economic, Regulatory and Technical Aspects. S.O. Duke (ed.). CRC Press, Inc., Boca Raton, Florida.

Pariza, M.W. and E.A. Johnson. 2001. Evaluating the safety of microbial enzyme preparations used in food processing: Update for a new century. Regulatory Toxicology and Pharmacology 33: 173-186.

Pattison, R.J. and A. Amtmann. 2009. N-glycan production in the endoplasmic reticulum of plants. Trends in Plant Science 14: 92-99.

Perry, T.W. 1988. Corn as a livestock feed. Pages 941-963 in Corn and Corn Improvement. Third Edition. G.F. Sprague and J.W. Dudley (eds.). American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc., Madison, Wisconsin.

Pilgeram, A.L., D.C. Sands, D. Boss, N. Dale, D. Wichman, P. Lamb, C. Lu, R. Barrows, M. Kirkpatrick, B. Thompson and D.L. Johnson. 2007. *Camelina sativa*, a Montana omega-3 and fuel crop. Pages 129-131 in Issues in New Crops and New Uses. J. Janick and A. Whipkey (eds.). ASHS Press, Alexandria, Virginia.

RFA. 2010. Industry resources: Co-products. Renewable Fuels Association, Washington, D.C. http://www.ethanolrfa.org/pages/industry-resources-coproducts [Accessed February 24, 2011].

Rice, E.A., A. Khandelwal, R.A. Creelman, C. Griffith, J.E. Ahrens, J.P. Taylor, L.R. Murphy, S. Manjunath, R.L. Thompson, M.J. Lingard, S.L. Back, H. Larue, B.R. Brayton, A.J. Burek, S. Tiwari, L. Adam, J.A. Morrell, R.A. Caldo, Q. Huai, J.-L.K. Kouadio, R. Kuehn, A.M. Sant, W.J. Wingbermuehle, R. Sala, M. Foster, J.D. Kinser, R. Mohanty, D. Jiang, T.E. Ziegler, M.G. Huang, S.V. Kuriakose, K. Skottke, P.P. Repetti, T.L. Reuber, T.G. Ruff, M.E. Petracek and P.J. Loida. 2014. Expression of a truncated ATHB17 protein in maize increases ear weight at silking. PLoS ONE 9: e94238.

Ridley, W.P., G.G. Harrigan, M.L. Breeze, M.A. Nemeth, R.S. Sidhu and K.C. Glenn. 2011. Evaluation of compositional equivalence for multitrait biotechnology crops. Journal of Agricultural and Food Chemistry 59: 5865-5876.

Ritchie, S.W., J.J. Hanway and G.O. Benson. 1997. How a corn plant develops. Special Report No. 48. Iowa State University of Science and Technology, Cooperative Extension Service, Ames, Iowa.

Rooney, L.W. and S.O. Serna-Saldivar. 1987. Food uses of whole corn and dry-milled fractions. Pages 399-429 in Corn: Chemistry and Technology. S.A. Watson and P.E. Ramstad (eds.). American Association of Cereal Chemists, St. Paul, Minnesota.

Ruberti, I., G. Sessa, S. Lucchetti and G. Morelli. 1991. A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. EMBO Journal 10: 1787-1791.

Safawo, T., N. Senthil, M. Raveendran, S. Vellaikumar, K.N. Ganesan, G. Nallathambi, S. Saranya, V.G. Shobhana, B. Abirami and E.V. Gowri. 2010. Exploitation of natural variability in maize for β-carotene content using HPLC and gene specific markers. Electronic Journal of Plant Breeding 1: 548-555.

Salomon, S. and H. Puchta. 1998. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO Journal 17: 6086-6095.

Sandring, S., M.-A. Riihimäki, O. Savolainen and J. Ågren. 2007. Selection on flowering time and floral display in an alpine and a lowland population of *Arabidopsis lyrata*. Journal of Evolutionary Biology 20: 558-567.

Sawa, S., M. Ohgishi, H. Goda, K. Higushi, Y. Shimada, S. Yoshida and T. Koshiba. 2002. The *HAT2* gene, a member of the HD-Zip gene family, isolated as an auxin inducible gene by DNA microarray screening, affects auxin response in *Arabidopsis*. The Plant Journal 32: 1011-1022.

Sessa, G., G. Morelli and I. Ruberti. 1993. The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. EMBO Journal 12: 3507-3517.

Severini, A.D., L. Borrás, M.E. Westgate and A.G. Cirilo. 2011. Kernel number and kernel weight determination in dent and popcorn maize. Field Crops Research 120: 360-369.

Sidorov, V. and D. Duncan. 2009. *Agrobacterium*-mediated maize transformation: Immature embryos versus callus. Pages 47-58 in Methods in Molecular Biology: Transgenic Maize - Methods and Protocols. M.P. Scott (ed.). Humana Press, Inc, Totowa, New Jersey.

Silvanovich, A., M.A. Nemeth, P. Song, R. Herman, L.Tagliani and G.A. Bannon. 2006. The value of short amino acid sequence matches for prediction of protein allergenicity. Toxicological Sciences 90: 252-258.

Simpson, G.G. and W. Filipowicz. 1996. Splicing of precursors to mRNA in higher plants: Mechanism, regulation and sub-nuclear organisation of the spliceosomal machinery. Plant Molecular Biology 32: 1-41.

Sjoblad, R.D., J.T. McClintock and R. Engler. 1992. Toxicological considerations for protein components of biological pesticide products. Regulatory Toxicology and Pharmacology 15: 3-9.

Sorin, C., M. Salla-Martret, J. Bou-Torrent, I. Roig-Villanova and J.F. Martínez-García. 2009. ATHB4, a regulator of shade avoidance, modulates hormone response in Arabidopsis seedlings. The Plant Journal 59: 266-277.

Stalker, D.M., C.M. Thomas and D.R. Helinski. 1981. Nucleotide sequence of the region of the origin of replication of the broad host range plasmid RK2. Molecular and General Genetics 181: 8-12.

Steindler, C., A. Matteucci, G. Sessa, T. Weimar, M. Ohgishi, T. Aoyama, G. Morelli and I. Ruberti. 1999. Shade avoidance responses are mediated by the ATHB-2 HD-Zip protein, a negative regulator of gene expression. Development 126: 4235-4245.

Thomas, K., M. Aalbers, G.A. Bannon, M. Bartels, R.J. Dearman, D.J. Esdaile, T.J. Fu, C.M. Glatt, N. Hadfield, C. Hatzos, S.L. Hefle, J.R. Heylings, R.E. Goodman, B. Henry, C. Herouet, M. Holsapple, G.S. Ladics, T.D. Landry, S.C. MacIntosh, E.A. Rice, L.S. Privalle, H.Y. Steiner, R. Teshima, R. van Ree, M. Woolhiser and J. Zawodny. 2004. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. Regulatory Toxicology and Pharmacology 39: 87-98.

Thomas, K., G. Bannon, S. Hefle, C. Herouet, M. Holsapple, G. Ladics, S. MacIntosh and L. Privalle. 2005. In silico methods for evaluating human allergenicity to novel proteins: International Bioinformatics Workshop Meeting Report, 23-24 February 2005. Toxicological Sciences 88: 307-310.

University of Utah. 2014. The evolution of corn. University of Utah, Salt Lake City, Utah. http://learn.genetics.utah.edu/content/selection/corn/ [Accessed February 18, 2014].

USDA-FAS. 2013a. Corn: Production (1000 MT), by country. U.S. Department of Agriculture, Foreign Agricultural Service, Washington, D.C.

USDA-FAS. 2013b. Corn: Production (1000 MT), world. U.S. Department of Agriculture, Foreign Agricultural Service, Washington, D.C. http://apps.fas.usda.gov/psdonline/psdResult.aspx [Accessed January 8, 2014].

USDA-FAS. 2014. World agricultural production. WAP 4-14. U.S. Department of Agriculture, Foreign Agricultural Service, Washington, D.C.

Vassilopoulou, E., N. Rigby, F.J. Moreno, L. Zuidmeer, J. Akkerdaas, I. Tassios, N.G. Papadopoulos, P. Saxoni-Papageorgiou, R. van Ree and C. Mills. 2006. Effect of *in vitro* gastric and duodenal digestion on the allergenicity of grape lipid transfer protein. Journal of Allergy and Clinical Immunology 118: 473-480.

Vieira, J. and J. Messing. 1987. Production of single-stranded plasmid DNA. Methods in Enzymology 153: 3-11.

Vieths, S., J. Reindl, U. Müller, A. Hoffmann and D. Haustein. 1999. Digestibility of peanut and hazelnut allergens investigated by a simple in vitro procedure. European Food Research and Technology 209: 379-388.

Vitale, A. and J. Denecke. 1999. The endoplasmic reticulum - Gateway of the secretory pathway. The Plant Cell 11: 615-628.

Watson, S.A. 1988. Corn marketing, processing, and utilization. Pages 881-940 in Corn and Corn Improvement. Third Edition. G.F. Sprague and J.W. Dudley (eds.). American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc., Madison, Wisconsin.

Wych, R.D. 1988. Production of hybrid seed corn. Pages 565-607 in Corn and Corn Improvement. Third Edition. G.F. Sprague and J.W. Dudley (eds.). American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc., Madison, Wisconsin.

Yagami, T., Y. Haishima, A. Nakamura, H. Osuna and Z. Ikezawa. 2000. Digestibility of allergens extracted from natural rubber latex and vegetable foods. Journal of Allergy and Clinical Immunology 106: 752-762.

Yang, L., C. Wang, A. Holst-Jensen, D. Morisset, Y. Lin and D. Zhang. 2013. Characterization of GM events by insert knowledge adapted re-sequencing approaches. Scientific Reports 3: 2839.

Yates, B., A. De Soyza, R. Harkawat and C. Stenton. 2008. Occupational asthma caused by *Arabidopsis thaliana*: A case of laboratory plant allergy. European Respiratory Journal 32: 1111-1112.

Zaidi, P.H., G. Srinivasan and C. Sanchez. 2003. Morpho-physiological traits associated with variable field performance of different types maize germplasm across multiple environments. Maydica 48: 207-220.

Zambryski, P., A. Depicker, K. Kruger and H.M. Goodman. 1982. Tumor induction by *Agrobacterium tumefaciens*: Analysis of the boundaries of T-DNA. Journal of Molecular and Applied Genetics 1: 361-370.

Zhao, Y., Y. Zhou, H. Jiang, X. Li, D. Gan, X. Peng, S. Zhu and B. Cheng. 2011. Systematic analysis of sequences and expression patterns of drought-responsive members of the HD-Zip gene family in maize. PLoS ONE 6: e28488.

Zhou, J., G.G. Harrigan, K.H. Berman, E.G. Webb, T.H. Klusmeyer and M.A. Nemeth. 2011. 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. Journal of Agricultural and Food Chemistry 59: 8822-8828.