SUBMISSION COUNTRY: **Australia**

EVENT NAME: **Soy MON 87751**

SUBMISSION TYPE: **Food and Feed (FSANZ)**

DATE: **Jan 6, 2015 (first draft complete)**
Jan 7 – 20, 2015 (AU team & SDC review)
Jan 21 – 26, 2015 (SDC Revision)
Jan 27 – 30, 2015 (Reference & EndNote)
Feb 2 – 6 (Final Revision)
February 6, 2015 (send to cRAM for the submission)

PREPARED BY: [redacted]

Application to Food Standards Australia New Zealand for the Inclusion of Soy MON 87751 in Standard 1.5.2 - Food Derived from Gene Technology

Submitted by:

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

February 2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

UNPUBLISHED REPORTS BEING SUBMITTED

[redacted] 2014. Amended Report for MSL0025106: Segregation Analysis of the Coding Sequences Present in Insect Protected Soybean MON 87751 Across Multiple Generations. **MSL0025712**. Monsanto Company.

[redacted] 2013. Assessment of the in vitro Digestibility of Cry1A.105 Protein in Simulated Gastric and Simulated Intestinal Fluids. **MSL0024977**. Monsanto Company.

[redacted] 2005. An Acute Oral Toxicity Study in Mice with Cry1A.105 Protein. **CRO-2005-050**. Monsanto Company.

[redacted] 2006. An Acute Oral Toxicity Study in Mice with Cry2Ab2 Protein. **CRO-2005-049**. Monsanto Company.

[redacted] 2013. Characterization of the Cry1A.105 Protein Purified from the Soybean Seed of MON 87751 and Comparison of the Physicochemical and Functional Properties of the Plant-Produced and *Escherichia coli*-Produced Cry1A.105 Proteins. **MSL0025088**. Monsanto Company.

[redacted] 2014 Amended Report for MSL0025312: Molecular Characterization of Insect Protected Soybean MON 87751. **MSL0025901**. Monsanto Company.

[redacted]. 2013. Effect of Heat Treatment on the Functional Activity of *Escherichia coli* (*E. coli*)-Produced MON 87751 Cry1A.105 Protein. **MSL0025092**. Monsanto Company.

[redacted]. 2013. Effect of Heat Treatment on the Functional Activity of *Escherichia coli* (*E. coli*)-Produced MON 87751 Cry2Ab2 Protein. **MSL0025102**. Monsanto Company.

[redacted]. 2013. Bioinformatics Evaluation of the Transfer DNA Insert in MON 87751 Utilizing the AD_2013, TOX_2013 and PRT_2013 Databases. **MSL0024971**. Monsanto Company.

[redacted] 2013. Bioinformatics Evaluation of DNA Sequences Flanking the 5' and 3' Junctions of Inserted DNA in MON 87751. Assessment of Putative Polypeptides. **MSL0024956**. Monsanto Company.

[redacted] 2013. Bioinformatics Evaluation of the Cry1A.105 Protein in MON 87751 Utilizing the AD_2013, TOX_2013 and PRT_2013 Databases. **MSL0024925**. Monsanto Company.

[redacted] 2013. Bioinformatics Evaluation of the Cry2Ab2 Protein in MON 87751 Utilizing the AD_2013, TOX_2013 and PRT_2013 Databases. **MSL0024926**. Monsanto Company.

[redacted]. 2014. Amended Report for MSL0024609: Compositional Analyses of Soybean Forage and Seed Collected from MON 87751 Grown in the United States during the 2012 Season. **MSL0026251**. Monsanto Company.

[redacted] 2013. Amended Report for MSL0024805: Assessment of Cry1A.105 and Cry2Ab2 Protein Levels in Soybean Tissues Collected from MON 87751 Produced in U.S. Field Trials during 2012. **MSL0025199**. Monsanto Company.

[redacted]. 2013. Assessment of the in vitro Digestibility of Cry2Ab2 Protein in Simulated Gastric and Simulated Intestinal Fluids. **MSL0025099**. Monsanto Company.

[redacted]. 2013. Amended from MSL0024792: Characterization of the Cry2Ab2 Protein Purified from the Soybean Seed of MON 87751 and Comparison of the Physicochemical and Functional Properties of the Plant-Produced and Escherichia coli-Produced Cry2Ab2 Proteins. **MSL0025096**. 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 13*
3.1.3 Purpose of the application	*Page 13*
3.1.4 Justification for the application	*Page 14*
3.1.5 Information to support the application	*Volumes 1 - 2*
3.1.6 Assessment procedure	*Page 15*
☑ 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 16*
3.1.9 International and Other National Standards	*Page 16*
3.1.10 Statutory Declaration	*Page 158*
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 17*
☑ A.2 History of use of host and donor organisms	*Page 19*
☑ A.3 Nature of genetic modification	*Page 26*
☑ A.4 Analytical method for detection	*Page 70*
☑ B.1 Equivalence studies	*Page 71*
☑ B.2 Antibiotic resistance marker genes (if used)	*Page 94*
☑ B.3 Characterisation of novel protein(s)/substances	*Page 95*
☑ B.4 Potential toxicity of novel protein(s)/substances	*Page 104*
☑ B.5 Potential allergenicity of novel protein(s)	*Page 130*
☑ B.6 Toxicity of novel herbicide metabolites	*Page 134*
☑ B.7 Compositional Analyses	*Page 134*
☑ C.1 Nutritional impact of GM food	*Page 157*
☑ C.2 Animal feeding studies (if available)	*Page 157*

ABBREVIATIONS AND DEFINITIONS[1]

~	approximately
ADF	acid detergent fiber
ANOVA	analysis of variance
APHIS	Animal and Plant Health Inspection Service
bp	base pairs
Bt	*Bacillus thuringiensis*
CTAB	hexadecyltrimethylammonium bromide
DAP	days after planting
dATP	deoxyadenosine triphosphate
DDI	daily dietary intake
DEEM-FCID	Dietary Exposure Evaluation Model-Food Commodity Intake Database
DHB	2,5-dihydroxybenzoic acid
DTT	dithiothreitol
dw	dry weight
DWCF	dry weight conversion factor
E. coli	*Escherichia coli*
ELISA	enzyme-linked immunosorbent assay
EPA	Environmental Protection Agency
EUP	experimental use permit
ETS	Excellence Through Stewardship
FA	fatty acid
FDA	U.S. Food and Drug Administration
FMOC	fluorenylmethyl chloroformate
FSANZ	Food Standards Australia and New Zealand
fw	fresh weight
Gb	gigabases
ha	hectares
HPLC	high-performance liquid chromatography
HRP	horseradish peroxidase
ILSI-CCDB	International Life Sciences Institute-Crop Composition Database
IPM	integrated pest management
IRM	insect resistance management
JSC	junction sequence class
kDa	kilodalton
LOD	limit of detection
LOQ	limit of quantitation
MEEC	maximum expected environmental concentration
Mg/ha	megagrams/hectare
MMT	million metric tons
MOA	mode-of-action
MOE	margin of exposure
n	number of samples
NDF	neutral detergent fiber
NFDM	nonfat dry milk

[1] [redacted]. 2003. Handbook of Technical Writing, 7th edn., pp. 2-7. Bedford/St. Martin's, Boston, MA.

NGS/JSA	Next Generation Sequencing/Junction Sequence Analysis
NOAELs	no observable adverse effect levels
NOEC	no observable effect concentration
nt	nucleotide
NTO	non-target organism
OECD	Organisation for Economic Co-operation and Development
OM	organic matter
OPA	o-phthalaldehyde
OSL	over season leaf
PBST	phosphate buffered saline containing 0.05% (v/v) Tween
PCR	polymerase chain reaction
PIP	plant incorporated protectant
PTH-AA	phenylthiohydantoin-amino acid
SBV	soybean variant
SD	standard deviation
SDS	sodium dodecyl sulfate
S.E.	standard error
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
sp.	species
TDF	total dietary fiber
T-DNA	transfer DNA
TFA	trifluoroacetic acid
TSSP	tissue-specific site pool
USDA	United States Department of Agriculture
UTR	untranslated region
UV	ultraviolet
v/v	volume to volume

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)	12 / 600 St Kilda Road, Melbourne, Victoria, 3004 PO Box 6051, St Kilda Road Central, Victoria, 8008
(d) Telephone and facsimile numbers	[redacted]
(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	Not applicable

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 MON 87751 soybean and products containing MON 87751 soybean (hereafter referred to as MON 87751) to the Table to Clause 2 (see below).

Food derived from gene technology	Special requirements
Food derived from insect protected soybean line MON 87751	None

1.3 Justification for the Application

1.3(a) The need for the proposed change

Monsanto Company has developed insect-protected soybean MON 87751 that produces the CrylA.105 and Cry2Ab2 insecticidal (Cry) proteins. Cry1A.105 is a modified Cry1A protein derived from *Bacillus thuringiensis*. Cry2Ab2 is derived from *B. thuringiensis* subsp. *kurstaki*. The CrylA.105 and Cry2Ab2 proteins provide protection from feeding damage caused by targeted lepidopteran insect pests. Studies conducted with MON 87751 demonstrated efficacy against key soybean pests including *Crocidosema aporema* (bean shoot moth), *Rachiplusia nu* (sunflower looper) and *Spodoptera frugiperda* (fall armyworm). Cry1A.105 and Cry2Ab2 are also known to be active against lepidopteran soybean pests such as *Anticarsia gemmatalis* (velvetbean caterpillar), *Chrysodeixis includens* (soybean looper) and *Helicoverpa zea* (corn earworm). The season-long expression pattern of Cry1A.105 and Cry2Ab2 in MON 87751 is expected to control target insects that are heterozygous for resistance genes specific to one of the proteins and provide an effective tool in managing potential insect resistance, thus prolonging the durability of this product. MON 87751 is expected to provide benefits to growers similar to those obtained by use of other lepidopteran-protected crop varieties, including reduced use of broad spectrum insecticides, increased yield protection and increased worker safety.

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

MON 87751 will be combined, through traditional breeding methods, with other deregulated biotechnology-derived traits to provide additional protection against lepidopteran soybean pests as well as tolerance to multiple herbicides. These next generation combined-trait soybean products will offer the ability to maximize grower choice, improve production efficiency and increase pest control and weed 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 87751 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 soybean products.

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

Industry and business in general:

- Sellers of processed foods containing soybean derivatives would benefit as foods derived from soybean MON 87751 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 soybean products or imported foods manufactured using soybean derivatives.
- Possible cost to food industry as some food ingredients derived from soybean MON 87751 would be required to be labelled.

Government:

- Benefit that if soybean MON 87751 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 soybean MON 87751 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.

1.4(b) Impact on international trade

If the draft variation to permit the sale and use of food derived from MON 87751 was rejected it would result in the requirement for segregation of any products containing MON 87751 from those containing approved soybean, which would be likely to increase the costs of imported soybean-derived foods.

It is important to note that if the draft variation is approved, MON 87751 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 soybean nutrients and anti-nutrients based on conventional commercial soybean varieties (OECD, 2012).

1.7(b) Other national standards or regulations

Monsanto has submitted a food and feed safety and nutritional assessment summary for MON 87751 to the United States Food and Drug Administration (FDA) and has also requested a Determination of Nonregulated Status for MON 87751, including all progenies derived from crosses between MON 87751 and other soybean, from the Animal and Plant Health Inspection Service (APHIS) of the U.S. Department of Agriculture (USDA). MON 87751 has received USDA deregulation on October 17th 2014.

Applications have also been submitted to Canadian Food Inspection Agency (CFIA) and Health Canada (HC), Korea Ministry of Food and Drugs Safety (MFDS) for food, and Rural Development Administration (RDA) for feed use, Japan's Ministry of Health, Labour, and Welfare (MHLW) for food use and Japan's Ministry of Agriculture, Forestry, and Fisheries for feed use, Ministry of Agriculture, People's Republic of China, Taiwan's Department of Health, Argentina's National Advisory Commission on Agriculture Biotechnology (CONABIA) and National Service of Agriculture and Cattle Sanitary and Food Safety (SENASA) for food and feed, Brazil's National Biosafety Technical Committee (CTNBio) for food and the European Food Safety Authority (EFSA).

CFIA and HC had granted approval for food and feed use of MON 87751 on October 31st 2014.

Regulatory submissions will be made to countries that import significant soybean or food and feed products derived from countries where MON 87751 soybean 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 Mexico's Ministry of Health and Philippine Bureau of Plant Industry, Department of Agricultureas well as to regulatory authorities in other soybean 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

Monsanto Company has developed insect-protected soybean MON 87751 that produces the CrylA.105 and Cry2Ab2 insecticidal (Cry) proteins derived from *B. thuringiensis*. Cry1A.105 is a modified Cry1A protein and Cry2Ab2 is derived from *B. thuringiensis* subsp. *kurstaki*. The CrylA.105 and Cry2Ab2 proteins provide protection from feeding damage caused by targeted lepidopteran insect pests. Studies conducted with MON 87751 demonstrated that this transformation event was not damaged by infestations of *Crocidosema aporema* (bean shoot moth), *Rachiplusia nu* (sunflower looper) and *Spodoptera frugiperda* (fall armyworm). At the same time, neither Cry1A.105 nor Cry2Ab2 exhibits biological activity against coleopteran, hymenopteran or hemipteran insects at exposure concentrations well above field exposure levels (U.S. EPA, 2010). The season-long expression pattern for Cry1A.105 and Cry2Ab2 in MON 87751 is expected to control target insects that are heterozygous for resistance genes specific to one of the proteins thereby providing an effective tool in managing potential insect resistance and prolonging product durability. MON 87751 would be most efficacious in soybean production areas where insecticides are typically applied to control lepidopteran insects. Accordingly, MON 87751 is expected to provide benefits similar to those provided by existing lepidopteran-protected crops, including improved control of lepidopteran pests, reduced use of broad spectrum insecticides (Brookes and Barfoot, 2012), increased yield protection and increased worker safety.

Several insect-protected crops derived from biotechnology, including maize MON 89034, expressing both Cry1A.105 and Cry2Ab2 as well as lepidopteran protected cotton MON 15985, expressing Cry2Ab2, have been deregulated and registered for commercial release in the U.S. since 2002. The Cry1A.105 and Cry2Ab2 proteins expressed in MON 87751 have greater than 99% and 97% amino acid identity, respectively, to the Cry1A.105 and Cry2Ab2 expressed proteins in MON 89034. Furthermore, the protease-resistant core domains of the modified Cry1A.105 and Cry2Ab2 proteins expressed in MON 87751 share 100% deduced amino acid identity to MON 89034 expressed core domains of the Cry1A.105 and Cry2Ab2 proteins. These domains are responsible for insecticidal activity and specificity (Gill *et al.*, 1992; Widner and Whiteley, 1989). Therefore, MON 87751, though a new event, contains the proteins and associated modes of action which have already received a determination of non-regulated status (USDA-APHIS, 2008) and approved (A595) into the Food Standards Australia and New Zealand Standards 1.5.2 in December 2008 (FSANZ, 2008). MON 87751 had also received a determination of non-regulated status on October 2014 (USDA-APHIS, 2014).

MON 87751 will be combined, through traditional breeding methods, with other deregulated biotechnology-derived traits to provide additional protection against lepidopteran soybean pests as well as tolerance to herbicides. These next generation combined-trait soybean

products will offer the ability to maximize grower choice, improve production efficiency and increase pest and weed control durability.

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 87751 has been assigned the unique identifier MON-87751- 7.

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

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. Soybean containing the transformation event MON 87751 will be produced in Argentina and Brazil. There are currently no plans to produce this product in Australia and New Zealand.

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

MON 87751 is intended primarily for use as a broad-acre commodity (or field) soybean and not for vegetable, garden, or food-grade soybean that are generally used to produce tofu, soybean sprouts, soymilk, green soybean (*e.g.*, edamame) or other similar food items. Vegetable and food-grade soybean generally have a different size, flavor, texture and other characteristics than field soybean, and are more easily cooked. Other than the introduction of the insect protection trait, MON 87751 is not materially different from conventional field soybeans and can be processed into a wide variety of food products as described in Section A2(b)(iii) and Section A2(b)(iv). Field soybean is a blended commodity that is highly processed before being consumed by humans. Thus, most food products derived from MON 87751 would likely be blended with those derived from other commercial soybean varieties before entering the human food supply.

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

MON 87751 contains the CrylA.105 and Cry2Ab2 insecticidal (Cry) proteins. Cry1A.105 is a modified Cry1A protein derived from *Bacillus thuringiensis*. Cry2Ab2 is derived from *B. thuringiensis* subsp. *kurstaki*. The taxonomy of this organism is as follows:

Kingdom: Eubacteria

Phylum: Firmicutes

Class: Bacilli

Order: Bacillales

Family: Bacillaceae

Genus: Bacillus

Species: thuringiensis

Subspecies: kurstaki, aizawai

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

There is no evidence of human or animal pathogenicity for any of the donor organisms of the coding and non-coding DNA sequences present in MON 87751. DNA has always been present in food and, upon consumption, is quickly degraded by restriction nucleases present in the gastrointestinal tract of humans and animals to nucleic acids. According to the U.S. FDA (U.S. FDA, 1992), nucleic acids are present in the cells of every living organism, do not raise concerns as a component of food, and are generally recognized as safe. Results from an International Life Sciences Institute (ILSI) workshop on safety considerations of DNA in food were reported (Jonas *et al.*, 2001) and confirmed that: 1) all DNA, including recombinant DNA, is composed of the same four nucleotides; 2) there are no changes to the chemical characteristics or the susceptibility to degradation by chemical or enzymatic hydrolysis of recombinant DNA as compared to non-recombinant DNA; and 3) there is no evidence that DNA from dietary sources has ever been incorporated into the mammalian genome. Additionally, the European Food Safety Authority (EFSA) has reported that a large number of experimental studies have shown that recombinant DNA consumed by livestock has not been subsequently detected in tissues, fluids, or edible products of these farm animals (EFSA, 2007).

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

The donor organism for *cry1A.105* and *cry2Ab2*, *B. thuringiensis,* has been used commercially in the United States to produce microbial-derived products with insecticidal activity. Applications of sporulated *B. thuringiensis* have a long history of safe use for pest control in agriculture, especially in organic farming (Cannon, 1993; U.S. EPA, 1988; WHO,

1999). Microbial pesticides containing *B. thuringiensis* Cry proteins have been subjected to extensive toxicity testing showing no adverse effects to human health (Baum *et al.*, 1999; Betz *et al.*, 2000; McClintock *et al.*, 1995; Mendelsohn *et al.*, 2003; U.S. EPA, 2001; U.S. EPA, 2005). Additionally, there are no confirmed cases of allergic reactions to Cry proteins in microbial-derived *B. thuringiensis* products during more than 50 years of use. The history of safe use of *Bacillus thuringiensis* has been previously reviewed and approved by FSANZ as a part of the safety assesessment of MON 89034 corn (A595) which expresses *cry1A.105* and *cry2Ab2*.

A2(b) Description of the host organism

A2(b)(i) Phenotypic information

Cultivated soybean, *Glycine max* (L.) Merr., is a diploidized tetraploid (2n=40), which belongs to the family Fabaceae, the subfamily Papilionoideae, the tribe Phaseoleae, the genus *Glycine* Willd., and the subgenus *Soja* (Moench) F.J. Herm.

Family: Fabaceae

Subfamily: Papilionoideae

Tribe: Phaseoleae

Genus: *Glycine* Willd.

Subgenus: *Soja* (Moench) F.J. Herm.

Species: *Glycine max* (L.) Merr.

The genus *Glycine* Willd. is of Asian and Australian origin and is divided into two subgenera, *Glycine* and *Soja* (Moench) F.J. Herm. The subgenus *Glycine* consists of 22 wild perennial species, which are indigenous to Australia, West, Central and South Pacific Islands, China, Russia, Japan, Indonesia, Korea, Papua New Guinea, the Philippines, and Taiwan (Hymowitz, 2004). The subgenus *Soja* includes the cultivated soybean, *G. max* (L.) Merr. and its wild annual relatives from Asia, *G. soja* Sieb. and Zucc. The list of species in the genus *Glycine* Willd. is presented in Table 1.

Table 1. List of Species in the Genus Glycine Willd., 2n Chromosome Number, Genome Symbol and Distribution

Genus	2n	Genome[1]	Distribution
Subgenus *Glycine*			
1. *G. albicans* Tind. & Craven	40	I1	Australia
2. *G. aphyonota* B. Pfeil	40	--[2]	Australia
3. *G. arenaria* Tind.	40	HH	Australia
4. *G. argyrea* Tind.	40	A2A2	Australia
5. *G. canescens* F.J. Herm.	40	AA	Australia
6. *G. clandestina* Wendl.	40	A1A1	Australia
7. *G. curvata* Tind.	40	C1C1	Australia
8. *G. cyrtoloba* Tind.	40	CC	Australia
9. *G. dolichocarpa* Tateishi and Ohashi	80	--	Taiwan
10. *G. falcate* Benth.	40	FF	Australia
11. *G. hirticaulis* Tind. & Craven	40	H1H1	Australia
	80	--	Australia
12. *G. lactovirens* Tind. & Craven.	40	I1I1	Australia
13. *G. latifolia* (Benth.) Newell & Hymowitz	40	B1B1	Australia
14. *G. latrobeana* (meissn.) Benth.	40	A3A3	Australia
15. *G. microphylla* (Benth.) Tind.	40	BB	Australia
16. *G. peratosa* B. Pfeil & Tind.	40	--	Australia
17. *G. pindanica* Tind. & Craven	40	H3H2	Australia
18. *G. pullenii* B. Pfeil, Tind. & Craven	40	--	Australia
19. *G. rubiginosa* Tind. & B. Pfeil	40	--	Australia
20. *G. stenophita* B. Pfeil & Tind.	40	B3B3	Australia
21. *G. tabacina* (Labill.) Benth.	40	B2B2	Australia
	80	Complex[3]	Australia, West Central and South Pacific Islands
22. *G. tomentella* Hayata	38	EE	Australia
	40	DD	Australia, Papua New Guinea
	78	Complex[4]	Australia, Papua New Guinea
	80	Complex[5]	Australia, Papua New Guinea, Indonesia, Philippines, Taiwan
Subgenus *Soja* (Moench) F.J. Herm.			
23. *G. soja* Sieb. & Zucc.	40	GG	China, Russia, Taiwan, Japan, Korea (Wild Soybean)
24. *G. max* (L.) Merr.	40	GG	Cultigen (Soybean)

[1] Genomically similar species carry the same letter symbols.
[2] Genome designation has not been assigned to the species.
[3] Allopolyploids (A and B genomes) and segmental allopolyploids (B genomes).
[4] Allopolyploids (D and E, A and E, or any other unknown combination).
[5] Allopolyploids (A and D genomes, or any other unknown combination).
Note: Table is adapted from Hymowitz, (2004).

Glycine soja grows wild in China, Japan, Korea, the Russian Far East, and Taiwan, and is commonly found in fields, hedgerows, roadsides and riverbanks (Lu, 2004). The plant is an annual, slender in build with narrow trifoliolate leaves. The purple or very rarely white flowers are inserted on short, slender racemes. The pods are short and tawny with hirsute pubescence, producing oval-oblong seeds (Hymowitz, 2004).

Glycine max (L.) Merr., the cultivated soybean, is an annual that generally exhibits an erect, sparsely branched, bush-type growth habit with trifoliolate leaves. The leaflets are broadly ovate, and the purple, pink, or white flowers are borne on short axillary racemes or reduced peduncles. The pods are either straight or slightly curved, and ovoid to subspherical seeds are produced in the pods (Hymowitz, 2004).

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

Soybean is an annual crop that is planted in late spring from April to May in the northern hemisphere, and from November to February in the southern hemisphere including second cropping. The system of soybean growth stages divides plant development into vegetative (V) and reproductive (R) stages (Pedersen, 2004). The vegetative stages begin with VE, which designates emergence. V stages continue and are numbered according to how many fully-developed trifoliate leaves are present (*i.e.* V1, V2, etc.). The reproductive (R) stages begin at flowering (R1) and include pod development and plant maturation. Full maturity is designated as R8 stage.

Soybean seed germinates when the soil temperature reaches 10 °C and emerges in a 5-7 day period under favorable conditions (OECD, 2000). Soybean roots normally become infected with *Bradyrhizobium japonicum* bacteria, which cause formation of round- or oval-shaped root growths known as nodules. There is a mutual benefit in the symbiotic relationship between *Bradyrhizobium japonicum* bacteria and the soybean plant (Pedersen, 2004). Much of the nitrogen (N) required by the soybean plant is supplied through a process called N-fixation, in which atmospheric N_2 is converted into a usable form for the plant through the symbiotic relationship between *Bradyrhizobium japonicum* bacteria and the soybean plant. Soybean vegetative development phase lasts about 40 days, during which time the root nodules develop slowly, but do not become fully functional. Soybeans grow most rapidly when air temperatures are between 25 °C and 30 °C (OECD, 1993). Pods typically develop in late summer, and harvest occurs in the fall. The life cycle of soybean is approximately 100 to 160 days, depending on the variety and the region it is cultivated. Harvesting may begin when the plants are completely dry and the seeds are fully matured within the pods.

Soybean cultivars are identified based on bands of adaptation that run east-west, determined by latitude and day length (OECD, 2000). They are classified into one of 13 maturity groups (000, 00, 0, I to X), which determines the area where plants will be most productive (Palmer *et al.*, 1987). As soybean is a short day plant, time to maturity is strongly influenced by photoperiod. Cultivars of maturity group X are adapted to southern latitudes (*e.g.* 0-10° N equatorial zones), where the summer photoperiod is short, whereas cultivars of maturity group 000 are adapted to the highest latitudes (*e.g.* >45° N) where the summer photoperiod is longest. Cultivars most adapted to production in Europe are of maturity groups 00 to early

group II, in U.S. are of groups 0 to V and in South America are groups IV, V (Argentina), VI to IX (Brazil).

There are two types of stem growth habits in soybean: indeterminate and determinate. Flowering activity of indeterminate soybeans is spread over a three to five week period once the critical day length has occurred. In indeterminate types, flowering begins on the third to sixth nodes of the main sten , depending on the vegitative stage at the time of flowering and progresses upward and downward (Pedersen, 2004). The longer flowering period allows these types to adjust to the effects of short-term stress or unfavorable environmental conditions. Indeterminate soybean has terminal buds that continue to grow several weeks after flowering. Determinate soybean types have terminal buds that cease to grow when the plant starts to flower, thus they have a short flowering period. At maturity, the determinate plants have a rather dense cluster of pods on a terminal raceme. Indeterminate varieties are almost always taller than determinate varieties when maturity is approximately the same. Most varieties in maturity Groups 000 to IV are indeterminate and most varieties in maturity Groups V to X are determinate.

Soybean is a self-pollinated species, propagated by seed (OECD, 2000). The papilionaceous (legume family) flower consists of a tubular calyx of five sepals, a corolla of five petals, one pistil, and nine fused stamens with a single separate posterior stamen. The stamens form a ring at the base of the stigma and elongate one day before pollination, at which time the elevated anthers form a ring around the stigma (OECD, 2000). The soybean flower stigma is receptive to pollen approximately 24 hours before anthesis and remains receptive for 48 hours after anthesis. The anthers mature in the bud and directly pollinate the stigma of the same flower. As a result, soybean is considered to be a highly self-pollinated species, with cross-pollination to adjacent plants occurring at very low frequency (0 to 6.3%) (Caviness, 1966; Ray *et al.*, 2003; Yoshimura *et al.*, 2006). Pollination typically takes place on the day the flower opens. The pollen naturally comes in contact with the stigma during the process of anthesis. Anthesis normally occurs in late morning, depending on the environmental conditions. The pollen usually remains viable for two to four hours, and no viable pollen can be detected by late afternoon. Natural or artificial cross-pollination can only take place during the short time when the pollen is viable.

See also Section A2(b)(iii) and Section A2(b)(iv).

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

There is a longstanding history of safe use and consumption as food and/or feed of conventional soybean and processed products. Soybean is grown as a commercial crop in over 35 countries. Domestication occurred as early as 1000 B.C. and is now the most widely grown oilseed crop in the world, with approximately 260 million metric tons of harvested seed produced in 2009-2010, which represented 58.5% of world oilseed seed production that year.

Soybean as a food source

In general, soyfoods can be roughly classified into four major categories (Liu, 2004).

Traditional soyfoods: they are primarily made from whole soybean. The non-fermented traditional soyfoods include soymilk, tofu, and soybean sprouts, whereas the fermented soyfoods include soybean paste (miso), soybean sauce, natto and tempeh.

Soybean oil: Soybean oil constitutes approximately 27% of global consumption of vegetable oil, and is the second largest source of vegetable oil worldwide (ASA, 2014). Refined, bleached, and deodorized (RDB) soybean oil can be further processed and utilized in the manufacture of cooking oils, shortening, margarine, mayonnaise, salad dressings and a wide variety of products that are either based entirely on fats and oils or contain fat or oil as a principal ingredient. Soybean oil is comprised primarily of five major fatty acids: saturated fatty acids;16:0 palmitic and 18:0 stearic acids, monounsaturated; 18:1 oleic acid and the PUFAs; 18:2 linoleic and 18:3 linolenic acids. These five major fatty acids have very different oxidative stabilities and chemical functionalities. Conventional soybean oil typically contains 60-65% PUFAs, mostly in the form of 18:2 linoleic acid. This composition makes soybean oil unsuitable for certain food applications since the high concentrations of PUFAs in the oil are susceptible to oxidation and degradation at high temperature resulting in off-flavors and odors. Therefore, hydrogenation of soybean oil is necessary to reduce levels of polyunsaturated fatty acids by converting them to saturated fatty acids resulting in higher stability oil suitable for a range of food uses. However, the hydrogenation process produces *trans* fatty acids (TFAs) that are linked to increased cardiovascular risk.

Soybean protein products: Soybean protein products are made from defatted soybean flakes, and include soybean flour, soybean protein concentrate, and soybean protein isolate. Soybean flour has a protein content of approximately 50% and is used mainly as an ingredient in the bakery industry. Soybean protein concentrate has a protein content of approximately 70% and is used widely in the meat industry as a key ingredient of meat alternative products such as soybean burgers and meatless “meatballs”. Soybean protein isolate has a protein content of 90%, and possesses many functional properties such as gelation and emulsification. As a result, it can be used in a wide range of food applications, including soups, sauce bases, energy bars, nutritional beverages, infant formula and dairy replacements.

Dietary supplements: Soybean is a rich source of certain phytochemicals used as dietary supplements, which include isoflavones and tocopherols. Isoflavones have been shown to inhibit the growth of cancer cells, lower cholesterol levels, and inhibit bone resorption (Messina, 1999). Tocopherols have long been recognized as a free radical scavenging antioxidant whose deficiency impairs mammalian fertility. In addition, new biological activities have been reported for the desmethyl tocopherols, such as γ-tocopherol, which possess anti-inflammatory, antineoplastic, and natriuretic functions[2] (Hensley *et al.*, 2004; Schafer *et al.*, 2003). Detailed reviews of soybean fractions and soybean-derived phytochemicals have been published (Liu, 2004).

---

[2] International Food Information Council Foundation - *Food insight*: http://www.foodinsight.org/Resources/Detail.aspx?topic=Functional_Foods_Fact_Sheet_Soy – Accessed on March 22, 2012.

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

There is only limited animal feed use and no food use for unprocessed soybeans, since they contain anti-nutrient factors, such as trypsin inhibitors and lectins. Trypsin inhibitors are heat-labile anti-nutrients that interfere with the digestion of proteins and result in decreased animal growth (Liener, 1994). Lectins are also heat-labile, and lectins can inhibit growth and cause death to animals if raw soybean is consumed (Liener, 1994). Both trypsin inhibitors and lectins are inactivated during processing of soybean protein products or SBM, and if processed appropriately, the final edible soybean fractions contain minimal levels of these anti-nutrients. Detailed information on the processing of soybean is provided in Section A2(b)(iii).

The three major soybean commodity products are beans, oil and meal (or cake). Soybeans are a good source of protein, producing more protein per unit of land than most major plant or animal sources used today. Soybean has a high content of essential amino acids, particularly lysine, leucine and isoleucine (Fehily, 2003).

Whole processed soybeans are used to produce baked soybeans, roasted soybeans, full fat soy flour and the traditional soy foods (miso, soy milk, soy sauce and tofu) (OECD, 2001).

As described in Section A2(b)(iii), soybean-derived products are used in a wide range of foods and feeds. The primary use of soybean is as a heat-processed (toasted) defatted (TD) meal for protein supplementation of animal feeds. Soybean oil is the main food ingredient derived from soybean. Soybean oil is extensively used in products such as cooking and salad oils, salad dressings, shortening and oleo margarine. Lecithin, a phosphatide removed from crude soybean oil, is used as a natural emulsifier, lubricant and stabilising agent. Glycerol, fatty acids and sterols are also derived from soybean oil (OECD, 2001).

The food industry is increasingly using products derived from soybeans as ingredients, as they can influence the physical structure, stability and texture of foods. Defatted, dehulled soybean flakes, minimally heat treated to retain solubility, are employed as a starting material for a wide variety of soybean food products (Snyder and Wilson, 2003). Soybean protein products, manufactured from defatted soybean flakes, are added to a number of meat, dairy, bakery and cereal products. Soybean protein isolate is the protein source for soybean-based infant formula. Soybean protein concentrates are incorporated in some meat products as an extender, but also in a textured form to simulate meat. Soybean flour products, either relecithinated or refatted, manufactured from defatted soybean flakes, are added to many baked goods. Soybean hulls are a high natural source of dietary fibre: they can be processed into fibre bran breads, cereals and snacks (OECD, 2001).

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

Soybean is the most widely grown oilseed in the world, with approximately 284.1 million metric tons (MMT) of harvested seed produced in 2013, which represented 56% of world oilseed production that year (ASA, 2014). Soybean is grown as a commercial crop in over 35 countries. The major producers of soybean are the U.S., Brazil, Argentina, China, India, Paraguay and Canada, which accounted for approximately 94% of the global soybean

production in 2013 (ASA, 2014); also see Table 2. Approximately 32% of the 2013 world soybean production was produced in the U.S. (ASA, 2014), with 1.580 billion bushels exported, the U.S. was second to Brazil (1.635 billion bushels) in world soybean exports. Approximately 46 MMT of soybeans were crushed in the U.S. in 2013 and used to supply the feed industry for livestock use or the food industry for edible vegetable oil and soybean protein isolates (ASA, 2014).

Table 2. World Soybean Production in 2013

Country	Production (million metric tons)
U.S.	89.5
Brazil	87.5
Argentina	54.0
China	12.2
India	11.0
Paraguay	8.1
Canada	5.2
Other	16.5

Source: Soy Stats, World Soybean Production (ASA, 2014).

It is anticipated that MON 87751 will be generally consumed in soybean products entering Australia and New Zealand. As described in Section A2(b)(iii), soybean is highly versatile and can be processed into a wide variety of food products including soybean oil, traditional soyfoods, soybean protein products, modern soyfoods, soybean-enriched foods, and functional soybean ingredients/dietary supplements. Estimates of soybean consumption are available from the WHO Global Environmental Monitoring System – Food Contamination Monitoring and Assessment Programme (GEMS/FOOD) (http://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 U.S. Canada and several other countries.

A3 The Nature of the Genetic Modification

A3(a) Method used to transform host organism

MON 87751 was developed through *Agrobacterium*-mediated transformation of soybean, based on the method described by Martinell *et al.* (2011), which allows for the generation of transformed plants without the utilization of callus. Briefly, meristem tissues were excised from the embryos of germinated conventional seed. After co-culturing with *Agrobacterium* carrying the transformation construct, the meristems were placed on selection medium

containing spectinomycin, carbenicillin disodium salt, cefotaxime sodium salt and ticarcillin disodium salt/potassium clavulanate mixture to inhibit the growth of untransformed plant cells and excess *Agrobacterium*. The meristems were then placed in media conducive to shoot development followed by a liquid overlay of selection medium and a transfer to a Jiffy Carefree propagation plug for root development. Rooted plants with normal phenotypic characteristics were selected and transferred to soil for growth and further assessment. As demonstrated in this request, the use of disarmed *A. tumefaciens* strain AB30, a designated plant pest, as the transformation vector has not imparted plant pest characteristics to MON 87751.

The R_0 plants generated through this transformation process were self-pollinated to produce R_1 seed, and the unlinked insertions of T-DNA I and T-DNA II were segregated (see Section A3(c)(i) for definition of T-DNA I and II). The *splA* scorable phenotype and *aadA* coding sequence polymerase chain reaction (PCR) analyses were used to eliminate any seeds or plants from further development that contain T-DNA II. Subsequently, R_1 plants homozygous for T-DNA I were selected for further development and their progenies were subjected to further molecular analysis, insect efficacy and phenotypic assessments. As is typical of a commercial event production and selection process, hundreds of different transformation events (regenerants) were generated in the laboratory using PV-GMIR13196 and related vectors. After careful selection and evaluation of these events in the laboratory, greenhouse and field, MON 87751 was selected as the lead event based on superior agronomic, phenotypic, and molecular characteristics. Studies on MON 87751 were initiated to further characterize the genetic insertion and the expressed products, and to establish the food and feed safety and low environmental risk compared to commercial soybean. The major development steps of MON 87751 are depicted in Figure 1.

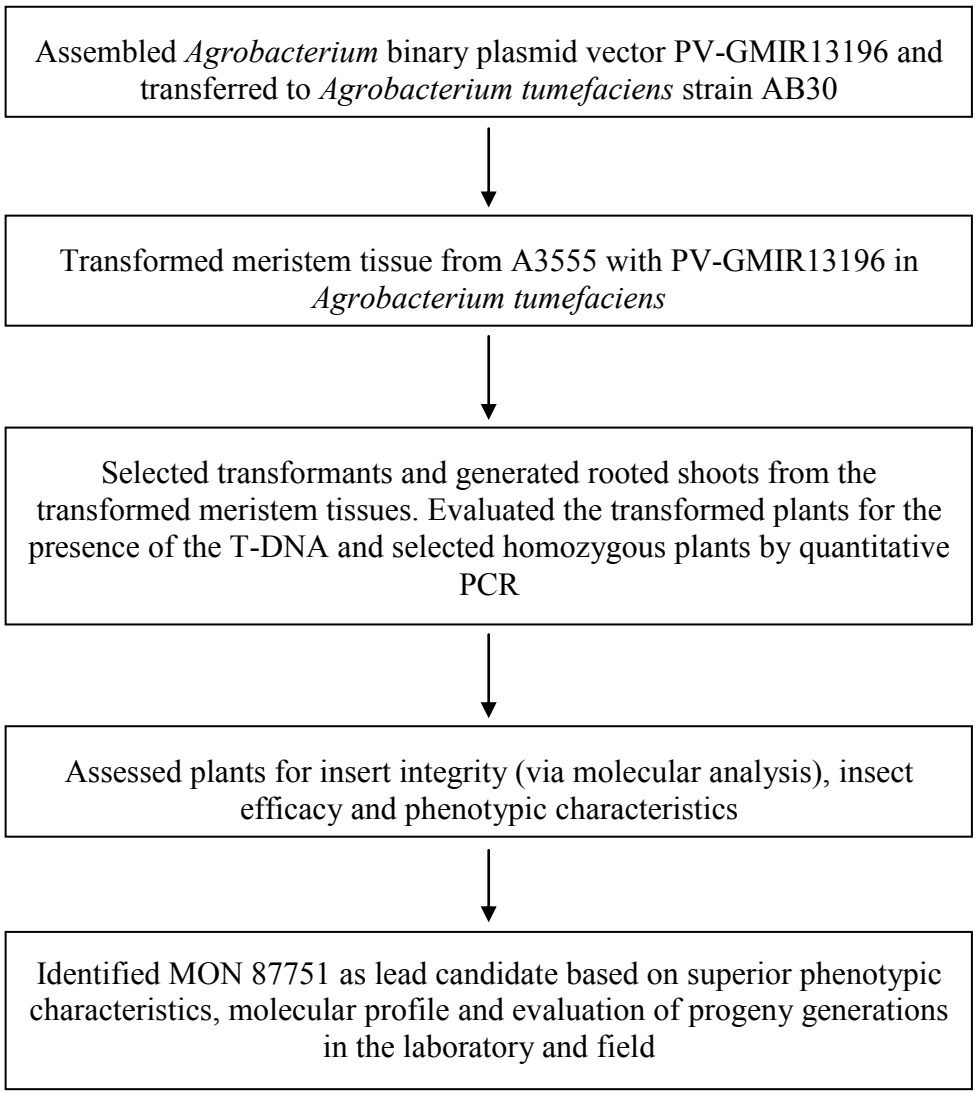


Figure 1. Schematic of the Development of MON 87751

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

A disarmed strain of *Agrobacterium tumefaciens* was the intermediate host used to transfer the T-DNA containing the *cry1A.105* and *cry2Ab2* expression cassettes from plasmid PV-GMIR13196 into soybean cells to produce MON 87751.

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

PV-GMIR13196 was used for the transformation of conventional soybean to produce MON 87751 and its plasmid map is shown in Figure 4. The elements included in this plasmid vector are described in Table 3. PV-GMIR13196 is approximately 24.5 kb and contains two separate T-DNAs, each delineated by Left and Right Border regions. The first T-DNA, designated as T-DNA I, contains the *crylA.105* and *cry2Ab2* expression cassettes. The second T-DNA, designated as T-DNA II, contains the *splA* and *aadA* expression cassettes. During transformation, both T-DNAs were inserted into the soybean genome (Section A3(a)). Subsequently, traditional breeding, segregation, selection and screening were used to isolate those plants that contain the *crylA.105* and *cry2Ab2* expression cassettes (T-DNA I) and do not contain the *splA* and *aadA* expression cassettes (T-DNA II).

splA and *aadA* act as selectable markers to allow selection of transformed plants. *aadA* encodes an aminoglycoside-modifying enzyme that confers spectinomycin and streptomycin resistance (Fling *et al.*, 1985) and allows selection of transformed tissue. *splA* encodes the sucrose phosphorylase enzyme (Piper *et al.*, 1999). When *splA* is expressed during embryo development, it interferes with sucrose metabolism, leading to a recognizable seed phenotype to provide a visual demonstration that T-DNA II is absent.

The backbone region of PV-GMIR13196, located outside both of the T-DNAs, contains two origins of replication for maintenance of the plasmid vector in bacteria (*ori-pRi*, *ori-pBR322*), a bacterial selectable marker gene (*nptII*), and a coding sequence for repressor of primer (ROP) protein for the maintenance of the plasmid vector copy number in *Escherichia coli* (*E. coli*). A description of the genetic elements and their prefixes (e.g. B-, P-, TS-, CS-, T-, and OR-) in PV-GMIR13196 is provided in Table 3.

The *cry1A.105* expression cassette, encodes a 142 kDa CrylA.105 protein consisting of a single polypeptide of 1265 amino acids (Figure 2). The *cry1A.105* coding sequence is the coding sequence from *B. thuringiensis* that encodes the CrylA.105 protein. The presence of CrylA.105 protein in soybean provides insect resistance.

The *cry2Ab2* expression cassette, encodes a 79 kDa Cry2Ab2 protein consisting of a single polypeptide of 713 amino acids (Figure 3). The *cry2Ab2* coding sequence is the coding sequence from *B. thuringiensis* that encodes the Cry2Ab2 protein (Donovan, 1991). The presence of Cry2Ab2 protein in soybean provides insect resistance.

The *crylA.105* coding sequence in T-DNA I is under the regulation of the *RbcS4* promoter, *RbcS4* chloroplast targeting sequence and the *Pt1* 3' untranslated region. The *RbcS4* promoter is the promoter for the *rbcS* gene family of *A. thaliana* (De Almeida *et al.*, 1989; Krebbers *et al.*, 1988), which functions to direct transcription in plant cells. The targeting sequence from *A. thaliana rbcS* gene family encodes the small subunit *ats1A* (Wong *et al.*,

1992) that directs transport of the protein to the chloroplast The *Pt1* 3' untranslated region is the 3' untranslated region of the *Pt1* gene from *M. truncatula* encoding phosphate transporter that directs polyadenylation of mRNA (Liu *et al.*, 1998).

The *cry2Ab2* coding sequence in T-DNA I is under the regulation of the *Act2* promoter, *CTP2* targeting sequence and the *Mt* 3' untranslated region. The *Act2* promoter is the promoter for the *act2* gene of *A. thaliana* (An *et al.*, 1996), which functions to direct transcription in plant cells. The *CTP2* targeting sequence is the targeting sequence of the *ShkG* gene from *A. thaliana* encoding the EPSPS transit peptide region (Herrmann, 1995; Klee *et al.*, 1987), which functions to direct transport of the protein to the chloroplast. The *Mt* 3′ untranslated region is the 3' untranslated region of the *Mt* gene from *O. sativa* encoding metallothionein-like protein that directs polyadenylation of mRNA (Hunt, 1994).

T-DNA II contains the *splA* coding sequence under the regulation of the *Usp* promoter, *Usp* leader, *Usp* enhancer, and the *nos* 3' untranslated region. The *Usp* promoter is the promoter consisting leader, promoter, and enhancer sequences from *V. faba* (broad bean) encoding a seed protein (Bäumlein *et al.*, 1991), which functions to direct transcription in plant cells. The *nos* 3' untranslated region is the 3' untranslated region from *A. tumefaciens* pTi nopaline synthase (*nos*) gene encoding NOS that directs polyadenylation (Bevan *et al.*, 1983; Fraley *et al.*, 1983).

T-DNA II also contains the *aadA* coding sequence under the regulation of the *FMV* enhancer, *EF-1α* promoter, the *CTP2* targeting sequence, and the *T-E9* untranslated region. The *EF-1α* promoter is the promoter consisting leader, promoter, and intron sequences from *A. thaliana* encoding elongation factor *EF-1α* (Axelos *et al.*, 1989), which functions to direct transcription in plant cells. The *FMV* enhancer is the enhancer sequence of the 35S RNA of figwort mosaic virus (FMV) (Richins *et al.*, 1987), which enhances transcription in most plant cells (Rogers, 2000). The *CTP2* targeting sequence is the targeting sequence of the *ShkG* gene from *A. thaliana* encoding the EPSPS transit peptide region (Klee *et al.*, 1987; Herrmann, 1995), which functions to direct transport of the protein to the chloroplast. The *E9* 3' untranslated region is the 3' untranslated region from *P. sativum* (pea) *rbcS* gene family encoding the small subunit of ribulose bisphosphate carboxylase protein (Coruzzi *et al.*, 1984), which functions to direct polyadenylation of the mRNA.

PV-GMIR13196 contains Left and Right Border regions (Figure 4 and Table 3) 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 soybean genome. Because PV-GMIR13196 is a 2T-DNA vector, it contains two Left Border regions and two Right Border regions, where one border region set flanks T-DNA I and the other border region set flanks T-DNA II. As demonstrated in this request, the use of genetic elements from *A. tumefaciens*, a designated plant pest, has not imparted plant pest characteristics to MON 87751.

Genetic elements that exist outside of the T-DNA border regions are those that are essential for the maintenance or selection of PV- GMIR13196 in bacteria and are referred to as

plasmid backbone. The origin of replication, *ori pRi*, is required for the maintenance of the plasmid in *Agrobacterium* and is derived from the broad host plasmid *pRi* (Ye *et al.*, 2011). The origin of replication, *ori-pBR322*, is required for the maintenance of the plasmid in *E. coli* and is derived from the plasmid vector pBR322 (Sutcliffe, 1979). 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 *rrn* promoter is the promoter for the ribosomal RNA operon from *A. tumefacien*s (Bautista-Zapanta *et al.*, 2002). The selectable marker *nptII* is the coding sequence for an enzyme from transposon Tn5 that confers neomycin and kanamycin resistance (Fraley *et al.*, 1983) in *E. coli* and *Agrobacterium* during molecular cloning. Because these elements are outside the border regions, they are not expected to be transferred into the soybean genome. The absence of the backbone and other unintended plasmid sequence in MON 87751 was confirmed by sequencing and bioinformatic analyses (see Section A3(d)(ii)).

Table 3. Summary of Genetic Elements in PV–GMIR13196

Genetic Element	Location in Plasmid	Function (Reference)
T-DNA I		
B[1]-Right Border Region	1-285	DNA region from *A. tumefaciens* containing the right border sequence used for transfer of the T-DNA (Depicker *et al.*, 1982; Zambryski *et al.*, 1982)
Intervening Sequence	286-337	Sequence used in DNA cloning
P[2]-*Act2*	338-1545	Promoter, leader and intron sequences from the *act2* gene of *A. thaliana* (An *et al.*, 1996) that directs transcription in plant cells
Intervening Sequence	1546-1555	Sequence used in DNA cloning
TS[3]-*CTP2*	1556-1783	Targeting sequence of the *ShkG* gene from *A. thaliana* encoding the EPSPS transit peptide region that directs transport of the protein to the chloroplast (Klee *et al.*, 1987; Herrmann, 1995)
Intervening Sequence	1784-1792	Sequence used in DNA cloning
CS[4]-*cry2Ab2*	1793-3697	Coding sequence for the Cry2Ab2 protein of *B. thuringiensis* that provides insect resistance (Donovan, 1991)
Intervening Sequence	3698-3700	Sequence used in DNA cloning
T[5]-*Mt*	3701-4000	3' UTR sequence from *O. sativa* (rice) *Mt* gene encoding metallothionein-like protein that directs polyadenylation of mRNA (Hunt, 1994)
Intervening Sequence	4001-4045	Sequence used in DNA cloning
P-*RbcS4*	4046-5768	Promoter and leader sequences from *A. thaliana rbcS* gene family encoding small subunit *ats1A* (Krebbers *et al.*, 1988; De Almeida *et al.*, 1989) that directs transcription in plant cells
TS-*RbcS4*	5769-6032	Targeting sequence from *A. thaliana rbcS* gene family encoding small subunit *ats1A* (Wong *et al.*, 1992) that directs transport of the protein to the chloroplast

Table 3. Summary of Genetic Elements in PV-GMIR13196 (continued)

Genetic Element	Location in Plasmid	Function (Reference)
CS-*cry1A.105*	6033-9566	Coding sequences for the Cry1Ab, Cry1F, and Cry1Ac proteins of *B. thuringiensis* to produce a chimeric protein that provides insect resistance (Monsanto unpublished data)
Intervening Sequence	9567-9569	Sequence used in DNA cloning
T-*Pt1*	9570-9969	3' UTR sequence from *M. truncatula* PT1 gene encoding phosphate transporter that directs polyadenylation of mRNA (Liu *et al.*, 1998)
Intervening Sequence	9970-10088	Sequence used in DNA cloning
B-Left Border Region	10089-10530	DNA region from *A. tumefaciens* containing the left border sequence used for transfer of the T-DNA (Barker *et al.*, 1983)
Backbone		
Intervening Sequence	10531-10739	Sequence used in DNA cloning
CS-*nptII*	10740-11534	Coding sequence of the *neo* gene from transposon Tn5 of *E. coli* encoding neomycin phosphotransferase II (Beck *et al.*, 1982) that confers neomycin and kanamycin resistance (Fraley *et al.*, 1983)
P-*rrn*	11535-11759	Promoter of the ribosomal RNA operon from *A. tumefaciens* (Bautista-Zapanta *et al.*, 2002)
Intervening Sequence	11760-11835	Sequence used in DNA cloning
OR[6]-*ori-pBR322*	11836-12424	Origin of replication from plasmid pBR322 for maintenance of plasmid in *E. coli* (Sutcliffe, 1979)
Intervening Sequence	12425-12851	Sequence used in DNA cloning
CS-*rop*	12852-13043	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	13044-13231	Sequence used in DNA cloning
OR-*ori-pRi*	13232-17345	Origin of replication from plasmid pRi for maintenance of plasmid in *Agrobacterium* (Ye *et al.*, 2011)
Intervening Sequence	17346-17352	Sequence used in DNA cloning

Table 3. Summary of Genetic Elements in PV-GMIR13196 (continued)

Genetic Element	Location in Plasmid	Function (Reference)
T-DNA II		
B-Left Border Region	17353-17671	DNA region from *A. tumefaciens* containing the left border sequence used for transfer of the T-DNA (Barker *et al.*, 1983)
Intervening Sequence	17672-17703	Sequence used in DNA cloning
T-*nos*	17704-17956	3' UTR sequence of the nopaline synthase (*nos*) gene from *A. tumefaciens* pTi encoding NOS that directs polyadenylation (Bevan *et al.*, 1983; Fraley *et al.*, 1983)
Intervening Sequence	17957-17972	Sequence used in DNA cloning
CS-*splA*	17973-19430	Coding sequence of the *splA* gene from *A. tumefaciens* strain C58 encoding the sucrose phosphorylase protein that catalyzes the conversion of sucrose to fructose and glucose-1-phosphate (Piper *et al.*, 1999)
Intervening Sequence	19431-19442	Sequence used in DNA cloning
P-*Usp*	19443-20621	5' UTR leader, promoter, and enhancer sequence from *V. faba* (broad bean) *Usp* gene encoding a seed protein that directs transcription in plant cells (Bäumlein *et al.*, 1991)
Intervening Sequence	20622-20672	Sequence used in DNA cloning
T-*E9*	20673-21315	3' UTR sequence from *P. sativum* (pea) *rbcS* gene family encoding the small subunit of ribulose bisphosphate carboxylase protein (Coruzzi *et al.*, 1984) that directs polyadenylation of the mRNA
Intervening Sequence	21316-21330	Sequence used in DNA cloning
CS-*aadA*	21331-22122	Bacterial coding sequence for an *aminoglycoside-modifying enzyme, 3''(9)-O-nucleotidyltransferase* from the transposon Tn7 (Fling *et al.*, 1985) that confers spectinomycin and streptomycin resistance

Table 3. Summary of Genetic Elements in PV-GMIR13196 (continued)

Genetic Element	Location in Plasmid	Function (Reference)
TS-*CTP2*	22123-22350	Targeting sequence of the *ShkG* gene from *A. thaliana* encoding the EPSPS transit peptide region that directs transport of the protein to the chloroplast (Klee *et al.*, 1987; Herrmann, 1995)
Intervening Sequence	22351-22359	Sequence used in DNA cloning
P-*EF-1α*	22360-23507	Promoter, leader, and intron sequences of the *EF-1α* gene from *A. thaliana* encoding elongation factor EF-1α that directs transcription in plant cells (Axelos *et al.*, 1989)
Intervening Sequence	23508-23530	Sequence used in DNA cloning
E[7]-*FMV*	23531-24067	Enhancer from the 35S RNA of figwort mosaic virus (Richins *et al.*, 1987) that enhances transcription in most plant cells (Rogers, 2000)
Intervening Sequence	24068-24117	Sequence used in DNA cloning
B-Right Border Region	24118-24474	DNA region from *A. tumefaciens* containing the right border sequence used for transfer of the T-DNA (Zambryski *et al.*, 1982; Depicker *et al.*, 1982)
Backbone		
Intervening Sequence	24475-24489	Sequence used in DNA cloning

[1]B, Border

[2]P, Promoter

[3]TS, Targeting Sequence

[4]CS, Coding Sequence

[5]T, Transcription Termination Sequence

[6]OR, Origin of Replication

[7]E, Enhancer

```
   1 MASSMLSSAT MVASPAQATM VAPFNGLKSS AAFPATRKAN NDITSITSNG GRVNCMQVWP
  61 PIGKKKFETL SYLPDLTDSG GRVNCMQAMD NNPNINECIP YNCLSNPEVE VLGGERIETG
 121 YTPIDISLSL TQFLLSEFVP GAGFVLGLVD IIWGIFGPSQ WDAFLVQIEQ LINQRIEEFA
 181 RNQAISRLEG LSNLYQIYAE SFREWEADPT NPALREEMRI QFNDMNSALT TAIPLFAVQN
 241 YQVPLLSVYV QAANLHLSVL RDVSVFGQRW GFDAATINSR YNDLTRLIGN YTDHAVRWYN
 301 TGLERVWGPD SRDWIRYNQF RRELTLTVLD IVSLFPNYDS RTYPIRTVSQ LTREIYTNPV
 361 LENFDGSFRG SAQGIEGSIR SPHLMDILNS ITIYTDAHRG EYYWSGHQIM ASPVGFSGPE
 421 FTFPLYGTMG NAAPQQRIVA QLGQGVYRTL SSTLYRRPFN IGINNQQLSV LDGTEFAYGT
 481 SSNLPSAVYR KSGTVDSLDE IPPQNNNVPP RQGFSHRLSH VSMFRSGFSN SSVSIIRAPM
 541 FSWIHRSAEF NNIIASDSIT QIPLVKAHTL QSGTTVVRGP GFTGGDILRR TSGGPFAYTI
 601 VNINGQLPQR YRARIRYAST TNLRIYVTVA GERIFAGQFN KTMDTGDPLT FQSFSYATIN
 661 TAFTFPMSQS SFTVGADTFS SGNEVYIDRF ELIPVTATLE AEYNLERAQK AVNALFTSTN
 721 QLGLKTNVTD YHIDQVSNLV TYLSDEFCLD EKRELSEKVK HAKRLSDERN LLQDSNFKDI
 781 NRQPERGWGG STGITIQGGD DVFKENYVTL SGTFDECYPT YLYQKIDESK LKAFTRYQLR
 841 GYIEDSQDLE IYSIRYNAKH ETVNVPGTGS LWPLSAQSPI GKCGEPNRCA PHLEWNPDLD
 901 CSCRDGEKCA HHSHHFSLDI DVGCTDLNED LGVWVIFKIK TQDGHARLGN LEFLEEKPLV
 961 GEALARVKRA EKKWRDKREK LEWETNIVYK EAKESVDALF VNSQYDQLQA DTNIAMIHAA
1021 DKRVHSIREA YLPELSVIPG VNAAIFEELE GRIFTAFSLY DARNVIKNGD FNNGLSCWNV
1081 KGHVDVEEQN NQRSVLVVPE WEAEVSQEVR VCPGRGYILR VTAYKEGYGE GCVTIHEIEN
1141 NTDELKFSNC VEEEIYPNNT VTCNDYTVNQ EEYGGAYTSR NRGYNEAPSV PADYASVYEE
1201 KSYTDGRREN PCEFNRGYRD YTPLPVGYVT KELEYFPETD KVWIEIGETE GTFIVDSVEL
1261 LLMEE
```

Figure 2. Deduced Amino Acid Sequence of the RbcS4 Targeting Sequence and Cry1A.105 Protein

The chloroplast transit peptide RbcS4 is underlined. Accumulation of the Cry1A.105 protein is targeted to the chloroplasts using cleavable RbcS4, the transit peptide of the *A. thaliana* small subunit *ats1A* protein. The amino acid sequence of the Cry1A.105 protein and RbcS4 targeting sequence was deduced from the full-length coding nucleotide sequence present in PV-GMIR13196.

```
   1 MAQVSRICNG VQNPSLISNL SKSSQRKSPL SVSLKTQQHP RAYPISSSWG LKKSGMTLIG
  61 SELRPLKVMS SVSTACMLAM DNSVLNSGRT TICDAYNVAA HDPFSFQHKS LDTVQKEWTE
 121 WKKNNHSLYL DPIVGTVASF LLKKVGSLVG KRILSELRNL IFPSGSTNLM QDILRETEKF
 181 LNQRLNTDTL ARVNAELTGL QANVEEFNRQ VDNFLNPNRN AVPLSITSSV NTMQQLFLNR
 241 LPQFQMQGYQ LLLLPLFAQA ANLHLSFIRD VILNADEWGI SAATLRTYRD YLKNYTRDYS
 301 NYCINTYQSA FKGLNTRLHD MLEFRTYMFL NVFEYVSIWS LFKYQSLLVS SGANLYASGS
 361 GPQQTQSFTS QDWPFLYSLF QVNSNYVLNG FSGARLSNTF PNIVGLPGST TTHALLAARV
 421 NYSGGISSGD IGASPFNQNF NCSTFLPPLL TPFVRSWLDS GSDREGVATV TNWQTESFET
 481 TLGLRSGAFT ARGNSNYFPD YFIRNISGVP LVVRNEDLRR PLHYNEIRNI ASPSGTPGGA
 541 RAYMVSVHNR KNNIHAVHEN GSMIHLAPND YTGFTISPIH ATQVNNQTRT FISEKFGNQG
 601 DSLRFEQNNT TARYTLRGNG NSYNLYLRVS SIGNSTIRVT INGRVYTATN VNTTTNNDGV
 661 NDNGARFSDI NIGNVVASSN SDVPLDINVT LNSGTQFDLM NIMLVPTNIS PLY
```

Figure 3. Deduced Amino Acid Sequence of the CTP2 Targeting Sequence and Cry2Ab2 Protein

The transit peptide CTP2 is underlined. Accumulation of the Cry2Ab2 protein is targeted to the chloroplasts using cleavable CTP2, the transit peptide of the *A. thaliana* EPSPS protein. The amino acid sequence of the Cry2Ab2 protein and CTP2 targeting sequence was deduced from the full-length coding nucleotide sequence present in PV-GMIR13196.

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

Plasmid map with locations of genetic elements are shown in Figure 4.

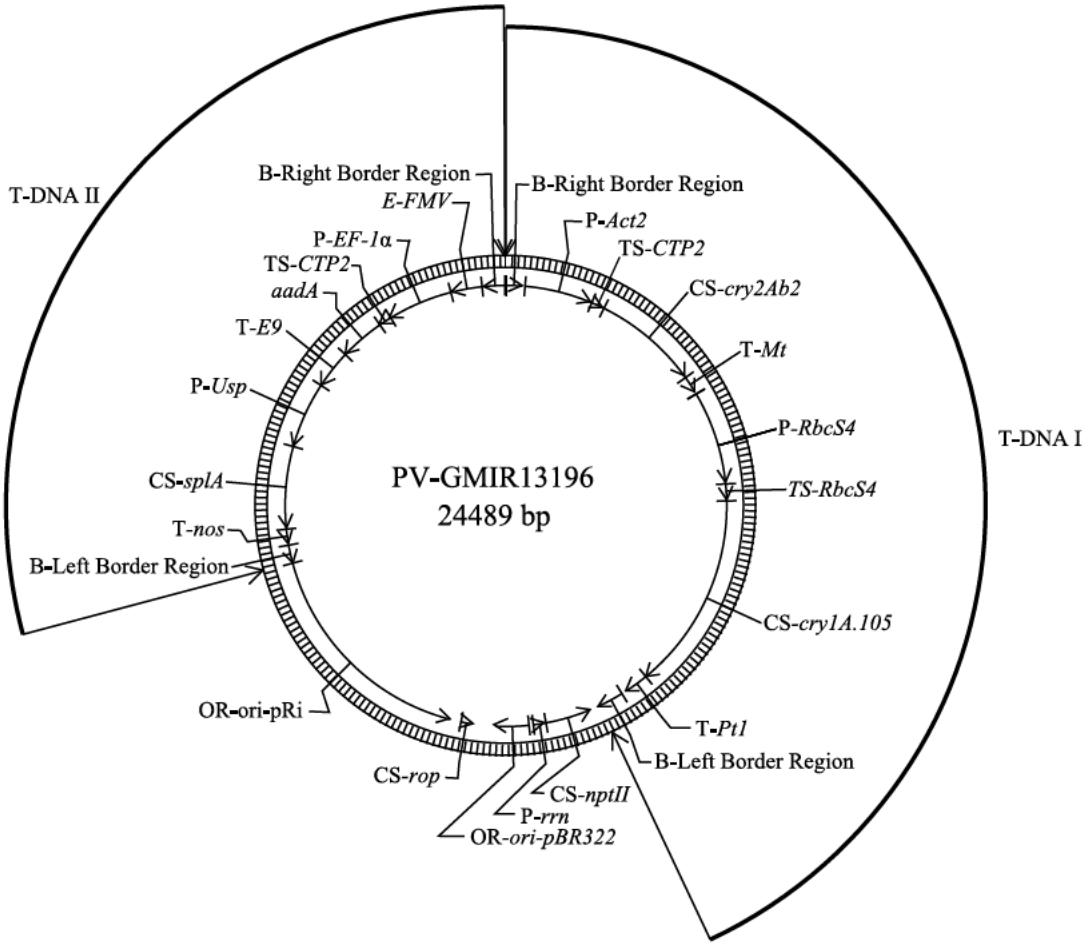


Figure 4. Circular Map of Plasmid Vector PV–GMIR13196

A circular map of the plasmid vector PV-GMIR13196 used to develop MON 87751 is shown. PV-GMIR13196 contains two T-DNAs, designated as T-DNA I and T-DNA II. Genetic elements are shown on the exterior of the map

A3(d)(i) Molecular characterisation including identification of GM elements

MON 87751 was produced by *Agrobacterium tumefaciens*–mediated transformation of soybean tissue using the transformation plasmid vector PV–GMIR13196 (which is illustrated in Figure 4 and whose elements are described in Table 3). PV–GMIR13196 contains two separate T–DNAs that are each delineated by left and right border regions. The first T–DNA, designated as T–DNA I, contains the Cry1A.105 coding sequence regulated by the *RbcS4* promoter and *Pt1* terminator; and the Cry2Ab2 coding sequence regulated by the *Act2* promoter and *Mt* terminator. The second T–DNA, designated as T–DNA II, contains the *aadA* coding sequence regulated by the *FMV* enhancer, *EF–1α* promoter and *E9* terminator; and the *splA* coding sequence regulated by the *Usp* promoter and *nos* terminator. During transformation, both T–DNAs were inserted into the soybean genome. Subsequently, traditional breeding, segregation, selection and screening were used to isolate those plants that contain the Cry1A.105 and Cry2Ab2 expression cassettes (T–DNA I) and do not contain the *aadA* and *splA* expression cassettes (T–DNA II), resulting in the production of marker–free, MON 87751.

In this study, bioinformatic analysis of next–generation sequence data (comprehensively covering the genome) was used to characterize the insert number of any DNA sequences from the transformation plasmid vector PV–GMIR13196. The data indicated that MON 87751 contains a single integration locus and a single copy of PV-GMIR13196 T–DNA I, and is devoid of plasmid backbone or T–DNA II sequence. Directed sequence data (generated from PCRs targeted at a specific locus) was used to demonstrate that the single DNA insert in MON 87751 contains the T–DNA I from the transformation plasmid vector PV–GMIR13196 and confirmed the integrity of the Cry1A.105 and Cry2Ab2 expression cassettes in the T–DNA I insert. The directed sequencing data also identified the 5' and 3' insert–to–flank junctions in MON 87751 and characterized the insertion site in conventional soybean. Furthermore, bioinformatic analysis of next–generation sequence data (comprehensively covering the genome) was used to evaluate the stability of the DNA insert and demonstrated that the DNA insert in MON 87751 was stably maintained across five breeding generations.

The purpose of this study was to characterize the inserted DNA in MON 87751. Bioinformatic evaluation of next–generation sequence data (comprehensively covering the genome) and directed sequence data (targeted at a specific locus) was used to demonstrate the following molecular characteristics of MON 87751:

- The insert and copy number of DNA sequences from the transformation plasmid vector PV–GMIR13196 (by analysis of next–generation sequence data)
- The presence or absence of backbone (and T–DNA II, or other unintended sequence) from plasmid PV–GMIR13196 (by analysis of next–generation sequence data)
- The sequence and intactness of the inserted DNA (by analysis of directed sequencing data)

- The integrity and organization of the insertion site compared to the locus in the near isogenic conventional control (by analysis of directed sequencing data)
- The stability of the inserted DNA across five breeding generations (by analysis of next–generation sequence data)

Characterization of the DNA insert number in MON 87751 was conducted using a combination of Next–Generation Sequencing technologies (NGS) and Junction Sequence Analysis (JSA) bioinformatics (Kovalic *et al.*, 2012; DuBose *et al.*, 2013). In addition, directed sequencing (locus specific PCR and DNA sequence analyses) was used to obtain the sequences of the DNA insert and the adjacent flanking DNA.

The NGS/JSA strategy to determine the number of insertion loci and the number of copies of the integrated plasmid DNA was designed to ensure that all transgenic segments would be identified. Genomic DNA from MON 87751 and the conventional control were used to generate short (~100 bp) randomly distributed sequence fragments (sequencing reads) in sufficient numbers to ensure comprehensive coverage of the soybean genomes. The depth of coverage (the median number of times each base of the genome is independently sequenced) was ≥75× for each genome. It has previously been demonstrated that 75× coverage of the soybean genome is adequate to provide comprehensive coverage and ensure detection of inserted DNA (Kovalic *et al.*, 2012). The level of sensitivity of this method was demonstrated by detection of a positive control spiked at 1/10th copy –per–genome equivalent.

Taken together, the characterisation of the genetic modification in MON 87751 demonstrates that a single copy of the intended T-DNA I was stably integrated at a single locus of the soybean genome and that no plasmid backbone or T-DNA II sequences are present in MON 87751.

A schematic representation of the NGS/JSA methodology and the basis of the characterization using NGS/JSA and directed sequencing is presented in Figure 5 below (Kovalic et al. 2012).

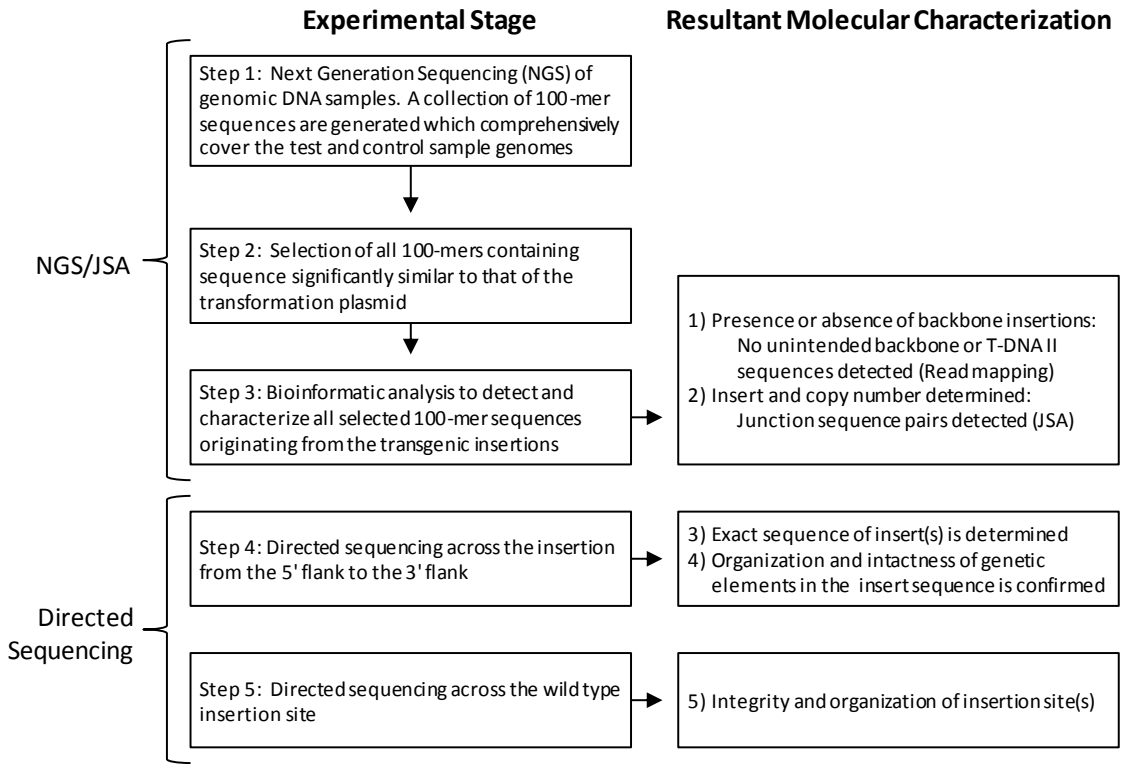


Figure 5. Molecular Characterization using Sequencing and Bioinformatics

Genomic DNA from the test and the conventional control was sequenced using technology that produces a set of short, randomly distributed sequence reads (each approximately 100 bp long) that comprehensively cover both genomes (Step 1). Utilizing 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 and T–DNA II 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 characterization of the inserted DNA and insertion site(s).

The number of DNA inserts was determined by analyzing sequences for novel junctions. The junctions of the DNA insert and the flanking DNA are unique for each insertion, an example is shown in Figure 6 below (Kovalic *et al.*, 2012). Therefore, insertion sites can be recognized by analyzing for sequence reads containing such junctions.

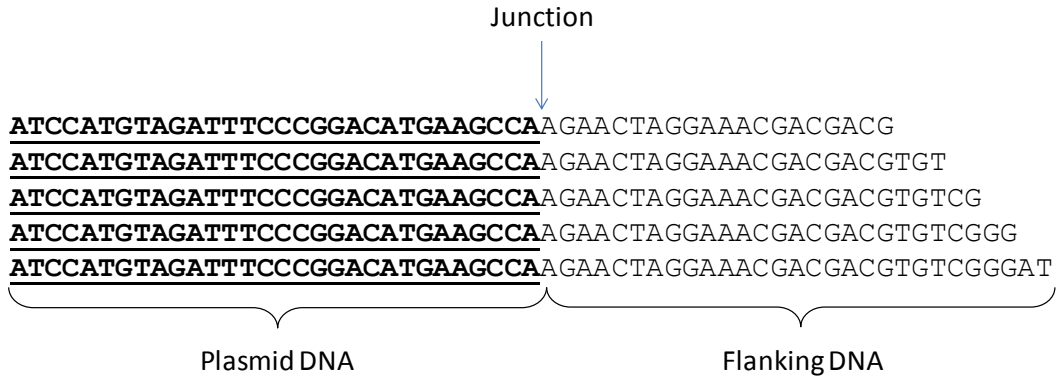


Figure 6. Junctions and Junction Sequences

Depicted above are five junction sequences with text formatted and labeled to indicate the plasmid/flanking DNA portions of the sequences and with the junction point indicated. Junctions are detected by examining the NGS data for reads having portions of plasmid sequences that span less than the full read. Detected junctions are typically characteristic of plasmid insertions in the genome.

Each insertion will produce two unique junction sequence classes characteristic of the genomic locus, one at the 5' end of the insert (illustrated in Figure 7 below, and named junction sequence class A, or JSC–A, in this case) and one at the 3' end of the insert (similarly class B, or JSC–B) (Kovalic *et al.*, 2012).

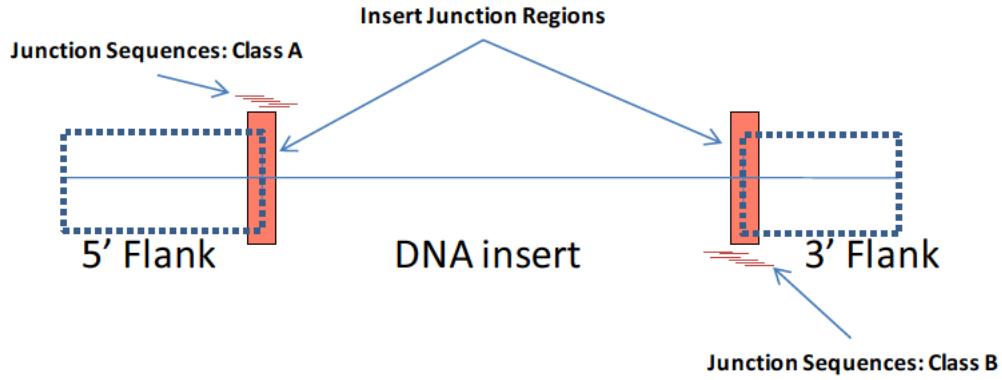


Figure 7. Two Unique Junction Sequence Classes are Produced by the Insertion of a Single Plasmid Region

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 (Class A at the 5' end and Class B at the 3' end) are represented.

By evaluating the number and the sequences of the unique junction classes detected, the number of plasmid sequence insertions into the genome and the T–DNA copy number can be determined. For a single copy of a T–DNA insert two junction sequence classes are expected, each originating from the ends of the insert, both containing portions of T–DNA and plant genomic flanking sequence. Through mapping the sequence reads obtained from the MON 87751 and conventional control to the transformation plasmid, the presence or absence of inserted backbone sequence and T-DNA II can be determined. Results are described in Section A3(d)(ii).

Directed sequencing (locus–specific PCR and DNA sequencing analyses, Figure 5, step 4) complements the NGS/JSA analyses. Sequencing of the insert and flanking DNA determines the complete sequence of the insert and flanks; if the sequence of the insert is identical to the corresponding sequence in the plasmid vector and if each genetic element in the insert is intact. Furthermore, the genomic organization at the insertion site was assessed by comparing the insert and flanking sequence to the sequence of the insertion site in conventional soybean (Figure 5, step 5).

The stability of the T–DNA present in MON 87751 across multiple breeding generations was evaluated by NGS/JSA analyses. Sequence reads of genomic DNA from five breeding generations of MON 87751 were mapped to the transformation plasmid and assayed for the

number and sequences of all unique junction classes as described above. This information was used to determine the number and identity of the DNA inserts. As described earlier, for a single copy T–DNA insert, two junction sequence classes are expected. In the case of an event where a single locus is stably inherited over multiple breeding generations, two identical junction sequence classes would be detected in all the breeding generations tested. Results are described in Section A3(f). [redacted]

A3(d)(ii) Determination of number of insertion sites, and copy number

The number of inserted DNA sequences from PV–GMIR13196 in MON 87751 was assessed by performing NGS/JSA on MON 87751 genomic DNA. A single genomic DNA insertion is expected to produce two junction sequence classes and any additional integration sites would produce additional junction sequence classes. A plasmid map of PV–GMIR13196 is shown in Figure 4. A schematic representation of the insert and flanking sequences in MON 87751 is shown in Figure 8. The generations used in this study are depicted in the breeding history diagram shown in Figure 9. The analysis of the R3 generation is described below with sequencing data presented in Figure 10; the other generations, Figure 9, that were used in the generational stability analysis are shown with the results of their analysis in Section A3(f).

Genomic DNA from MON 87751 and the conventional control were isolated and prepared for sequencing (Illumina, TruSeq library protocol). These genomic DNA libraries were used to generate short (~100 bp) randomly distributed sequence fragments (sequencing reads) in sufficient numbers to ensure comprehensive coverage of the genomes. The depth of coverage (the median number of times each base of the genome is independently sequenced) was ≥75× for each sample, the median coverage of the MON 87751 (R3) was 75× and the A3555 control was sequenced to 126×. It had previously been demonstrated that 75× coverage of the soybean genome is adequate to provide comprehensive coverage of the soybean genome (Kovalic *et al.*, 2012). The level of sensitivity achieved in this study was demonstrated by detection of a positive control spiked–in at 1/10th copy –per–genome equivalent as described below. The sequencing coverage was sufficient in all samples which ensured comprehensive JSA data analytics.

To confirm sufficient sequence coverage in both MON 87751 (R3) and the control, the 100–mer sequence reads from MON 87751 (R3) and the control were analyzed to determine the coverage of a known single–copy endogenous gene, (*Glycine max* lectin (*Le1*), GenBank accession.version: K00821.1). The analysis showed that *Le1* was covered by 100–mers at ≥75×median depth in each sample. These data established that the sequencing produced in the study was adequate for a comprehensive analysis of the sample genomes including the inserted T–DNA I sequence in MON 87751.

In order to confirm the method's ability to detect any sequences derived from the transformation plasmid, a plasmid spike–in control was analyzed as described in Section 3.12.5. The resultant sequence reads were mapped to the know plasmid sequence and the mapped sequence showed 100% nucleotide identity over 100% of the known sequence of

PV–GMIR13196 at one genome equivalent. This result demonstrates that all bases of the transformation plasmid are observed by the sequencing and bioinformatics performed in this study. Also, observed coverage was complete and correct down to a level of at least 1/10th genomic equivalent (100% coverage at 100% identity for the 1/10th copy spike sample and hence a detection limit of at least 1/10th copy number was achieved for the plasmid DNA sequences in this study.

Table 4. Summary of Genetic Elements in MON 87751

Genetic Element[1]	Location in Sequence[2]	Function (Reference)
Flanking DNA	1-1334	DNA adjacent to 5' of the DNA insert
B[3]-Right Border Region[r1]	1335-1404	DNA region from *A. tumefaciens* containing the right border sequence used for transfer of the T DNA (Depicker *et al.*, 1982; Zambryski *et al.*, 1982)
Intervening Sequence	1405-1456	Sequence used in DNA cloning
P[4]-*Act2*	1457-2664	Promoter, leader and intron sequences from the *act2* gene of *A. thaliana* (An *et al.*, 1996) that directs transcription in plant cells
Intervening Sequence	2665-2674	Sequence used in DNA cloning
TS[5]-*CTP2*	2675-2902	Targeting sequence of the *ShkG* gene from *A. thaliana* encoding the EPSPS transit peptide region that directs transport of the protein to the chloroplast (Herrmann, 1995; Klee *et al.*, 1987)
Intervening Sequence	2903-2911	Sequence used in DNA cloning
CS[6]-*cry2Ab2*	2912-4816	Coding sequence for the Cry2Ab2 protein of *B. thuringiensis* that provides insect resistance (Donovan, 1991)
Intervening Sequence	4817-4819	Sequence used in DNA cloning
T[7]-*Mt*	4820-5119	3' UTR sequence from *O. sativa* (rice) *Mt* gene encoding metallothionein like protein that directs polyadenylation of mRNA (Hunt, 1994)
Intervening Sequence	5120-5164	Sequence used in DNA cloning
P-*RbcS4*	5165-6887	Promoter and leader sequences from *A. thaliana rbcS* gene family encoding small subunit *ats1A* (De Almeida *et al.*, 1989; Krebbers *et al.*, 1988) that directs transcription in plant cells
TS-*RbcS4*	6888-7151	Targeting sequence from *A. thaliana rbcS* gene family encoding small subunit *ats1A* (Wong *et al.*, 1992) that directs transport of the protein to the chloroplast
CS-*cry1A.105*	7152-10685	Coding sequences for the Cry1Ab, Cry1F, and Cry1Ac proteins of *B. thuringiensis* to produce a chimeric protein that provides insect resistance (Monsanto unpublished data)
Intervening Sequence	10686-10688	Sequence used in DNA cloning

Table 4. Summary of Genetic Elements in MON 87751 (continued)

Genetic Element[1]	Location in Sequence[2]	Function (Reference)
T-*Pt1*	10689-11088	3' UTR sequence from *M. truncatula PT1* gene encoding phosphate transporter that directs polyadenylation of mRNA (Liu *et al.*, 1998)
Intervening Sequence	11089-11207	Sequence used in DNA cloning
B-Left Border Region[r1]	11208-11453	DNA region from *A. tumefaciens* containing the left border sequence used for transfer of the T-DNA (Barker *et al.*, 1983)
Flanking DNA	11454-12640	DNA adjacent to 3' of the DNA 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 87751 and adjacent DNA.
[3]B, Border
[4]P, Promoter
[5] TS, Targeting Sequence
[6]CS, Coding Sequence
[7] T, Transcriptional Terminator
[r1]Superscript in Left and Right Border Regions indicate that the sequence in MON 87751 was truncated compared to the sequences in PV-GMIR13196.

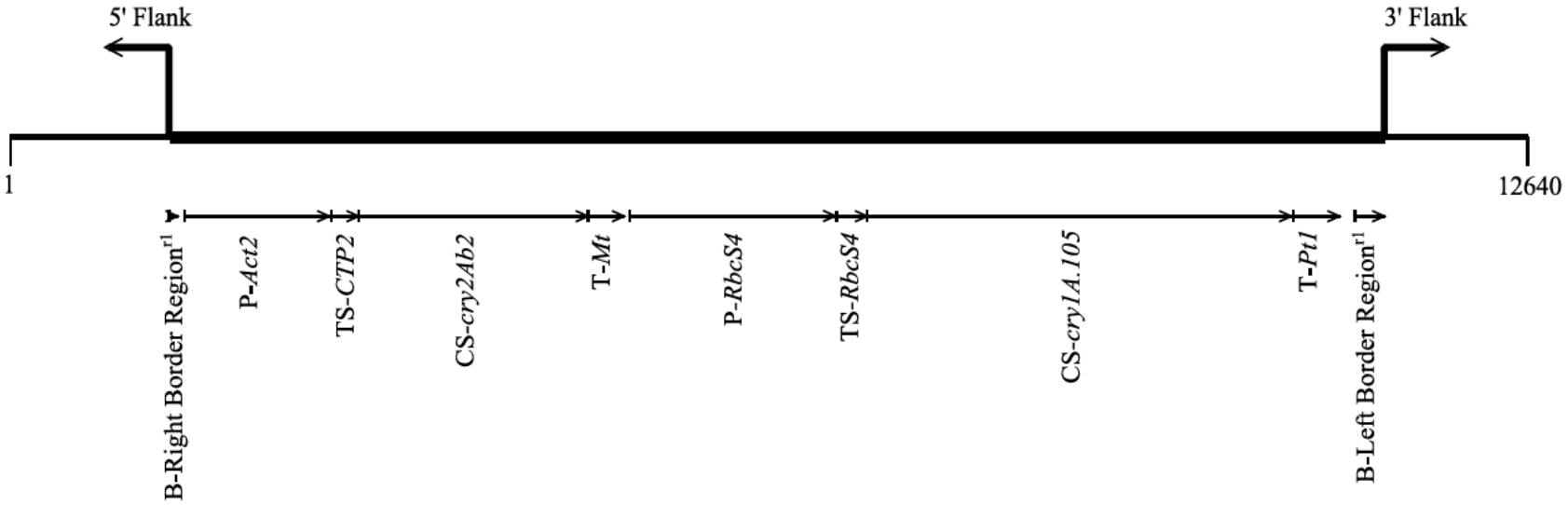


Figure 8. Schematic Representation of the Insert and Flanking Sequences in MON 87751

DNA derived from T-DNA I of PV-GMIR13196 integrated in MON 87751. Right-angled arrows indicate the ends of the integrated T-DNA I and the beginning of the flanking sequence. Identified on the map are genetic elements within the insert. This schematic diagram is drawn to scale.

[r1]Superscript in Left and Right Border Regions indicate that the sequence in MON 87751 was truncated compared to the sequences in PV-GMIR13196.

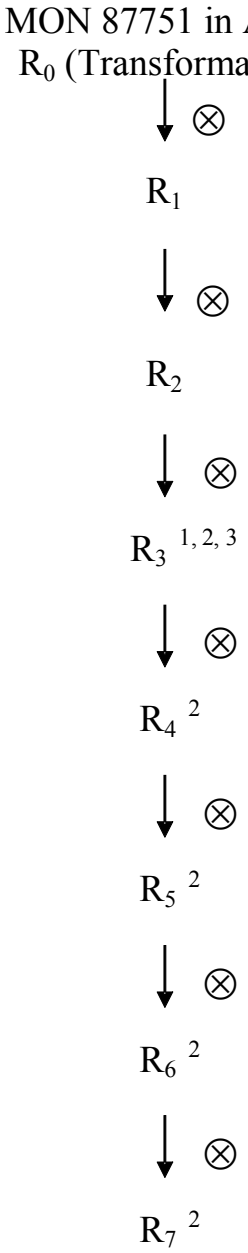


Figure 9. Breeding History of MON 87751

R_0 corresponds to the transformed plant, ⊗ designates self-pollination.
[1]Generation used for full molecular characterization
[2] Generations used to confirm insert stability
[3] Generation used for commercial development of MON 87751

Characterization of insert number in MON 87751 using Bioinformatic Analysis

The number of insertion sites of DNA from PV-GMIR13196 in MON 87751 was assessed by performing NGS/JSA on MON 87751 genomic DNA using the R3 generation (Figure 9).

Selection of Sequence Reads Containing Sequence of the Transformation Plasmid.

Since PV–GMIR13196 was utilized to transform MON 87751, any DNA inserted into MON 87751 will consist of sequences that have similarities to the PV–GMIR13196. To fully characterize the DNA inserted in MON 87751, it is sufficient to analyze only the sequence reads that have similarities to the transformation plasmid. In order to analyze the sequence data for insert number, all sequences that have significant sequence similarity to PV–GMIR13196 were selected from the full sequencing datasets. Due to the high overall sequence coverage, on average, any area of the genome will be covered by at least 75 of the 100–mer sequences; this ensures that sequences from PV–GMIR13196 inserted into the genome will be detected by the analysis.

Using previously established criteria, reads similar to the transformation plasmid were selected from MON 87751 and the conventional control sequence datasets and were mapped and then used as input data for bioinformatic junction sequence analysis.

Mapping and Junction Sequence Analysis (JSA) to Determine Presence or Absence of Backbone and T–DNA II sequence, the Number of Insertions and T–DNA I Copy Number

By mapping sequence reads relative to the transformation plasmid sequence and evaluating the number and sequence of unique junction classes, the presence or absence of backbone and T–DNA II sequence and the number of T–DNA I insertions and copies of the T–DNA I can be determined. For a single copy T–DNA I insert sequence at a single genomic locus, few if any reads aligning with plasmid backbone and T–DNA II, and a pair of junction sequences, each originating from the end of the T-DNA I insert are expected.

Thirty reads from the conventional control dataset and a total of 143 reads from the MON 87751 R3 dataset aligned with the plasmid backbone or T-DNA II sequence. Of the 143 reads, there were 126 sequence reads that aligned with the CTP2 sequence that is contained in T–DNA I and in T-DNA II. Since the CTP2 sequence contained in the T–DNA I is identical to that found in the T–DNA II and as described later, no junction sequence containing CTP2 was identified, the 126 reads were determined to originate from T-DNA I. Therefore, 17 reads from the MON 87751 R3 dataset aligned with the plasmid backbone or T–DNA II sequence. In contrast to the aforementioned 30 and 17 reads aligning that are scattered across the length of the backbone or T–DNA II, greater than 10,000 reads from MON 87751 R3 dataset aligned with the plasmid T–DNA I sequence. In addition, no junction sequence containing backbone or T-DNA II was identified as described below. When combined, results indicate that MON 87751 R3 does not contain inserted sequence from the transformation plasmid backbone or T–DNA II.

To determine the insert number in MON 87751 R3 generation, 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 the table below.

Table 5. Unique Junction Sequence Class Results

Sample	*Junction Sequence Classes Detected*
MON 87751	2
A3555	0

Detailed information of the junction sequences resulting from JSA is shown in Figure 10. The location and orientation of the junction sequences relative to the DNA insert in MON 87751 are shown in Figure 10. As shown in the figure, there are two junction classes identified in MON 87751. Junction sequence class A and class B (JSC–A and JSC–B) both contain the T–DNA border sequence joined to flanking sequence, indicating that they represent the sequences at the junctions of the intended T–DNA insert and the soybean genome. No JSCs were detected in the control sample.

Considered together, the absence of plasmid backbone or T–DNA II and the presence of two junction sequence classes (joining T–DNA border and flanking sequences) indicate this single pair of JSCs arises from the intended PV–GMIR13196 T–DNA I at a single locus in the genome of MON 87751. JSC–A represents the junction of the T–DNA left border sequence to the flank; JSC–B represents the junction of the T–DNA right border sequence to the flank. As shown by perfect and complete alignment of the JSCs to the full flank/insert sequence, both of these JSCs originate from the same locus of the MON 87751 genome and are linked by contiguous, known and expected DNA sequence. This result further demonstrates a single locus of integration. Based on this comprehensive NGS/JSA study it is concluded that MON 87751 contains one copy of the T–DNA inserted into a single locus, as shown in Figure 8. This conclusion is also confirmed by the sequencing and analysis of overlapping PCR products from this locus as described in Section A3(d)(iii).

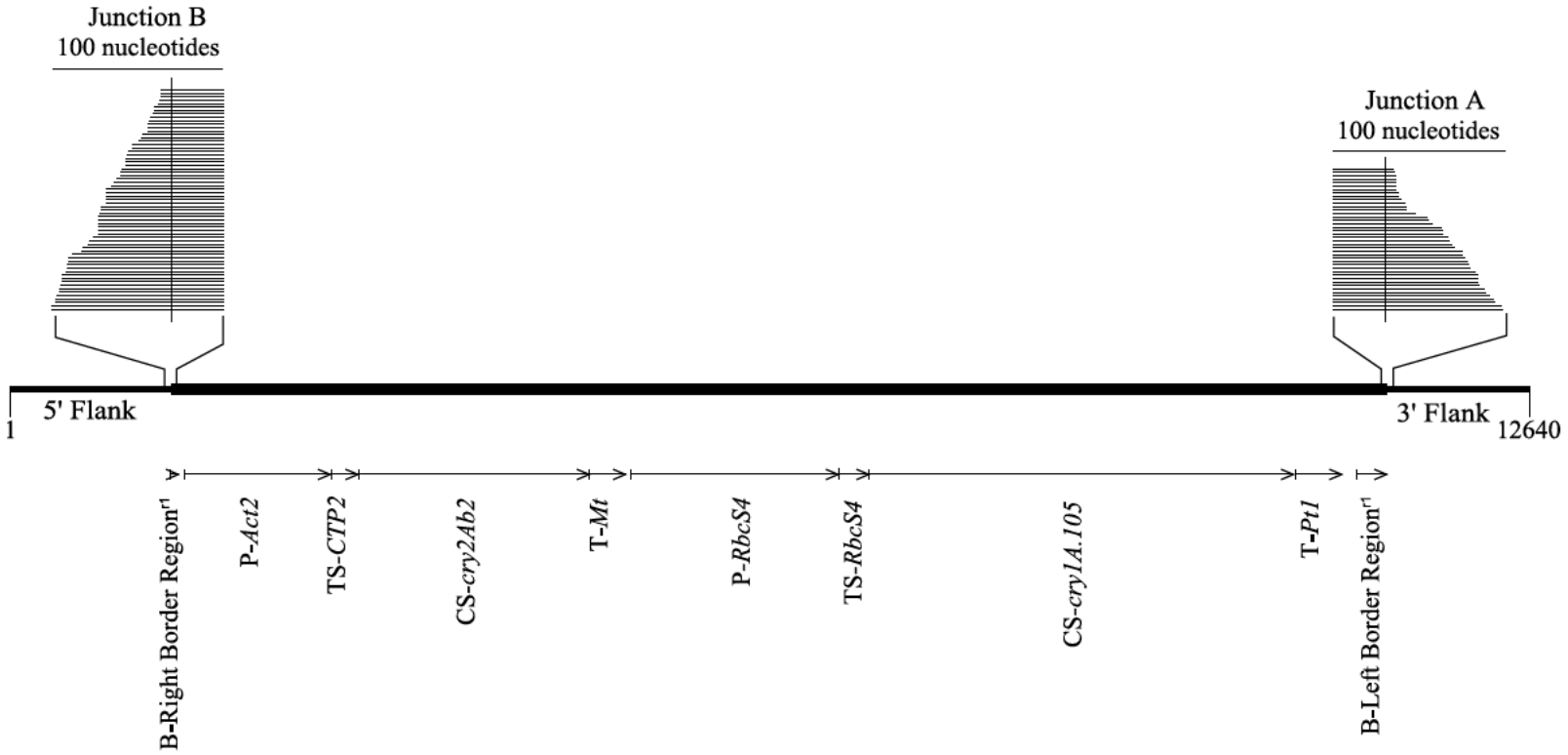


Figure 10. Junction Sequences Detected by NGS/JSA

Linear map of MON 87751 illustrating the relationship of the detected junction sequences to the insert locus. The individual junction sequences detected by JSA are illustrated as stacked bars.

[r1]Superscript in Left and Right Border Regions indicate that the sequence in MON 87751 was truncated compared to the sequences in PV-GMIR13196.

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

The intactness and organization of the elements within the MON 87751 insert was confirmed by using PCR to amplify and subsequently sequence four overlapping DNA amplicons that span the entire insert and the associated flanking DNA sequence. The positions of the PCR products relative to the insert, as well as the results of the PCR analyses, are shown in Figure 11.

As expected, PCR products were not obtained from the negative control reactions containing the conventional control A3555 DNA template (Figure 11, Lane 2, Lane 5, Lane 9, and Lane 13) or no template DNA (Figure 11, Lane 4, Lane 8, Lane 12, and Lane 15) with any of the primer sets. PV–GMIR13196 was used as a positive control template and produced the expected band at ~4.6 kb (Lane 6) and ~3.3 kb (Lane 10). In addition, there are unexpected faint bands at ~1.0 kb and ~0.5 kb (Figure 11, Lane 6 and Lane 10, respectively), which are present only in the positive control lanes and therefore have no effect on the study. The reactions containing MON 87751 DNA produced the expected bands at ~3.3 kb for Product A, ~4.6 kb for Product B, ~3.3 kb for Product C, and ~2.6 kb for Product D (Figure 11, Lane 3, Lane 7, Lane 11, and Lane 14, respectively). The generation of these overlapping PCR products confirms that the organization of the insert is as expected.

To determine the sequence of the MON 87751 insert and the adjacent DNA flanking the insert, Product A, Product B, Product C, and Product D (Figure 11, Lane 3, Lane 7, Lane 11, and Lane 14, respectively) were sequenced. The consensus sequence representing the insert in MON 87751 is described in Table 4. This consensus sequence was generated by compiling data from multiple sequencing reactions performed on the overlapping PCR products spanning the length of the insert and DNA sequence flanking the insert.

An alignment was performed between the MON 87751 sequence and the PV–GMIR13196 sequence to determine the intactness and organization of the integrated DNA and the 5′ and 3′ insert–to–flank DNA junctions of the insert. This analysis determined that the DNA sequence of the MON 87751 insert is 10,119 base pairs (bases 1335–11453). The MON 87751 sequence aligns to the PV–GMIR13196 T–DNA I sequence beginning at base 216 in the Right Border region and ending at base 10334 in the Left Border region of PV–GMIR13196 T–DNA I (Figure 4). This result confirms that the organization and the sequence of the genetic elements in the insert are consistent with that in PV–GMIR13196 T–DNA I, as described in Table 4 and is fully consistent with the results of the NGS dataset analyses.

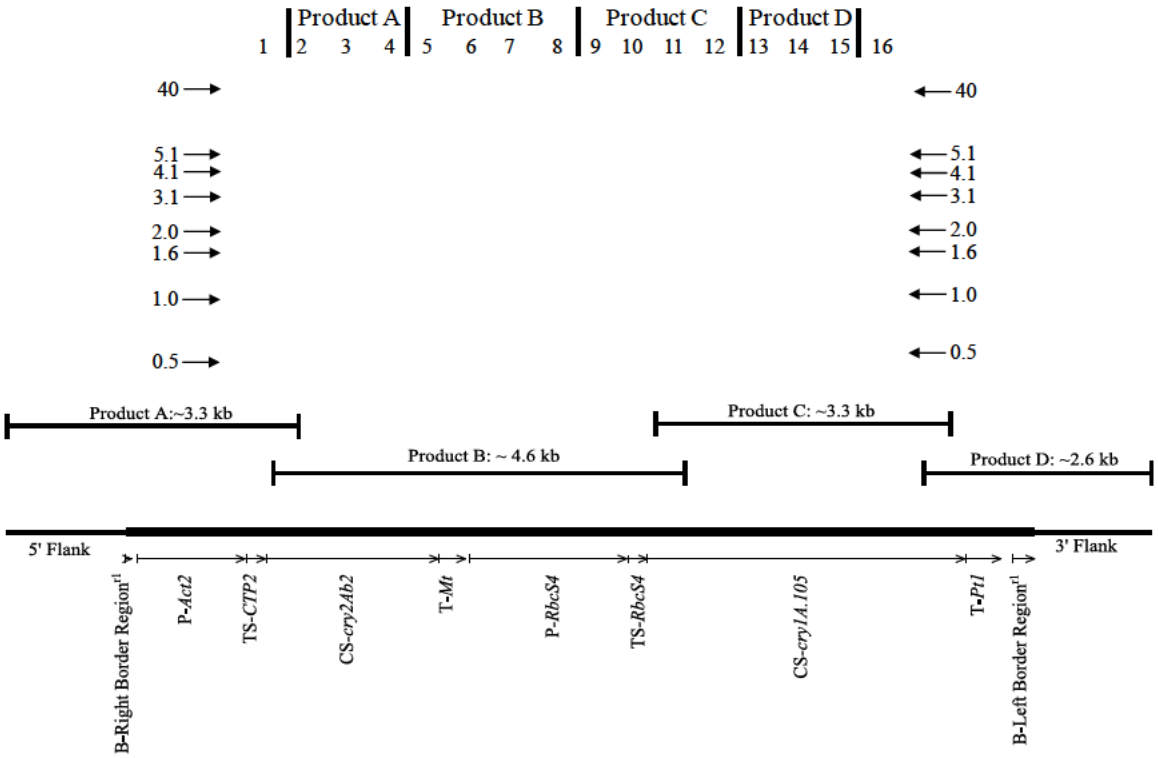


Figure 11. Analysis of Overlapping PCR Products Across the Insert in MON 87751

PCR was performed on both conventional control genomic DNA and MON 87751 genomic DNA using four pairs of primers to generate overlapping PCR fragments from MON 87751 for sequencing analysis. To verify the PCR products, 5 µl of each of the PCR reactions was loaded on the gel except lane 3 which had 10 µl loaded on the gel. The expected product size for each amplicon is provided in the illustration of the insert in MON 87751 that appears at the bottom of the figure. This figure is a representative of the data generated in the study. Lane designations are as follows:

Lane		Lane	
1	1 Kb DNA Extension Ladder	9	Conventional Control A3555
2	Conventional Control A3555	10	PV-GMIR13196
3	MON 87751 (10 µl)	11	MON 87751
4	No template control	12	No template control
5	Conventional Control A3555	13	Conventional Control A3555
6	PV-GMIR13196	14	MON 87751
7	MON 87751	15	No template control
8	No template control	16	1 Kb DNA Extension Ladder

Arrows next to the gel denote the size of the DNA, in kilobase pairs, obtained from the 1kb DNA Extension Ladder (Invitrogen) on the ethidium bromide stained gel.

[r1]Superscript in Left and Right Border Regions indicate that the sequence in MON 87751 was truncated compared to the sequences in PV-GMIR13196.

PCR and DNA sequence analyses were performed on genomic DNA extracted from MON 87751 and the conventional control A3555 to examine the DNA insertion site in MON 87751. PCR was performed using a forward primer specific to the DNA sequence flanking the 5′ end of the insert (Figure 12, Primer A) paired with a reverse primer specific to the DNA sequence flanking the 3′ end of the insert (Figure 12, Primer B). The results of the PCR analysis are shown in Figure 12.

The reaction containing the conventional control A3555 genomic DNA template generated a PCR product of ~2.6 kb (Figure 12, Lane 2). The control PCR reaction containing no template DNA (Figure 12, Lane 3) did not generate a PCR product, as expected.

The ~2.6 kb PCR product generated from the conventional control genomic DNA was sequenced. This consensus sequence was generated by compiling data from multiple sequencing reactions performed on the PCR product.

Alignments of the conventional control sequence to the 5′ and 3′ sequences flanking the MON 87751 insert were performed separately to determine the integrity and organization of the insertion site in MON 87751. The alignment between the sequence flanking the 5′ end of the MON 87751 insert and the conventional control sequence showed that base 15 to base 984 and base 985–1333 of the 5′ flanking sequence of the MON 87751 insert are identical to base 1 to base 970 and base 987–1335 of the conventional control sequence. The alignment between the 3′ end of the MON 87751 insert and the conventional control sequence showed that base 1343 to base 2529 of the conventional control sequence are identical to base 11454 to base 12640 of the sequence flanking the 3′ end of the MON 87751 insert .

The results of the alignment demonstrated that base 1336 to base 1342 of conventional control sequence are not found in the 5′ or 3′ sequences flanking the insert, indicating that 7 bases were likely deleted upon T–DNA integration and base 971 to base 986 of conventional control sequence are not found in the 5′ or 3′ sequences flanking the insert, indicating that 16 bases were likely deleted upon T–DNA integration. The results of the alignment also determined there was a one base pair insertion at position 1334 that occurred during integration of the insert in MON 87751.

In summary, the genomic organization at the insertion site was assessed by comparing the insert and flanking sequence to the sequence of the insertion site in conventional soybean, and this analysis identified both a 1 base pair insertion and a 7 base pair deletion at the insertion site, and also a 16 base pair deletion in the 5′ flanking sequence, that occurred during integration of the T–DNA I sequences. Such changes are common during plant transformation and these changes presumably resulted from double–stranded break repair mechanisms in the plant during *Agrobacterium*–mediated transformation process (Salomon and Puchta, 1998).

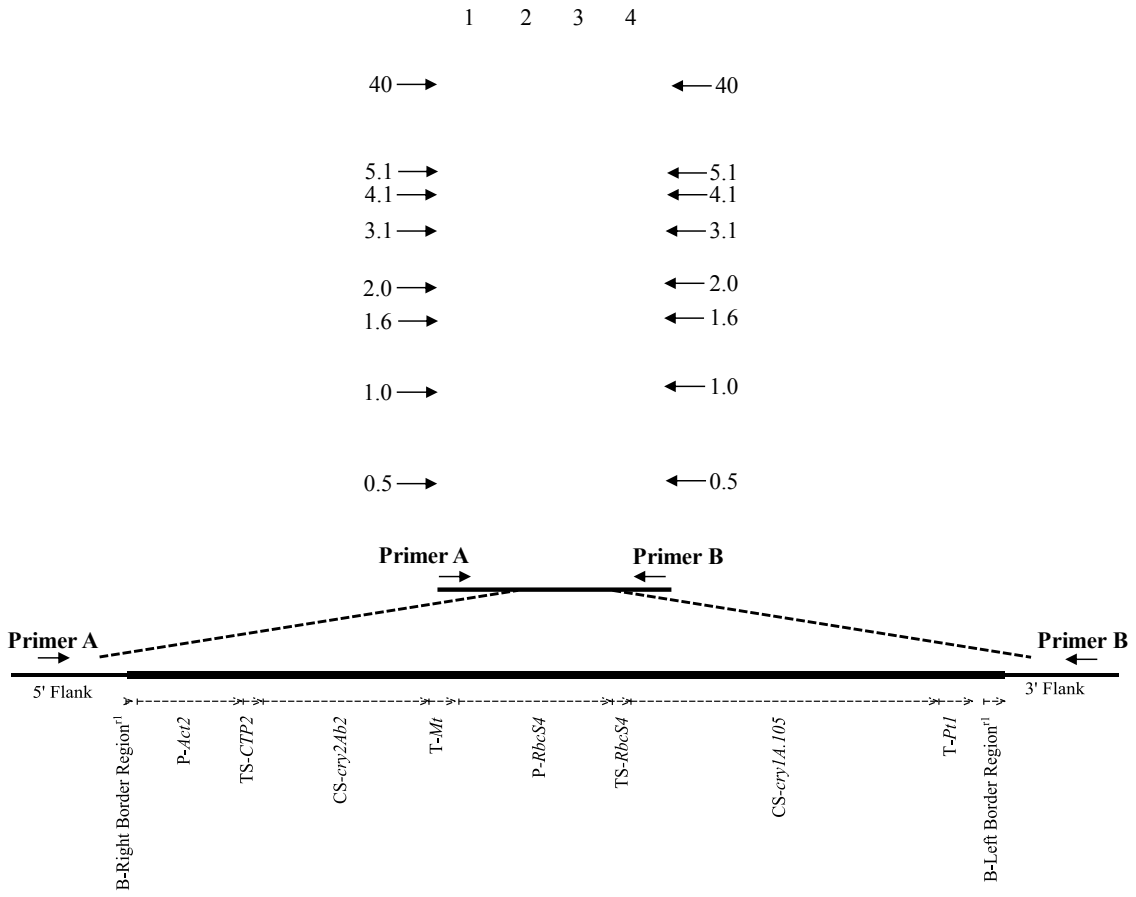


Figure 12. PCR Amplification of the MON 87751 Insertion Site

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 87751. The DNA generated from the conventional control PCR was used for sequencing analysis. This illustration depicts the MON 87751 insertion site in the conventional control (upper panel) and the MON 87751 insert (lower panel). Approximately 5 µl of each of the PCR reactions was loaded on the gel. This figure is representative of the data generated in the study. Lane designations are as follows:

Lane	Sample
1	1 Kb DNA Extension Ladder
2	Conventional Control A3555
3	No template DNA control
4	1 Kb DNA Extension Ladder

Arrows next to the gel denote the size of the DNA, in kilobase pairs, obtained from the 1 Kb DNA Extension 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 87751 was truncated compared to the sequences in PV-GMIR13196.

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

PCR and DNA sequence analyses performed on MON 87751 and the conventional control determined the organisation of the genetic elements within the insert as given in Figure 8.

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

Bioinformatic Assessment of Putative Open Reading Frames (ORFs) of MON 87751 Insert and Flanking Sequences

The 2009 Codex Alimentarius Commission guidelines for the safety assessment of food derived from biotechnology crops (Codex Alimentarius, 2009) advise an 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 ORFs at the plant-insert junction or alternative reading frames in the insert are capable of being transcribed or translated into a protein. Results from these bioinformatics analyses demonstrate that any putative polypeptides in MON 87751 are unlikely to exhibit allergenic, toxic or otherwise biologically adverse properties.

Bioinformatic analyses were performed on the MON 87751 insert and flanking genomic DNA sequences to assess the potential for allergenicity, toxicity, or biological activity of putative polypeptides encoded by all six reading frames present in the MON 87751 insert DNA as well as ORFs present in the 5' flanking-inserted T-DNA and inserted T-DNA 3' flanking sequence junctions. These various bioinformatic evaluations are depicted in Figure 13. The entire MON 87751 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 sequence was also subjected to bioinformatic analyses. Putative peptides/polypeptides from each reading frame of eight amino acids or greater were then compared to toxin, allergen and all proteins databases using bioinformatic tools. Similarly, ORFs spanning the 5' flanking sequence DNA-inserted DNA junctions, and 3' flanking sequence DNA-inserted DNA junctions were translated from stop codon to stop codon in all six reading frames (three forward reading frames and three reading frames in reverse orientation). There are no analytical data that indicate any putative polypeptides/proteins subjected to bioinformatic evaluation other than Cry1A.105 and Cry2Ab2 are produced. Moreover, the data generated from these analyses confirm that even in the highly unlikely occurrence that a translation product other than Cry1A.105 and Cry2Ab2 was derived from frames 1 to 6 of the insert DNA, or the ORFs spanning the insert junctions, it would not share a sufficient degree of sequence similarity with other proteins to indicate it 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 toxicity, allergenicity or biological activity of any putative peptides encoded by translation of reading frames 1 through 6 of the inserted DNA in MON 87751 (Figure 13).

The FASTA sequence alignment tool was used to assess structural relatedness between the query sequences and any protein sequences in the AD_2013[3], TOX_2013[4], and PRT_2013[5] databases. Structural similarities shared between each putative polypeptide and 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 an *E*-score less than 1e-5 (1×10^{-5}) are deemed significant because they may reflect shared structure and function among sequences (Ladics *et al.*, 2007). 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 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_2013 database.

The results of the search comparisons showed that no relevant structural similarity to known allergens or toxins were observed for any of the putative polypeptides when compared to proteins in the allergen (AD_2013) or toxin (TOX_2013) databases. Additionally, no alignment met or exceeded the Codex Alimentarius (2009) FASTA alignment threshold for potential allergenicity of 35% identity over 80 amino acids. Using the eight amino acid sliding window algorithm to search the AD_2013 database, the frame 4 translation identified one alignment with tropomyosin from the bivalve mollusk, *Solen strictus* (GI-219806602). Inspection of the alignment revealed that the aligned region in the query sequence is on the reverse compliment strand of the promoter from the *act2* gene. It is highly unlikely that an mRNA from the reverse complement strand of the *act2* promoter would be produced *in planta* as this would require downstream CTP or Cry2Ab2 coding sequence to act as a promoter. This putative mRNA in turn would result in the formation of a dsRNA as it

[3] The allergen, gliadin, and glutenin sequence database (AD_2013) was obtained from Food Allergy Research and Resource Program Database (FARRP_2013) (http://www.allergenonline.org) and was used for the evaluation of sequence similarities. It is referred to as the AD_2013 database and contains 1,630 sequences.

[4] The toxin database is a subset of sequences derived from the PRT_2013 database that was selected using a keyword search then filtered to remove likely non-toxin proteins. It is referred to as the TOX_2013 database and contains 8,881 sequences and was used for the evaluation of sequence similarities.

[5] A comprehensive collection of protein sequences that comprises GenBank release 193 was downloaded from NCBI and was used for the evaluation of sequence similarities. It is referred to as the PRT_2013 database and contains 27,998,271 sequences.

would be complementary to the mRNA that encodes Cry2Ab2. Double stranded RNA molecules are unstable and are therefore not translated.

When used to query the PRT_2013 database, translations of frames 1-3 and 5 yielded alignments with *E*-scores less than or equal to 1e-5. The frame 1 translation positively identified Cry2Ab protein (GI-27311145) with a significant *E*-score of 0 and 100% identity, and is consistent with the known insert structure in MON 87751. The frame 2 translation positively identified Cry1A.105 protein (GI-213504429) with a significant *E*-score of 0 and 100% identity, and is consistent with the known insert structure in MON 87751. The frame 3 translation yielded 22 alignments with significant *E*-scores. Visual inspection of these alignments revealed that although yielding *E*-scores that were below or equal to the significance threshold, the alignments were punctuated with numerous stop codons in the query sequences and required numerous gaps to optimize the alignments. As a result, it is unlikely they reflect conserved structure. The only significant alignment observed for the frame 5 query sequence was to an uncharacterized protein from *Arabidopsis thaliana* (GI-332642625). The alignment displays 100% identity over 47 amino acids with an *E*-score of 3e-12. Inspection of the alignment revealed that the aligned region in the query sequence is on the reverse compliment strand of the promoter from the *act2* gene of *Arabidopsis thaliana*. It is highly unlikely that an mRNA from the reverse complement strand of the *act2* promoter would be produced *in planta.* Although less than the *E*-score limit of 1e-5, this conceptual translation's alignment with frame 5 did not provide any indication of adverse biological activity.

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 87751. As a result, in the unlikely event that translation products other than Cry1A.105 and Cry2Ab2 were derived from reading frames 1 to 6, these putative polypeptides are not expected to be cross-reactive allergens, toxins, or display adverse biological activity.

[redacted]

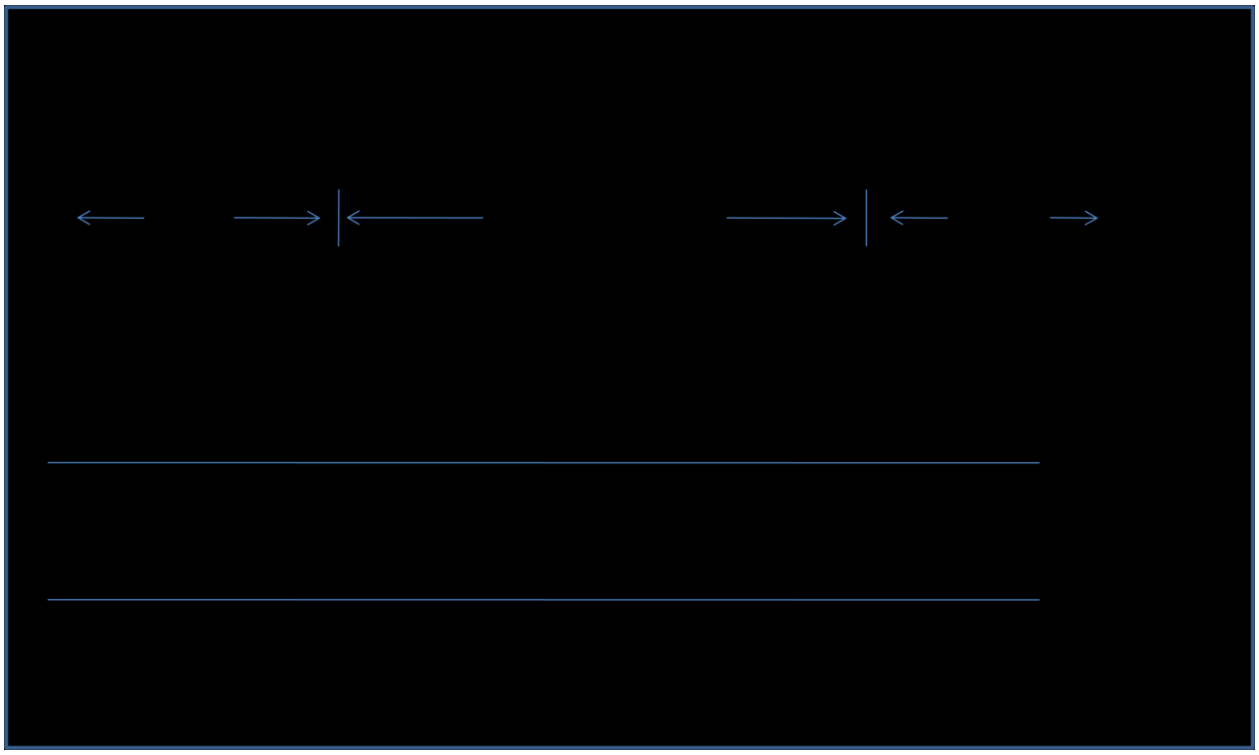
Insert Junction Open Reading Frame Bioinformatics Analysis

Analyses of putative polypeptides encoded by DNA spanning the 5' and 3' genomic junctions of the MON 87751 inserted DNA were performed using a bioinformatic comparison strategy (Figure 13). The purpose of the assessment is to evaluate the potential for putative polypeptides predicted from novel ORFs to have homology to known allergens, toxins, or proteins that display adverse biological activity. Sequences spanning the 5′ flanking sequence DNA-inserted DNA and the inserted DNA-3′ flanking sequence DNA (Figure 13) were translated from stop codon (TGA, TAG, TAA) to stop codon in all six reading frames. The resulting putative polypeptides from each reading frame that were eight amino acids or greater in length were compared to AD_2013, TOX_2013, and PRT_2013 databases using FASTA and to the AD_2013 database using an eight amino acid sliding window search. A total of 12 putative polypeptide sequences were subjected to bioinformatics analysis.

The FASTA sequence alignment tool was used to assess structural relatedness between the query sequences and protein sequences in the AD_2013, TOX_2013, and PRT_2013 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) thresholds for FASTA searches of the AD_2013 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 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_2013 database.

The bioinformatic analysis performed using the 12 putative sequences 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 12 putative flank junction derived sequences and allergens, toxins or biologically active proteins. As a result, in the unlikely occurrence that any of the 12 peptides 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.

[redacted]



AD= AD_2013; TOX= TOX_2013 and PRT= PRT_2013 (GenBank release 193): 8-mer = the eight amino acid sliding window search. POI Coding sequences correspond to the *cryA.105* and *cry2Ab2* coding sequence.

Figure 13. Schematic Summary of MON 87751 Bioinformatic Analyses

A3(e) Family tree or breeding process

The conventional soybean variety A3555, used as the recipient for the *cry1A.105* and *cry2Ab2* expression cassette insertion that produced MON 87751, was developed by Asgrow Seed Company. A3555 is a mid−maturity group III soybean variety.

Soybean variety A3555 is the parental line of MON 87751 and was used as the conventional soybean comparator (hereafter referred to as the conventional control) in the safety assessment of MON 87751. MON 87751 and A3555 have similar genetic backgrounds with the exception of the *cry1A.105*and *cry2Ab2* expression cassettes, thus, the effect of the *cry1A.*105 and *cry2Ab2* expression cassettes could be assessed in an unbiased manner in the comparative safety assessment. In addition, reference varieties, where appropriate, were used to establish ranges of natural variability or responses representative of commercial soybean varieties. The commercial reference varieties used at each location were selected based on their availability and agronomic fit for the respective geographic region.

The breeding history of MON 87751 is presented in Figure 9, and the specific generations tested are indicated in the figure legend. The MON 87751 (R_3) generation was used for the molecular characterization analyses discussed in this section and shown in Figure 9. To assess stability, four additional generations were evaluated by NGS/JSA analysis as previously described in Section A3(f), and compared to the fully characterized MON 87751 (R_3) generation. The conventional control used for the generational stability analysis was A3555, which has a genetic background similar to the other generations in Figure 9 and represents the original transformation line.

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

In order to demonstrate the genetic stability of the T-DNA present in MON 87751 through multiple breeding generations, mapping relative to the transformation plasmid and JSA analyses were performed using NGS reads from DNA obtained from five breeding generations of MON 87751. The breeding history of MON 87751 is presented in Figure 9 and the specific generations tested are indicated in the figure legend. The MON 87751 (R3) generation was used for the molecular characterization analyses described in Section 4.1 and shown in Figure 10. To assess stability, NGS reads from four additional generations were evaluated by mapping and JSA as previously described in this section, and compared to the fully characterized MON 87751 (R3) generation. The conventional control used for the generational stability analysis was A3555, which included similar background to all generations tested. Genomic DNA isolated from each of the selected generations of MON 87751 and conventional control were used for mapping and JSA.

To determine the insert number in the MON 87751 samples, the sequences generated and selected as described above in Section 4.1 were analyzed using JSA. The number of any resultant unique JSC junctions containing the PV–GMIR13196 DNA sequence determined by this analysis and is shown in the table below.

Table 6. Junction Sequence Classes Detected

Sample	*Junction Sequence Classes Detected*
MON 87751 (R_3)	2
MON 87751 (R_4)	2
MON 87751 (R_5)	2
MON 87751 (R_6)	2
MON 87751 (R_7)	2
A3555	0

The table above shows the presence of two, and only two, identical junction sequence classes in each breeding generation (R3, R4, R5, R6 and R7). This single identical pair of JSCs is observed due to the insertion of a single copy of PV-GMIR13196 T–DNA at a single locus in the genome of MON 87751, and the consistency of these JSC data combined with the mapping data across all generations tested demonstrates that this single locus is stably maintained throughout the MON 87751 breeding process.

This result, therefore, demonstrates that the MON 87751 (R3) single locus of integration has been maintained through five generations of the MON 87751 breeding history; thereby confirming the stability of the insert. Based on this comprehensive sequence data and bioinformatic analysis (NGS/JSA) it is reasonable to conclude MON 87751 contains a single, stable, inserted T–DNA I.

Inheritance of the Genetic Insert in MON 87751

During development of MON 87751, segregation data were generated to assess the heritability and stability of T-DNA I in MON 87751. Chi square (χ^2) analysis was performed over several generations to confirm the segregation and stability of T-DNA I in MON 87751. The Chi square analysis is based on testing the observed segregation ratio to the segregation ratio expected based on Mendelian principles.

The MON 87751 breeding path for generating segregation data is described in Figure 14. The transformed R_0 plant was self-pollinated to produce R_1 seed. From the R_1 segregating population, which consisted of 42 total plants containing T-DNA I but not T-DNA II, an individual plant homozygous for the *cry1A.105* and *cry2Ab2* genes was identified via Real-Time TaqMan® PCR and Invader® analyses. Real-Time TaqMan PCR measures the accumulation of the fluorescent signals incorporated into the target sequence throughout the entire PCR assay. The Invader analysis is a signal amplification technology for quantitative analysis or detection of genetic variations. The Invader method uses Cleavase® enzymes, a structure-specific family of endonucleases, and fluorescence detection to detect specific target sequences.

The selected R_1 MON 87751 homozygous plant was self-pollinated to give rise to a population of R_2 plants, which were in turn self-pollinated to obtain the R_3 generation. At each generation, the fixed homozygous plants were tested for the expected segregation

pattern of 1:0 (positive: negative) for the MON 87751 T-DNA I using the Real-Time TaqMan® PCR and Invader® analyses.

Homozygous R_3 MON 87751 plants were crossed to a Monsanto proprietary soybean line (MonSoy8329) that did not contain the *cry1A.105* and *cry2Ab2* expression cassettes to produce F_1 hemizygous seed. A hemizygous F_1 plant was selected and then self-pollinated to produce F_2 seed. The resulting F_2 plants were tested for the presence of the T-DNA I by Real-Time TaqMan® PCR assay. This process of self-pollination of hemizygous plants and zygosity determination of the MON 87751 T-DNA I by Real-Time TaqMan PCR analysis was repeated for the F_2, F_3, and F_4 plants. Subsequently, assessment at each of these generations was based on zygosity, and MON 87751 T-DNA I was predicted to segregate at a 1:2:1 (homozygous positive: hemizygous positive: homozygous negative) ratio for progeny derived from a hemizygous parental plant according to Mendelian inheritance principles.

A Chi square (χ^2) analysis was used to compare the observed segregation ratios to the expected ratios according to Mendelian inheritance principles. The χ^2 was calculated as:

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

where o = observed frequency of the phenotype and e = expected frequency of the 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 87751 are presented in Table 7 The χ^2 value in the F_2, F_3, and F_4 generations indicated no statistically significant difference between the observed and expected 1:2:1 segregation ratio. These results support the conclusion that the MON 87751 T-DNA resides at a single locus within the soybean genome and is inherited according to Mendelian inheritance principles. These results are also consistent with the molecular characterization data that indicate MON 87751 contains single, intact copies of the *cry1A.105* and *cry2Ab2* expression cassettes that were inserted into the soybean genome at a single locus.

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

® Invader and Cleavase are registered trademarks of Third Wave Technologies, Inc.

Figure 14. Breeding Path for Generating Segregation Data for MON 87751

⊗Self pollinated.
* Chi-square analysis was conducted on segregation data from the F_2, F_3, and F_4 generations (bolded text).
†The soybean line used in the cross that did not contain the *cry1A.105* and *cry2Ab2* genes is MonSoy8329**.**

Table 7. Segregation of the Expression Cassette During the Development of MON 87751

					1:2:1 Segregation				
Generation	**Total Plants[2]**	**Observed Homozygous Positive**	**Observed Hemizygous**	**Observed Homozygous Negative**	**Expected Homozygous Positive**	**Expected Hemizygous**	**Expected Homozygous Negative**	**χ^2**	**Probability[3]**
F_2[1]	152	39	72	41	38	76	38	0.47	0.79
F_3[1]	214	49	114	51	53.5	107	53.5	0.95	0.62
F_4[1]	204	58	105	41	51	102	51	3.01	0.22

[1]Segregation was evaluated using Real-Time TaqMan analysis.
[2]"Total plants" refers to the total number of plants in which zygosity could be determined using the assay.
[3]Chi-square analysis was performed to analyze the segregation ratios ($p \leq 0.05$).

Characterization of the Genetic Modification Summary and Conclusion

Molecular characterization of MON 87751 by mapping and NGS/JSA analyses demonstrated that a single copy of T–DNA I from PV–GMIR13196 was integrated in MON 87751. This single locus included only sequence from the T–DNA I and was devoid of sequence from the backbone and T-DNA II.

The PCR and DNA sequence analyses performed on MON 87751 determined the complete sequence of the insert and adjacent flanking DNA, and the 5' and 3' insert–to–flank junctions. It confirmed that each genetic element (except for the border regions) in the insert is intact and the sequence of the insert is identical to the corresponding sequence in PV–GMIR13196. Furthermore, the genomic organization at the insertion site was assessed by comparing the insert and flanking sequence to the sequence of the insertion site in conventional soybean, and this analysis identified both a 1 base pair insertion and a 7 base pair deletion at the insertion site, and also a 16 base pair deletion in the 5′ flanking sequence, that occurred during integration of the T–DNA I sequences.

Generational stability analysis by NGS/JSA demonstrated that MON 87751 has been maintained through five breeding generations, thereby confirming the stability of the T-DNA I in MON 87751.

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

Generational Stability of Cry1A.105 and Cry2Ab2 Protein Expression in MON 87751

In order to assess the presence of the Cry1A.105 and Cry2Ab2 proteins in MON 87751 across multiple breeding generations, western blot analysis of MON 87751 was conducted on leaf tissue collected from generations R3, R4, R5, R6, and R7 of MON 87751, and on leaf tissue of the conventional control (A3555).

Generational Stability of Cry1A.105 Protein Expression in MON 87751

The presence of the Cry1A.105 protein was demonstrated in five breeding generations of MON 87751 using western blot analysis (Figure 15). An *E. coli*-produced Cry1A.105 standard was used as a reference for the identification of the Cry1.105 protein. The presence of Cry1A.105 protein in leaf tissues of MON 87751 was determined by visual comparison of the bands produced in multiple breeding generations (Figure 15, Lanes 6–10) to the Cry1A.105 reference standard (Figure 15, Lane 4). As expected, the Cry1A.105 protein was present in all five breeding generations of MON 87751 and migrated with a mobility indistinguishable from that of the *E. coli*-produced protein standard analyzed on the same western blot. Also, as expected, the Cry1A.105 protein was not detected in the conventional control leaf extract (Figure 15, Lane 5). An additional lower molecular weight protein band between 50 and 60 kDa appeared in Cry1A.105 protein reference standard (Figure 15, Lane 4) and MON 87751 leaf samples (Figure 15, Lanes 6-10). The lower molecular weight immunoreactive band most likely represents a degradation product of the Cry1A.105 protein. It has been shown that C-terminal regions of Cry1 proteins can be degraded by proteolytic enzymes present in the cells or released during protein purification (Gao *et al.*, 2006).

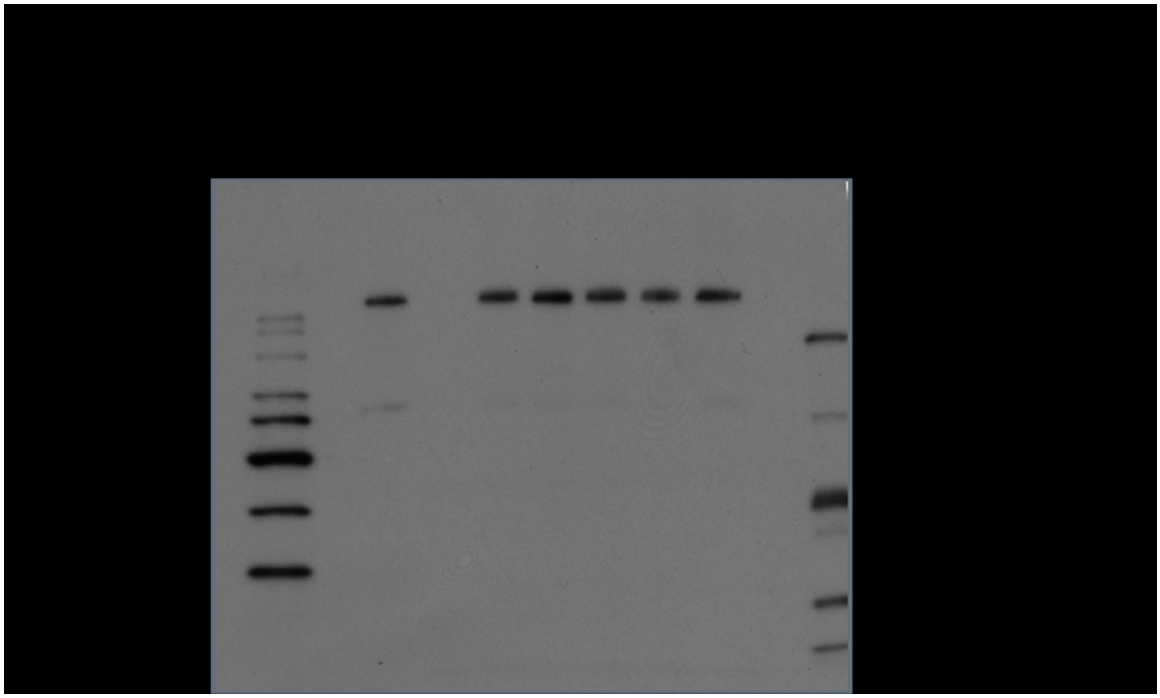


Figure 15. Presence of Cry1.A105 Protein in Multiple Generations of MON 87751

Extracts from five generations of MON 87751 leaf tissues, conventional control leaf tissue, *E. coli*-produced Cry1A.105 standard, and molecular weight markers were subjected to SDS-PAGE and electrotransferred to a polyvinylidene fluoride membrane. The membrane was incubated with goat anti-Cry1A.105 antibodies and immunoreactive bands visualized through the use of chemiluminescent reagents. The image represents a 10 minute exposure. Arrows denote the size of the protein in kiloDaltons obtained from MagicMark XP Standards (Invitrogen) transferred to the membrane. Lane designations are as follows:

Lane	Description	Amount Loaded on Gel
1	Blank	N/A
2	MagicMark XP Molecular Weight Marker	5 µl
3	Blank	10 µl
4	*E. coli*-produced Cry1A.105 protein (1.5 ng)	10 µl
5	Conventional Substance, Control	10 µl
6	Test Substance, R3	10 µl
7	Test Substance, R4	10 µl
8	Test Substance, R5	10 µl
9	Test Substance, R6	10 µl
10	Test Substance, R7	10 µl
11	Blank	N/A
12	Precision Plus Protein Dual Color Molecular Weight Marker	5 µl

Generational Stability of Cry2Ab2 Protein Expression in MON 87751

The presence of the Cry2Ab2 protein was demonstrated in five breeding generations of MON 87751 using western blot analysis (Figure 16). A *Bacillus thuringiensis* (*B.t.*)-produced Cry2Ab2 standard was used as a reference for the identification of the Cry2Ab2 protein. The presence of Cry2Ab2 protein in leaf tissue of MON 87751 was determined by visual comparison of the bands produced in the five breeding generations (Figure 16, Lanes 5–9) to the Cry2Ab2 reference standard (Figure 16, Lane 3). As expected, the Cry2Ab2 protein was present in all five breeding generations of MON 87751 tissue samples and migrated with a mobility indistinguishable from that of the *B.t.*-produced protein standard analyzed on the same western blot. Also, as expected, the Cry2Ab2 protein was not detected in the conventional control extract (Figure 16, Lane 4).

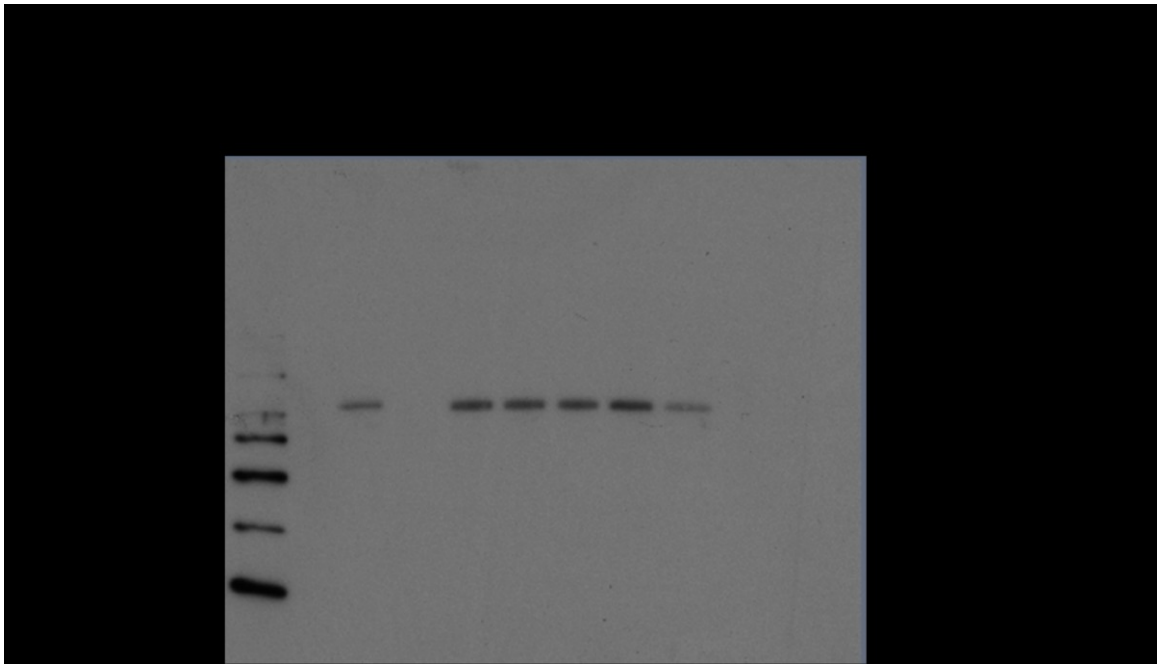


Figure 16. Presence of Cry2Ab2 Protein in Multiple Generations of MON 87751

Aliquots of extracts from five generations of MON 87751 leaf tissues, conventional control leaf tissues, *B.t.*-produced Cry2Ab2 protein standard, and molecular weight markers were subjected to SDS-PAGE and electrotransferred to a polyvinylidene fluoride membrane. The membrane was incubated with mouse anti-Cry2Ab2 antibody and immunoreactive bands visualized through the use of chemiluminescent reagents. The image represents a 10 minutes exposure. Arrows denote the size of the protein in kiloDaltons obtained from MagicMark XP Standards (Invitrogen) transferred to the membrane. Lane designations are as follows:

Lane	Description	Amount Loaded on Gel
1	MagicMark XP Molecular Weight Marker	5 µl
2	Blank	N/A
3	*B.t.*-produced Cry2Ab2 protein (0.2 ng)	10 µl
4	Conventional Substance, Control	10 µl
5	Test Substance, R3	10 µl
6	Test Substance, R4	10 µl
7	Test Substance, R5	10 µl
8	Test Substance, R6	10 µl
9	Test Substance, R7	10 µl
10	Blank	N/A
11	Precision Plus Protein Dual Color Molecular Weight Marker	5 µl
12	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 87751 in a collected sample.

B Information Related to the Safety of the GM Food

B1 Equivalence Studies

B1(a) Cry1A.105 protein identity and equivalence

The safety assessment of crops derived through biotechnology includes characterization of the physicochemical and functional properties of the protein(s) produced from the inserted DNA and confirmation of the safety of the introduced protein(s). For the safety data generated using *E. coli*-produced Cry1A.105 to be applied to Cry1A.105 protein produced in MON 87751, the equivalence of the plant- and *E. coli*-produced proteins must be assessed. To assess the equivalence between MON 87751-produced Cry1A.105 and *E. coli*-produced Cry1A.105 proteins a small quantity of the Cry1A.105 protein was purified from MON 87751 soybean. The MON 87751-produced Cry1A.105 protein purified from MON 87751 soybean was characterized and the equivalence of the physicochemical and functional properties between the MON 87751-produced and *E. coli*-produced Cry1A.105 proteins was established using a panel of analytical tests: 1) N-terminal sequence analysis of MON 87751-produced Cry1A.105 characterized the N-terminus; 2) MALDI-TOF MS analysis yielded peptide masses consistent with the expected peptide masses from the theoretical trypsin digest of the MON 87751-produced Cry1A.105 sequence; 3) western blot analysis with antibodies specific for Cry1A.105 protein demonstrated that the immunoreactive properties of the MON 87751-produced and *E. coli*-produced Cry1A.105 proteins were equivalent; 4) SDS-PAGE analysis showed that the electrophoretic mobility and apparent molecular weight of the MON 87751-produced and *E. coli*-produced Cry1A.105 proteins were equivalent; 5) MON 87751-produced and *E coli*-produced Cry1A.105 proteins were both determined to be non-glycosylated; and 6) functional activity analysis demonstrated that MON 87751-produced and *E. coli*-produced Cry1A.105 proteins had equivalent insecticidal activity. Taken together, these data provide a detailed characterization of the MON 87751-produced Cry1A.105 protein and establish its equivalence to the *E. coli*-produced Cry1A.105 protein.

[redacted]

B1(a)(i) Results of the N-terminal sequencing analysis

Fifteen cycles of N-terminal sequencing was performed on the MON 87751-produced Cry1A.105 protein. The sequence obtained corresponded to the N-terminal sequence for Cry1A.105 containing four amino acids derived from the CTP (Figure 17). Cysteine is shown in the predicted sequence at position one based on the coding sequence of the Cry1A.105 construct in MON 87751, but it was not observed in the experimental analysis. However, cysteine is unstable during the acid hydrolysis reaction used for N-terminal sequencing, and is usually not explicitly observed (Speicher *et al.*, 2009). The clear identification of amino acids in subsequent cycles of the sequencing analysis confirmed that an unidentified amino acid was present at position one. The N-terminal sequencing results for MON 87751-produced Cry1A.105 protein were consistent with the sequencing results for the *E. coli*-produced Cry1A.105 protein, which was engineered to contain a cysteine as the first

amino acid. Hence, the sequence information confirms the identity of the Cry1A.105 protein isolated from the seed of MON 87751.

Amino acid residue # from the N-terminus →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Expected Sequence →	C	M	Q	A	M	D	N	N	P	N	I	N	E	C	I
	|	|	|	|	|	|	|	|	|	|	|	|	|	|	|
Experimental Sequence →	X	M	Q	A	M	D	N	N	(P)	X	X	X	X	X	X

Figure 17. N-Terminal Sequence of the MON 87511 Cry1A.105 Protein

The experimental sequence obtained from the MON 87751-produced Cry1A.105 was compared to the expected sequence. The single letter IUPAC-IUB amino acid code is C, cysteine; M, methonine; Q, glutamine; A, alanine; D, aspatic acid; N, asparagine; P, proline; E, glutamic acid; I, isoleucine;. X indicates that the residue was not identified.
( )'s denote tenuous designations.

B1(a)(ii) Results of MALDI-TOF tryptic mass map analysis

The identity of the MON 87751-produced Cry1A.105 protein was confirmed by MALDI-TOF MS analysis of peptide fragments produced by the trypsin digestion of the MON 87751-produced Cry1A.105 protein. In general, protein identification made by proteolytic peptide mapping is considered to be reliable if the measured coverage of the sequence is 15% or higher with a minimum of five matched peptides (Jensen *et al.*, 1997). There were 60 unique peptides identified that corresponded to the expected masses (Table 8). The identified masses were used to assemble a coverage map of the Cry1A.105 protein (Figure 18). The experimentally determined coverage of the Cry1A.105 protein was 53.5% (632 out of 1181 amino acids). This analysis further confirms the identity of MON 87751-produced Cry1A.105 protein.

Table 8. Summary of the Tryptic Masses Identified for the MON 87751-produced Cry1A.105 Protein Using MALDI-TOF MS

Observed Mass	Expected Mass	Diff. [2]	Fragment	Sequence[3]
360.95	361.20	0.25	890-891	WR
515.34	515.34	0.00	529-532	ARIR
579.32	579.33	0.01	753-756	YQLR
589.29	589.31	0.02	229-232	DWIR*
	589.28	0.00	1027-1031	VCPGR*
621.35	621.37	0.02	1032-1036	GYILR
649.37	649.37	0.00	258-262	TYPIR
727.35	727.35	0.00	233-237	YNQFR
731.35	731.36	0.01	428-433	QGFSHR
764.40	764.39	0.01	92-97	IEEFAR
781.39	781.38	0.01	197-202	YNDLTR
804.46	804.46	0.00	263-269	TVSQLTR
816.39	816.40	0.01	222-228	VWGPDSR
832.34	832.48	0.14	670-676	ELSEKVK*
	832.48	0.14	742-748	IDESKLK*
854.40	854.41	0.01	1118-1124	SYTDGRR
907.46	907.46	0.00	178-185	DVSVFGQR
924.49	924.49	0.00	550-557	IFAGQFNK
940.51	940.51	0.00	365-372	TLSSTLYR
976.49	976.50	0.01	434-441	LSHVSMFR
1007.56	1007.55	0.01	541-549	IYVTVAGER
1038.50	1038.50	0.00	214-221	WYNTGLER
1066.44	1066.43	0.01	1125-1132	ENPCEFNR
1074.55	1074.55	0.00	286-296	GSAQ…GSIR
1089.57	1089.57	0.00	495-505	GPGF…DILR
1144.58	1144.57	0.01	454-462	APMFSWIHR
1203.69	1203.68	0.01	354-364	IVAQ…GVYR
1237.60	1237.60	0.00	186-196	WGFD…INSR
1253.65	1253.65	0.00	442-453	SGFS…SIIR
1258.65	1258.65	0.00	203-213	LIGN…HAVR
1269.70	1269.69	0.01	483-494	AHTL..TVVR
1303.67	1303.67	0.00	969-979	IFTA…YDAR
1352.73	1352.71	0.02	1136-1147	DYTP…YVTK
1398.68	1398.67	0.01	120-131	EWEA…PALR
1424.67	1424.65	0.02	998-1009	GHVD…NNQR
1551.83	1551.81	0.02	895-906	EKLE…IVYK
1576.87	1576.81	0.04	686-698	NLLQ…DINR*
	1576.87	0.00	627-641	AVNA…LGLK*
1598.87	1598.71	0.16	1124-1135	RENP…RGYR
1625.77	1625.70	0.07	984-997	NGDF…WNVK
1794.92	1794.87	0.05	703-720	GWGG…DVFK
1800.88	1800.87	0.01	757-771	GYIE…YSIR
1900.92	1900.91	0.01	270-285	EIYT…GSFR
1902.99	1902.96	0.03	104-119	LEGL…ESFR
1956.02	1956.01	0.01	1010-1026	SVLVV…QEVR
2088.97	2088.94	0.03	1099-1117	GYNE…YEEK
2098.17	2098.15	0.02	864-882	LGNL…ALAR
2108.08	2108.09	0.01	606-623	FELI…NLER
2125.16	2125.18	0.02	354-372	IVAQ…TLYR
2133.12	2133.11	0.01	507-526	TSGG…LPQR
2149.04	2149.05	0.01	408-427	SGTV…VPPR
2160.14	2160.16	0.02	463-482	SAEF…PLVK

Table 8. (continued). Summary of the Tryptic Masses[1] Identified for the MON 87751-produced Cry1A.105 Using MALDI-TOF MS

Observed Mass	Expected Mass	Diff. [2]	Fragment	Sequence[3]
2195.15	2195.16	0.01	239-257	ELTL…YDSR
2197.10	2197.11	0.01	297-315	SPHL…DAHR
2277.25	2277.10	0.15	757-775	GYIE…YNAK*
	2277.15	0.10	407-427	KSGT…VPPR*
2375.30	2375.24	0.06	776-798	HETV…PIGK
2616.39	2616.36	0.03	945-968	EAYL…LEGR
3363.75	3363.63	0.12	910-939	ESVD…ADKR
3729.07	3728.87	0.20	373-406	RPFN…AVYR

[1]The observed mass was collected from at least one of three matrices including a-cyano, DHB and sinapinic acid. The observed mass shown is the mass closest to the expected mass.
[2]The data represent the calculated difference between the expected mass and the observed mass
*The expected peptide masses are nearly identical (< 1 dalton). Because this analysis did not determine with certainty which expected peptide was actually observed, the peptides with an asterisk (*) were not included in determining sequence coverage.
[3]For peptide matches greater than nine amino acids in length the first 4 residues and last 4 residues are show separated by three dots (…)

```
   1  CMQAMDNNPN INECIPYNCL SNPEVEVLGG ERIETGYTPI DISLSLTQFL
  51  LSEFVPGAGF VLGLVDIIWG IFGPSQWDAF LVQIEQLINQ RIEEFARNQA
 101  ISRLEGLSNL YQIYAESFRE WEADPTNPAL REEMRIQFND MNSALTTAIP
 151  LFAVQNYQVP LLSVYVQAAN LHLSVLRDVS VFGQRWGFDA ATINSRYNDL
 201  TRLIGNYTDH AVRWYNTGLE RVWGPDSRDW IRYNQFRREL TLTVLDIVSL
 251  FPNYDSRTYP IRTVSQLTRE IYTNPVLENF DGSFRGSAQG IEGSIRSPHL
 301  MDILNSITIY TDAHRGEYYW SGHQIMASPV GFSGPEFTFP LYGTMGNAAP
 351  QQRIVAQLGQ GVYRTLSSTL YRRPFNIGIN NQQLSVLDGT EFAYGTSSNL
 401  PSAVYRKSGT VDSLDEIPPQ NNNVPPRQGF SHRLSHVSMF RSGFSNSSVS
 451  IIRAPMFSWI HRSAEFNNII ASDSITQIPL VKAHTLQSGT TVVRGPGFTG
 501  GDILRRTSGG PFAYTIVNIN GQLPQRYRAR IRYASTTNLR IYVTVAGERI
 551  FAGQFNKTMD TGDPLTFQSF SYATINTAFT FPMSQSSFTV GADTFSSGNE
 601  VYIDRFELIP VTATLEAEYN LERAQKAVNA LFTSTNQLGL KTNVTDYHID
 651  QVSNLVTYLS DEFCLDEKRE LSEKVKHAKR LSDERNLLQD SNFKDINRQP
 701  ERGWGGSTGI TIQGGDDVFK ENYVTLSGTF DECYPTYLYQ KIDESKLKAF
 751  TRYQLRGYIE DSQDLEIYSI RYNAKHETVN VPGTGSLWPL SAQSPIGKCG
 801  EPNRCAPHLE WNPDLDCSCR DGEKCAHHSH HFSLDIDVGC TDLNEDLGVW
 851  VIFKIKTQDG HARLGNLEFL EEKPLVGEAL ARVKRAEKKW RDKREKLEWE
 901  TNIVYKEAKE SVDALFVNSQ YDQLQADTNI AMIHAADKRV HSIREAYLPE
 951  LSVIPGVNAA IFEELEGRIF TAFSLYDARN VIKNGDFNNG LSCWNVKGHV
1001  DVEEQNNQRS VLVVPEWEAE VSQEVRVCPG RGYILRVTAY KEGYGEGCVT
1051  IHEIENNTDE LKFSNCVEEE IYPNNTVTCN DYTVNQEEYG GAYTSRNRGY
1101  NEAPSVPADY ASVYEEKSYT DGRRENPCEF NRGYRDYTPL PVGYVTKELE
1151  YFPETDKVWI EIGETEGTFI VDSVELLLME E
```

Figure 18. MALDI-TOF MS Coverage Map of the MON 87751-produced Cry1A.105 Protein

The amino acid sequence of the MON 87751-produced Cry1A.105 protein was deduced from the *cry1A.105* gene present in MON 87751 and the identified N-terminus of the MON 87751-produced Cry1A.105. Boxed regions correspond to peptides that were identified from the MON 87751-produced Cry1A.105 protein sample using MALDI-TOF MS. In total, 53.5% (632 out of 1181amino acids) of the expected protein sequence was identified.

B1(a)(iii) Results of Western blot analysis of the Cry1A.105 protein

Western blot analysis was conducted using goat anti-Cry1A.105 polyclonal antibody as additional means to confirm the identity of the Cry1A.105 protein isolated from the seed of MON 87751 and to assess the equivalence of the immunoreactivity of the MON 87751-produced and *E. coli*-produced Cry1A.105 proteins.

The results showed that immunoreactive bands migrating at the expected apparent MW were present in all lanes loaded with the MON 87751-produced (Figure 19, Lanes 2-7) or *E.coli*-produced (Figure 19, Lanes 9-14) Cry1A.105 proteins. For each amount loaded, comparable signal intensity was observed between the MON 87751- and *E. coli*-produced Cry1A.105 protein bands. As expected, the signal intensity increased with increasing load amounts of the MON 87751- produced and *E. coli*-produced Cry1A.105 proteins, thus, supporting identification of MON 87751-produced Cry1A.105 protein.

To compare the immunoreactivity of the MON 87751-produced and the *E. coli*-produced Cry1A.105 proteins, densitometric analysis was conducted on bands that migrated to the expected apparent MW for Cry1A.105 proteins (~130 kDa). The signal intensity (reported in OD × mm^2) of the band of interest in lanes loaded with MON 87751-produced and the *E. coli*-produced Cry1A.105 protein was measured (Table 9). Because the mean signal intensity of the MON 87751-produced Cry1A.105 protein band was within ±35% of the mean signal intensity of the *E. coli*-produced Cry1A.105 protein, the MON 87751-produced and *E. coli*-produced Cry1A.105 proteins were determined to have equivalent immunoreactivity.

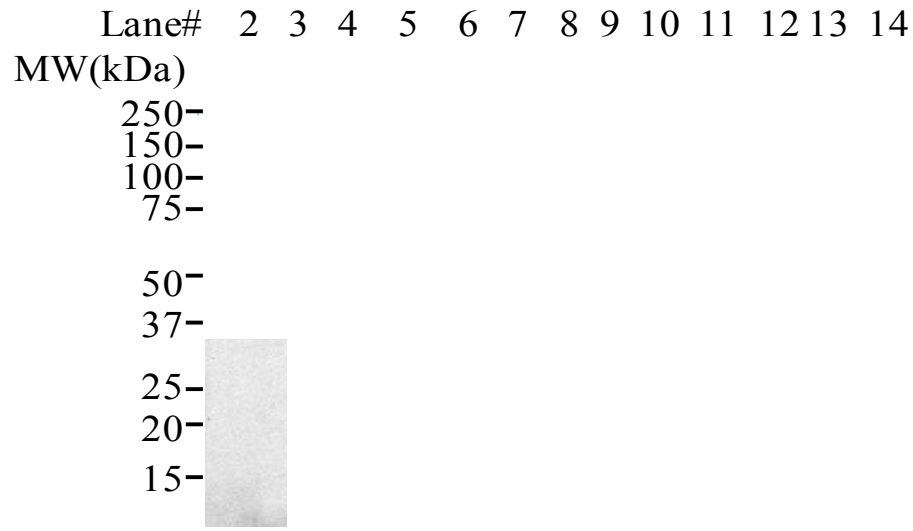


Figure 19. Western Blot Analysis of the MON 87751-produced and *E. coli* -produced Cry1A.105 Protein

Aliquots of the MON 87751-produced Cry1A.105 protein and the *E. coli*-produced Cry1A.105 protein were subjected to SDS-PAGE and electrotransferred to a PVDF membrane. Proteins were detected using anti-Cry1A.105 antibodies as the primary antibodies. Immunoreactive bands were visualized using HRP-conjugated secondary antibodies and an ECL system. The molecular weights (kDa) of the standards are shown on the left. The 4 min exposure is shown. Lanes 1 and 15 were cropped from the image. Lane designations are as follows:

Lane	Sample	Amount (ng)
1	Precision Plus Protein Standards Dual color	-
2	MON 87751-produced Cry1A.105	1
3	MON 87751-produced Cry1A.105	1
4	MON 87751-produced Cry1A.105	2
5	MON 87751-produced Cry1A.105	2
6	MON 87751-produced Cry1A.105	3
7	MON 87751-produced Cry1A.105	3
8	Blank	-
9	*E. coli*-produced Cry1A.105	1
10	*E. coli*-produced Cry1A.105	1
11	*E. coli*-produced Cry1A.105	2
12	*E. coli*-produced Cry1A.105	2
13	*E. coli*-produced Cry1A.105	3
14	*E. coli*-produced Cry1A.105	3
15	Blank	-

Table 9. Comparison of Immunoreactive Signals between MON 87751- and *E. coli* -produced Cry1A.105 Proteins

Mean signal intensity from MON 87751-produced Cry1A.105[1] (OD x mm^2)	Mean signal intensity from *E. coli* -produced Cry1A.105[1] (OD x mm^2)	Acceptance limits[2] (OD x mm^2)
2.67	3.33	2.16 – 4.50

[1]Each value represents the mean of six values (n=6)
[2] The acceptance limits are for the MON 87751-produced Cry1A.105 protein and are based on the interval between –35% (3.33 × 0.65 = 2.16) and +35% (3.33× 1.35 = 4.50) of the mean of the *E. coli*-produced Cry1A.105 signal intensity.

B1(a)(iv) Results of molecular weights analysis of the MON 87751 Cry1A.105 protein

The intact MON 87751-produced Cry1A.105 protein (Figure 20, Lanes 3-8) migrated to the same position on the gel as the *E. coli*-produced Cry1A.105 protein (Figure 20, Lane 2) and the apparent MW was calculated to be 132.9 kDa (Table 10). Because the experimentally determined apparent MW of the MON 87751-produced Cry1A.105 protein was within the acceptance limits for equivalence (Table 10), the MON 87751- and *E. coli*-produced Cry1A.105 proteins were determined to have equivalent apparent molecular weights.

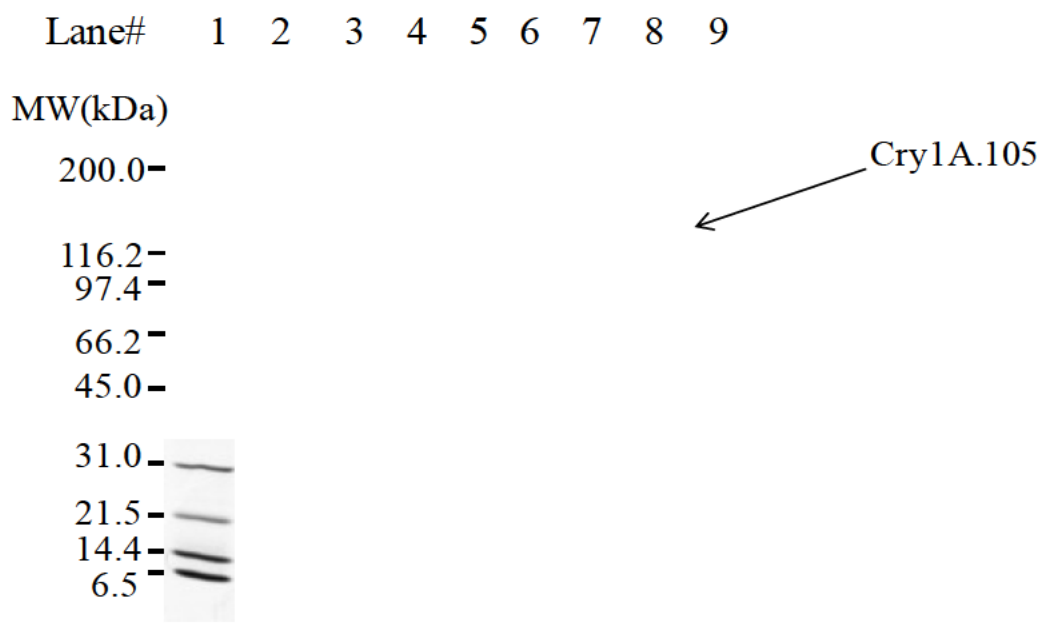


Figure 20. Molecular Weight Analysis of the MON 87751-produced Cry1A.105 Protein

Aliquots of the MON 87751-produced Cry1A.105 and the *E. coli*-produced Cry1A.105 proteins were subjected to SDS-PAGE and the gel was stained with Brilliant Blue G-Colloidal stain. The molecular weights (kDa) are shown on the left and correspond to the standards loaded in Lanes 1 and 9. Lane 10 was cropped from the image. Lane designations are as follows:

Lane	Sample	Amount (µg)
1	Broad Range MW Standards	4.5
2	*E. coli*-produced Cry1A.105	1.0
3	MON 87751-produced Cry1A.105	0.5
4	MON 87751-produced Cry1A.105	0.5
5	MON 87751-produced Cry1A.105	1.0
6	MON 87751-produced Cry1A.105	1.0
7	MON 87751-produced Cry1A.105	1.5
8	MON 87751-produced Cry1A.105	1.5
9	Broad Range MW Standards	4.5
10	Blank	-

Table 10. Molecular Weight Comparison Between the MON 87751- and *E. coli*-produced Cry1A.105 Proteins Based on SDS-PAGE

Apparent MW of MON 87751-produced Cry1A.105 Protein (kDa)[1]	Apparent MW of *E. coli*-produced Cry1A.105 Protein[2] (kDa)	Acceptance Limits (kDa)[3]
132.9	130.8	126.5 – 135.1

[1] The apparent MW represents the average of six loadings (n=6) of the MON 87751-produced Cry1A.105 (duplicate loadings of 0.5, 1.0 and 1.5 µgs of total protein). The final apparent MW was rounded to one decimal place

[2] The MW of the *E. coli* produced Cry1A.105 as reported on the Certificate of Analysis for lot 11349124.

[3]Calculated lower and upper bounds for this assay based on two-tailed 95% prediction interval derived from a historical data set of 8 apparent MW determinations for Cry1A.105

B1(a)(v) Results of assessing whether MON 87751-produced Cry1A.105 is glycosylated

Proteins expressed in eukaryotes can be post-translationally modified with carbohydrate moieties (Rademacher *et al.*, 1988). To test whether Cry1A.105 protein was glycosylated when expressed in the seed of MON 87751, the MON 87751-produced Cry1A.105 protein was analyzed using an ECL™ Glycoprotein Detection Module (GE Healthcare).

No glycosylation signal was observed in the molecular weight range of 130 kDa in the lanes containing the MON 87751-produced Cry1A.105 protein (Figure 21, Panel A, Lanes 6 and 7) or the *E. coli*-produced Cry1A.105 protein (Figure 21, Panel A, Lanes 8 and 9). There is a signal of ~40 kDa in the lanes of MON 87751-produced Cry1A.105 protein, which is not likely derived from Cry1A.105 since no immunoreactive signal was observed at this molecular weight range in the western blot analysis. This low molecular weight signal is likely from glycosylated soy protein that was co-purified with Cry1A.105 protein. A clear glycosylation signal was observed as expected in the lanes containing the positive control, transferrin, at its expected MW (~80 kDa) (Figure 21, Panel A, Lanes 2-5), demonstrating that the assay performed as expected.

To confirm that MON 87751-produced and *E. coli*-produced Cry1A.105 proteins were appropriately loaded for glycosylation analysis, a second membrane with identical loadings and transfer time was stained with Coomassie Blue R250 for protein detection (Figure 21, Panel B). Both the MON 87751-produced Cry1A.105 (Figure 21, Panel B, Lanes 6 and 7) and *E. coli*-produced Cry1A.105 (Figure 21, Panel B, Lanes 8 and 9) proteins were detected.

These data indicate that MON 87751-produced Cry1A.105 protein is not glycosylated and is equivalent to that of the *E. coli* -produced Cry1A.105 protein.

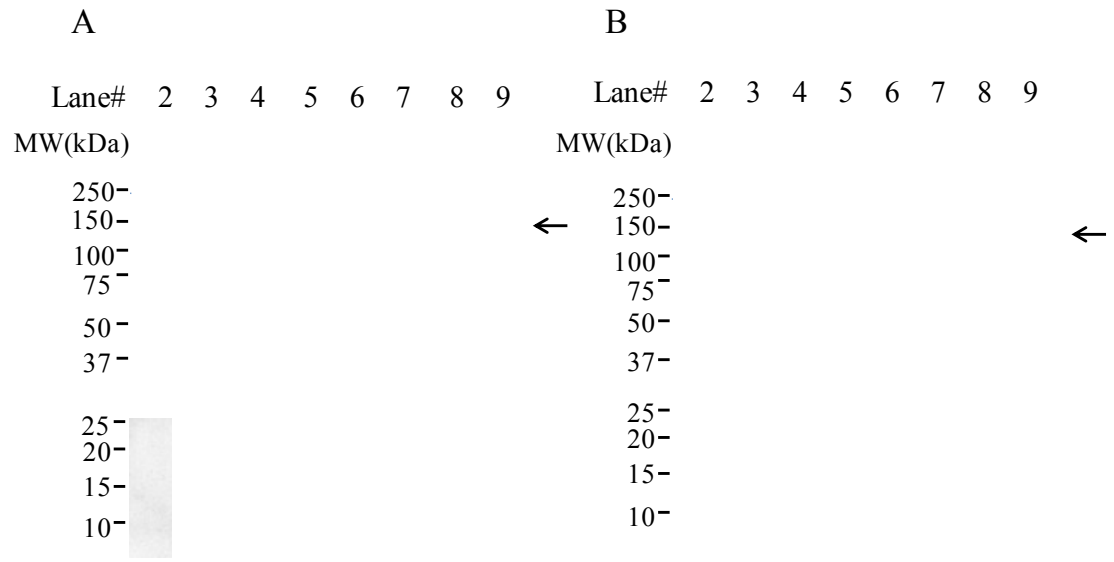


Figure 21. Glycosylation Analysis of the MON 87751-produced Cry1A.105 Protein

Aliquots of the transferrin (positive control), MON 87751-produced Cry1A.105 and *E. coli* -produced Cry1A.105 were subjected to SDS-PAGE and electrotransferred to PVDF membranes. The molecular weights (kDa) of the standards are shown on the left. The arrows show the expected migration of the MON 87751- and *E. coli*-produced Cry1A.105 protein. (A) Where present, the labeled carbohydrate moieties were detected by addition of streptavidin conjugated to HRP followed by a luminol-based detection using ECL reagents and exposure to Hyperfilm®. The 45 seconds exposure is shown. (B) An equivalent blot was stained with Coomassie Blue R250 to confirm the presence of proteins. Lanes 1 and 10 were cropped from both images. Lane designations are as follows:

Lane	Sample	Amount (ng)
1	Precision Plus Protein™ Standards	
2	Transferrin (positive control)	200
3	Transferrin (positive control)	150
4	Transferrin (positive control)	100
5	Transferrin (positive control)	50
6	MON 87751-produced Cry1A.105	100
7	MON 87751-produced Cry1A.105	200
8	*E. coli*-produced Cry1A.105	100
9	*E. coli*-produced Cry1A.105	200
10	Blank	

B1(a)(vi) Cry1A.105 results of the functional activity analysis

The functional activity of the MON 87751-produced and *E. coli*-produced Cry1A.105 protein was determined by corn earworm diet incorporation assay. In this assay, activity is expressed as $EC_{50,}$ µg/ml diet. The MON 87751- and *E. coli*-produced Cry1A.105 proteins were considered functionally equivalent if the EC_{50} of both were within acceptance limits of 0.0028 µg/ml diet to 0.0091 µg/ml diet; which is derived from the 95% prediction interval calculated from data obtained for the *E. coli*-produced Cry1A.105 protein activity. The EC_{50} of the MON 87751-produced and *E. coli*-produced Cry1A.105 proteins were determined to be 0.0035 µg/ml diet and 0.0032 µg/ml diet; respectively (Table 11). Because the EC_{50} of MON 87751-produced and *E. coli*-produced Cry1A.105 proteins were both within preset acceptance limits (Table 11), the proteins were determined to have equivalent functional activity.

Table 11. Cry1A.105 Functional Assay

MON 87751-produced Cry1A.105[1] EC_{50} (µg /ml)	*E. coli*-produced Cry1A.105[1] EC_{50} (µg /ml)	Acceptance Limits[2] EC_{50} (µg /ml)
0.0035	0.0032	0.0028 – 0.0091

[1]Value refers to mean based on n = 3.

[2] Calculated lower and upper bounds for this assay based on two-tailed 95% prediction interval derived from historical data sets of 11 EC_{50} determinations for Cry1A.105. Values in this column represent a 95% prediction interval developed from a series of 11 assays with *E. coli*- produced Cry1A.105 protein. Eight assays relied on a protein with an amino acid sequence identical to the MON 89034-produced Cry1A.105 protein and three relied on a protein with an amino acid sequence identical to the MON 87751-produced protein.

B1(a)(vii) Cry1A.105 protein identity and equivalence conclusion

A panel of analytical techniques was used to characterize the MON 87751-produced Cry1A.105 protein purified from seed of MON 87751. In addition, the equivalence of the MON 87751-produced and *E. coli*-produced Cry1A.105 proteins was evaluated by comparing their apparent MW, immunoreactivity with anti-Cry1A.105 antibodies, glycosylation status, and functional activity. The results demonstrate that the MON 87751-produced Cry1A.105 protein and the *E. coli*-produced Cry1A.105 protein are equivalent. Furthermore, the demonstration of functional activity equivalence of MON 87751- and *E. coli*-produced proteins allows for the existing Cry1A.105 protein evaluations on acute toxicity, NTOs and activity spectrum to be applicable to the Cry1A.105 protein produced by MON 87751.

[redacted]

B1(b) Cry2Ab2 protein identity and equivalence

The safety assessment of crops derived through biotechnology includes characterization of the physicochemical and functional properties of the protein(s) produced from the inserted DNA, and confirmation of the safety of the protein(s). For the safety data generated using *E. coli*-produced Cry2Ab2 to be applied to Cry2Ab2 protein produced in MON 87751, the

equivalence of the plant- and *E. coli*-produced proteins must be assessed. The MON 87751-produced Cry2Ab2 protein purified from soybean of MON 87751 was characterized and the equivalence of the physicochemical and functional properties between the MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins was established using a panel of analytical tests: 1) N-terminal sequence analysis of MON 87751-produced Cry2Ab2 characterized the N-terminus; 2) MALDI-TOF MS analysis yielded peptide masses consistent with the expected peptide masses from the theoretical trypsin digest of the MON 87751-produced Cry2Ab2 sequence; 3) western blot analysis with antibodies specific for Cry2Ab2 protein demonstrated that the immunoreactive properties of the MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins were equivalent; 4) SDS-PAGE analysis showed that the electrophoretic mobility and apparent molecular weight of the MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins were equivalent; 5) MON 87751-produced and *E coli*-produced Cry2Ab2 proteins were both determined to be non-glycosylated; and 6) functional activity analysis demonstrated that MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins had equivalent insecticidal activity. Taken together, these data provide a detailed characterization of the MON 87751-produced Cry2Ab2 protein and establish its equivalence to the *E. coli*-produced Cry2Ab2 protein.

B1(b)(i) Results of the N-terminal sequencing analysis

Fifteen cycles of N-terminal sequencing was performed on the MON 87751-produced Cry2Ab2 protein. The sequence obtained corresponded to the N-terminal sequence for Cry2Ab2 starting at position 16 relative to the first methionine of the predicted Cry2Ab2 sequence. The N-terminal sequencing results for MON 87751-produced Cry2Ab2 protein were consistent with the sequencing results for the E. coli-produced Cry2Ab2 protein, which was engineered to start at the same amino acid position as the MON 87751-produced Cry2Ab2 protein (see Figure 22). Hence, the sequence information confirms the identity of the Cry2Ab2 protein isolated from the seed of MON 87751.

Amino acid residue # from the N-terminus		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Expected Sequence	→	A	Y	N	V	A	A	H	D	P	F	S	F	Q	H	K
		|	|	|	|	|	|	|	|	|	|	|	|	|	|	|
Experimental Sequence	→	A	Y	N	V	A	A	H	D	P	F	S	X	Q	X	X

Figure 22. N-Terminal Sequence of the MON 87751 Cry2Ab2 Protein

The experimental sequence obtained from the MON 87751-produced Cry2Ab2 was compared to the expected sequence deduced from the *cry2Ab2* gene present in MON 87751. The amino acid at position 16 of the deduced sequence is shown as position 1 of the observed protein. The single letter IUPAC-IUB amino acid code is; A, alanine; D, aspatic acid; F, phenylalanine; H, histidine; K, lysine; N, asparagine; P, proline; Q, glutamine; S, serine; V, valine, Y tyrosine. X indicates that the residue was not identified.

B1(b)(ii) Results of MALDI-TOF tryptic mass map analysis

Peptide mass fingerprint analysis is a standard technique used for confirming the identity of proteins. The identity of the MON 87751-produced Cry2Ab2 protein was confirmed by MALDI-TOF MS analysis of peptide fragments produced by the trypsin digestion of the MON 87751-produced Cry2Ab2 protein. In general, protein identification made by proteolytic peptide mapping is considered to be reliable if the measured coverage of the sequence is 15% or higher with a minimum of five matched peptides (Jensen *et al.*, 1997). There were 39 unique peptides identified that corresponded to the expected masses (Table 12). The identified masses were used to assemble a coverage map of the Cry2Ab2 protein (Figure 23). The experimentally determined coverage of the Cry2Ab2 protein was 66% (406 out of 619 amino acids). Both acetylated and non-acetylated N-terminal peptides were present in the sample (Table 12)

Table 12. Summary of the Tryptic Masses Identified for the MON 87751-produced Cry2Ab2 Protein Using MALDI-TOF MS

Observed Mass[1]	Expected Mass	Diff.[2]	Fragment	Sequence[3]
439.05	439.23	0.18	193-195	TYR
506.22	506.25	0.03	82-85	ETEK
552.39	552.31	0.08	520-523	YTLR
560.39	560.32	0.07	219-223	GLNTR
646.32	646.32	0.00	421-425	NEDLR
677.38	677.37	0.01	86-90	FLNQR
709.38	709.36	0.02	392-398	SGAFTAR
724.39	724.39	0.00	496-501	TFISDEK
730.46	730.45	0.01	59-64	ILSELR
815.62	815.51	0.11	51-58	VGSLVGKR
878.44	878.40	0.04	23-28	EWTEWK
886.66	886.55	0.11	58-64	RILSELR
903.52	903.49	0.03	91-98	LNTDTLAR
958.55	958.50	0.05	193-199	TYRDYLK
993.50	993.48	0.02	502-510	FGNQGDSLR
1006.52	1006.50	0.02	23-29	EWTEWKK
1022.46	1022.45	0.01	362-370	SWLDSGSDR
1033.70	1033.56	0.14	535-544	VSSIGNSTIR
1053.65	1053.64	0.01	411-420	NISGVPLVVR
1060.59	1060.52	0.07	224-231	LHDMLEFR
1076.57	1076.53	0.04	448-456	AYMVSVHNR
1080.57	1080.51	0.06	511-519	FEQNNTTAR
1184.67	1184.60	0.07	435-447	NIAS...GGAR
1197.71	1197.65	0.06	426-434	RPLHYNEIR
1216.65	1216.61	0.04	116-125	QVDNFLNPNR
1270.71	1270.62	0.09	524-534	GNGN...LYLR
1492.74	1492.69	0.05	399-410	GNSN...YFIR**
	1492.75	0.01	193-203	TYRD...NYTR
1732.40	1731.82	0.58	1-15	AYNV….FQHK
1773.96	1773.82	0.14	1-15	[4]Ac-AYNV….FHQK
1844.05	1843.96	0.09	176-192	DVIL...ATLR
1904.09	1903.95	0.14	99-115	VNAE...EFNR
1919.11	1919.01	0.10	65-81	NLIF...DILR
2201.31	2201.20	0.11	30-49	NNHS...FLLK
2311.58	2311.06	0.52	551-572	VYTA...NGAR
2333.40	2333.23	0.17	126-146	NAVP...FLNR
2339.31	2339.15	0.16	371-391	EGVA...LGLR
2451.44	2451.34	0.10	302-325	LSNT...LAAR
4212.98	4212.04	0.94	458-495	NNIH...NQTR

[1]The observed mass was collected from at least one of three matrices including a-cyano, DHB and sinapinic acid. The observed mass shown is the mass closest to the expected mass.

[2]The data represent- the calculated difference between the expected mass and the observed mass

**The expected peptide masses are nearly identical (< 1 dalton). Because this analysis did not determine with certainty which expected peptide was actually observed, the peptides with an asterisk (*) were not included in determining sequence coverage (Figure 23).

[3]For peptide matches greater than nine amino acids in length the first 4 residues and last 4 residues are show separated by three dots (…). Fragment numbering is based on the observed N-terminus of the protein.

[4] AC is the abbreviation for acetylation

001 AYNVAAHDPF SFQHKSLDTV QKEWTEWKKN NHSLYLDPIV GTVASFLLKK

051 VGSLVGKRIL SELRNLIFPS GSTNLMQDIL RETEKFLNQR LNTDTLARVN

101 AELTGLQANV EEFNRQVDNF LNPNRNAVPL SITSSVNTMQ QLFLNRLPQF

151 QMQGYQLLLL PLFAQAANLH LSFIRDVILN ADEWGISAAT LRTYRDYLKN

201 YTRDYSNYCI NTYQSAFKGL NTRLHDMLEF RTYMFLNVFE YVSIWSLFKY

251 QSLLVSSGAN LYASGSGPQQ TQSFTSQDWP FLYSLFQVNS NYVLNGFSGA

301 RLSNTFPNIV GLPGSTTTHA LLAARVNYSG GISSGDIGAS PFNQNFNCST

351 FLPPLLTPFV RSWLDSGSDR EGVATVTNWQ TESFETTLGL RSGAFTARGN

401 SNYFPDYFIR NISGVPLVVR NEDLRRPLHY NEIRNIASPS GTPGGARAYM

451 VSVHNRKNNI HAVHENGSMI HLAPNDYTGF TISPIHATQV NNQTRTFISE

501 KFGNQGDSLR FEQNNTTARY TLRGNGNSYN LYLRVSSIGN STIRVTINGR

651 VYTATNVNTT TNNDGVNDNG ARFSDINIGN VVASSNSDVP LDINVTLNSG

601 TQFDLMNIML VPTNISPLY

Figure 23. MALDI-TOF MS Coverage Map of the MON 87751-produced Cry2Ab2 Protein

The amino acid sequence of the MON 87751-produced Cry2Ab2 protein was deduced from the *cry2Ab2* gene present in MON 87751, with the amino acid at position 16 of the deduced sequence shown as position 1 of the observed protein. Boxed regions correspond to peptides that were identified from the MON 87751-produced Cry2Ab2 protein sample using MALDI-TOF MS. In total, 66% (406 out of 619 amino acids) of the expected protein sequence was identified.

B1(b)(iii) Results of Western blot analysis of the Cry2Ab2 protein

Western blot analysis was conducted using anti-Cry2Ab2 monoclonal antibody as additional means to confirm the identity of the Cry2Ab2 protein isolated from the seed of MON 87751 and to assess the equivalence of the immunoreactivity of the MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins.

The results showed that immunoreactive bands migrating at the expected apparent MW were present in all lanes loaded with the MON 87751-produced (Figure 24, Lanes 9-14) or *E. coli*-produced (Figure 24, Lanes 2-7) Cry2Ab2 proteins. For each amount loaded, comparable signal intensity was observed between the MON 87751- and *E. coli*-produced Cry2Ab2 protein bands. As expected, the signal intensity increased with increasing load amounts of the MON 87751- produced and *E. coli*-produced Cry2Ab2 proteins, thus, supporting identification of MON 87751-produced Cry2Ab2 protein.

To compare the immunoreactivity of the MON 87751-produced and the *E. coli*-produced Cry2Ab2 proteins, densitometric analysis was conducted on bands that migrated to the expected apparent MW for Cry2Ab2 proteins (~60 kDa). The signal intensity (reported in OD × mm^2) of the band of interest in lanes loaded with MON 87751-produced and the *E. coli*-produced Cry2Ab2 protein was measured (Table 13). Because the mean signal intensity of the MON 87751-produced Cry2Ab2 protein band was within ±35% of the mean signal of the *E. coli*-produced Cry2Ab2 protein, the MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins were determined to have equivalent immunoreactivity.

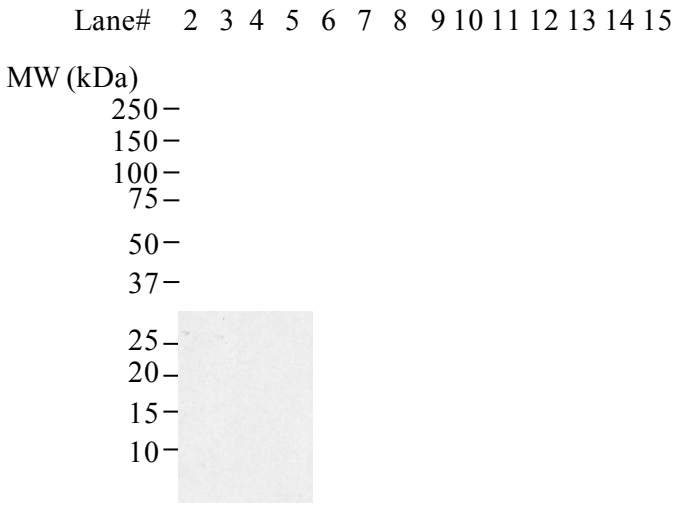


Figure 24. Western Blot Analysis of the MON 87751-produced and *E. coli* -produced Cry2Ab2 Protein

Aliquots of the MON 87751-produced Cry2Ab2 protein and the *E. coli*-produced Cry2Ab2 protein were subjected to SDS-PAGE and electrotransferred to a nitrocellulose membrane. Proteins were detected using anti-Cry2Ab2 antibodies as the primary antibodies. Immunoreactive bands were visualized using HRP-conjugated secondary antibodies and an ECL system. The molecular weights (kDa) of the standards are shown on the left. The 2 min exposure is shown. Lane 1 was cropped from the image. Lane designations are as follows:

Lane	Sample	Amount (ng)
1	Precision Plus Protein Standards Dual color	
2	*E. coli*-produced Cry2Ab2	1
3	*E. coli*-produced Cry2Ab2	1
4	*E. coli*-produced Cry2Ab2	2
5	*E. coli*-produced Cry2Ab2	2
6	*E. coli*-produced Cry2Ab2	3
7	*E. coli*-produced Cry2Ab2	3
8	Blank	-
9	MON 87751-produced Cry2Ab2	1
10	MON 87751-produced Cry2Ab2	1
11	MON 87751-produced Cry2Ab2	2
12	MON 87751-produced Cry2Ab2	2
13	MON 87751-produced Cry2Ab2	3
14	MON 87751-produced Cry2Ab2	3
15	Blank	

Table 13. Comparison of Immunoreactive Signals between MON 87751- and *E. coli* -produced Cry2Ab2 Proteins

Mean signal intensity from MON 87751-produced Cry2Ab2[1] (OD x mm^2)	Mean signal intensity from *E. coli*-produced Cry2Ab2[1] (OD x mm^2)	Acceptance limits[2] (OD x mm^2)
5.04	5.46	3.55-7.37

[1]Each value represents the mean of six values (n=6)
[2] The acceptance limits are for the MON 87751-produced Cry2Ab2 protein and are based on the interval between +35% (5.46 × 1.35 = 7.37) and −35% (5.46 × 0.65 = 3.55) of the mean of the *E. coli* -produced Cry2Ab2 signal intensity across all loads.

B1(b)(iv) Results of molecular weights analysis of the MON 87751 Cry2Ab2 protein

The intact MON 87751-produced Cry2Ab2 protein (Figure 25, Lanes 3-8) migrated to the same position on the gel as the *E. coli*-produced Cry2Ab2 protein (Figure 25, Lane 2) and the apparent MW was calculated to be 61.4 kDa (Table 14). Because the experimentally determined apparent MW of the MON 87751-produced Cry2Ab2 protein was within the acceptance limits for equivalence (Table 14), the MON 87751- and *E. coli*-produced Cry2Ab2 proteins were determined to have equivalent apparent molecular weights.

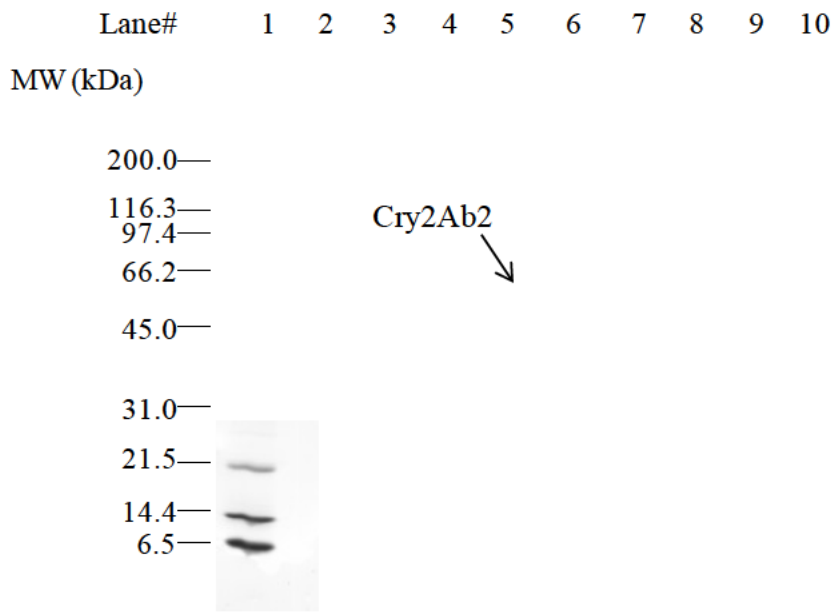


Figure 25. Molecular Weight Analysis of the MON 87751-produced Cry2Ab2 Protein

Aliquots of the MON 87751-produced Cry2Ab2 and the *E. coli*-produced Cry2Ab2 proteins were subjected to SDS-PAGE and the gel was stained with Brilliant Blue G-Colloidal stain. 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 (µg)
1	Broad Range MW Standards	4.5
2	*E. coli*-produced Cry2Ab2	1
3	MON 87751-produced Cry2Ab2	1
4	MON 87751-produced Cry2Ab2	1
5	MON 87751-produced Cry2Ab2	2
6	MON 87751-produced Cry2Ab2	2
7	MON 87751-produced Cry2Ab2	3
8	MON 87751-produced Cry2Ab2	3
9	Broad Range MW Standards	4.5
10	Blank	-

Table 14. Molecular Weight Comparison Between the MON 87751- and *E. coli*-produced Cry2Ab2 Proteins Based on SDS-PAGE

Apparent MW of MON 87751-produced Cry2Ab2Protein (kDa)[1]	Apparent MW of *E. coli*-produced Cry2Ab2 Protein[2] (kDa)	Acceptance Limits (kDa)[3]
61.4	60.1	58.7 - 61.5

[1] The apparent MW represents the average of six loadings (n=6) of the MON 87751-produced Cry2Ab2 (duplicate loadings of 1.0, 2.0 and 3.0 µgs of total protein). The final apparent MW was rounded to one decimal place

[2] The MW of the *E. coli* produced Cry2Ab2 as reported on the Certificate of Analysis for lot 11351673.

[3]Calculated lower and upper bounds for this assay based on two-tailed 95% prediction interval derived from a historical data set of eight apparent MW determinations for Cry2Ab2

B1(b)(v) Results of assessing whether MON 87751-produced Cry2Ab2 is glycosylated

Proteins expressed in eukaryotes can be post-translationally modified with carbohydrate moieties (Rademacher *et al.*, 1988). To test whether Cry2Ab2 protein was glycosylated when expressed in the seed of MON 87751, the MON 87751-produced Cry2Ab2 protein was analyzed using an ECL™ Glycoprotein Detection Module (GE Healthcare).

No glycosylation signal was observed in the molecular weight range of 61 kDa in the lanes containing the MON 87751-produced Cry2Ab2 protein (Figure 26, Panel A, Lanes 9 and 10) or the *E. coli*-produced Cry2Ab2 protein (Figure 26, Panel A, Lanes 7 and 8). A clear glycosylation signal was observed as expected in the lanes containing the positive control, transferrin, at its expected MW (~80 kDa) (Figure 26, Panel A, Lanes 2-5), demonstrating that the assay performed as expected.

To confirm that MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins were appropriately loaded for glycosylation analysis, a second membrane with identical loadings and transfer time was stained with Coomassie Blue R250 for protein detection (Figure 26, Panel B). Both the MON 87751-produced Cry2Ab2 (Figure 26, Panel B, Lanes 9 and 10) and *E. coli*-produced Cry2Ab2 (Figure 26, Panel B, Lanes 7 and 8) proteins were detected.

These data indicate that MON 87751-produced Cry2Ab2 protein is not glycosylated and is equivalent to that of the *E. coli*-produced Cry2Ab2 protein.

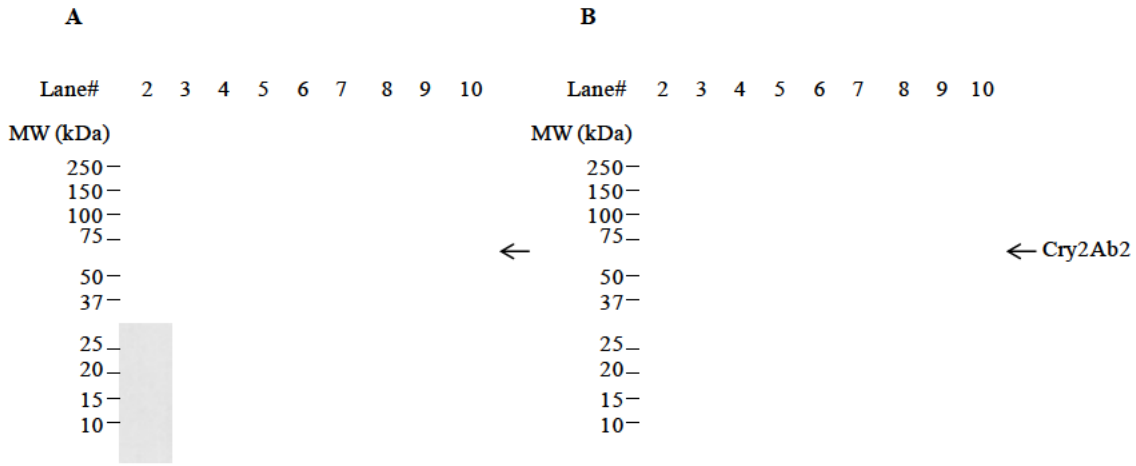


Figure 26. Glycosylation Analysis of the MON 87751-produced Cry2Ab2 Protein

Aliquots of the transferrin (positive control), *E. coli*-produced Cry2Ab2 and MON 87751-produced Cry2Ab2 were subjected to SDS-PAGE and electrotransferred to PVDF membranes. The MWs (kDa) correspond to the Precision Plus Protein™ Standards. Lane 1 in both images was loaded with MW standards and were cropped. The arrows show the expected migration of the MON 87751- and *E. coli*-produced Cry2Ab2 protein. (A) Where present, the labeled carbohydrate moieties were detected by addition of streptavidin conjugated to HRP followed by a luminol-based the detection using ECL reagents and exposure to Hyperfilm®. (B) An equivalent blot was stained with Coomassie Blue R250 to confirm the presence of proteins. Lane designations are as follows:

Lane	Sample	Amount (ng)
1	Precision Plus Protein™ Standards	
2	Transferrin (positive control)	50
3	Transferrin (positive control)	100
4	Transferrin (positive control)	150
5	Transferrin (positive control)	200
6	Blank	
7	*E. coli* -produced Cry2Ab2	100
8	*E. coli* -produced Cry2Ab2	200
9	MON 87751-produced Cry2Ab2	100
10	MON 87751-produced Cry2Ab2	200

B1(b)(vi) Cry2Ab2 results of the functional activity analysis

The functional activity of the MON 87751- produced and *E. coli*-produced Cry2Ab2 protein was determined by corn earworm diet incorporation assay. In this assay, activity is expressed as $EC_{50,}$ µg/ml diet. The MON 87751- and *E. coli*-produced Cry2Ab2 proteins were considered functionally equivalent if the EC_{50}, of both were within acceptance limits of 0.049 µg/ml diet to 0.204 µg/ml diet; the 95% prediction interval calculated from data obtained for the *E. coli*-produced Cry2Ab2 protein activity. The EC_{50} of the MON 877511-produced and *E. coli*-produced Cry2Ab2 proteins were determined to be 0.0734 µg/ml diet and 0.115 µg/ml diet; respectively (Table 15). Because the EC_{50} of MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins were both within preset acceptance limits (Table 15), the proteins were determined to have equivalent functional activity.

Table 15. Cry2Ab2 Functional Assay

MON 87751-produced Cry2Ab2[1] EC_{50} (µg Cry2Ab2/ml diet)	***E. coli*-produced Cry2Ab2[1] EC_{50} (µg Cry2Ab2/ml diet)**	**Acceptance Limits[2] EC_{50} (µg /ml diet)**
0.0734	0.1145	0.049 – 0.204

[1]Value refers to mean calculated based on n = 2.

[2] Calculated lower and upper bounds for this assay based on two-tailed 95% prediction interval derived from historical data sets of 11 EC_{50} determinations for Cry2Ab2. Values in this column represent a 95% prediction interval developed from a series of 11 assays with *E. coli*-produced Cry2Ab2 protein. Eight assays relied on a protein with an amino acid sequence identical to the MON 89034-produced Cry2Ab2 protein and three relied on a protein with an amino acid sequence identical to the MON 87751-produced protein.

B1(b)(vii) Cry2Ab2 protein identity and equivalence conclusion

A panel of analytical techniques was used to characterize the MON 87751-produced Cry2Ab2 protein purified from seed of MON 87751. In addition, the equivalence of the MON 87751-produced and *E. coli*-produced Cry2Ab2 proteins was evaluated by comparing their apparent MW, immunoreactivity with anti-Cry2Ab2 antibodies, glycosylation status, and functional activity. The results demonstrate that the MON 87751-produced Cry2Ab2 protein and the *E. coli*-produced Cry2Ab2 protein are equivalent.

[redacted]

B2 Antibiotic Resistance Marker Genes

MON 87751 does not contain genes that encode resistance to antibiotic markers. Molecular characterisation data presented in Section A demonstrate the absence of antibiotic resistance marker genes in MON 87751.

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 protein(s)

Cry1A.105 and Cry2Ab2 proteins are derived from *B. thuringiensis*, a ubiquitous gram-positive soil bacterium that accumulates crystal proteins during sporulation. Cry1A.105 and Cry2Ab2 proteins are members of the 3D-Cry family of insecticidal proteins (Crickmore, 2012). 3D-Cry proteins are subdivided into different groups based on the high specificity they have for their target category of insects. Because of their narrow spectrum of activity, they lack an impact on broader insect populations or other organisms. For example, both Cry1A.105 and Cry2Ab2 proteins have insecticidal activity specifically against lepidopteran insects (Höfte and Whiteley, 1989).

The generalized MOA for Cry proteins was described by English and Slatin (1992). It includes ingestion of the crystals by insects and solubilization of the crystals in the insect midgut, followed by activation through proteolytic processing of the soluble Cry protein by digestive enzymes in their midguts. The activated protein then binds to specific receptors on the surface of the midgut epithelium of target insects and inserts into the membrane, leading to pore formation and generalized disruption of the transmembrane gradients and, therefore, cell integrity. While alternate mechanisms have been proposed, a review of the available data has recently been published and the authors concluded that the original model, pore formation, is the most valid model for Cry protein MOA (Vachon *et al.*, 2012).

B3(a)(i) Mode-of-action of Cry1A.105 Protein from MON 87751

Cry1A.105 in MON 87751 is a protein consisting of a single polypeptide of 1181 amino acids containing three domains with an apparent molecular weight of approximately 133 kDa. Like other Cry proteins, it is synthesized as a prototoxin and is cleaved by digestive enzymes in the midgut of target organisms to an approximately 60 kDa activated protein (Bravo *et al.*, 2007). Cry1A.105 is a chimeric protein that consists of domains I and II from Cry1Ab or Cry1Ac[6], domain III from Cry1F, and the C-terminal domain from Cry1Ac (Figure 27). Cry1Ac, Cry1Ab and Cry1F are all well known and well characterized insecticidal proteins derived from the soil bacterium *B. thuringiensis*. Cry1A.105 was designed using a domain exchange strategy to achieve high levels of activity against target lepidopteran insect pests. Domain exchange is a well known occurring mechanism in nature, resulting in Cry protein diversities that have been described extensively in the literature (de Maagd *et al.*, 2003; de Maagd *et al.*, 2001). Domain exchange with modern molecular biological tools has been used to switch the functional domains of Cry1 proteins to develop microbial biopesticides with improved specificity to lepidopteran insect pests. (Gao *et al.*, 2006; Baum, 1998; Baum *et al.*, 1999). Domains I and II of Cry1A.105 are 100% identical to the respective domains of Cry1Ab or Cry1Ac. Domain III of Cry1A.105 is 99% identical to domain III of Cry1F. The C-terminal region of Cry1A.105 is 100% identical to that of Cry1Ac. The overall amino acid sequence identity of Cry1A.105 to Cry1Ac, Cry1Ab, and Cry1F is 93.6%, 90.0%, and 76.7 %, respectively.

---

[6] Cry1Ab and Cry1Ac share 100% amino acid sequence identity in domains I and II.

The Cry1A.105 protein expressed in MON 87751 is targeted to chloroplasts through the addition of a chloroplast transit peptide (CTP) coding sequence at the 5′ end of the coding sequence. Following translation and translocation into chloroplasts, the CTP is cleaved. Experimental analysis of the N-terminus of MON 87751-produced Cry1A.105 protein indicated the presence of four additional amino acids at the N-terminus compared to the Cry1A.105 protein in MON 89034 (Figure 28). The additional four amino acids are cysteine (C), methionine (M), glutamine (Q), and alanine (A). While the identities of methionine, glutamine, and alanine were clearly determined by N-terminal sequencing, the identity of the first amino acid, cysteine, was inferred based on the *RbcS*4 targeting sequence in MON 87751. The chemistry employed in N-terminal sequencing is known to degrade cysteine, preventing its clear identification. With the exception of the four additional CTP-derived amino acids, the deduced sequence of the Cry1A.105 protein that accumulates in MON 87751 shares 100% amino acid identity with the deduced sequence of the Cry1A.105 protein present in MON 89034. The presence of these four additional amino acids at the N-terminus of the MON 87751-produced Cry1A.105 protein are unlikely to impact protein specificity because they are not within the trypsin-resistant core that is responsible for target organism specificity and efficacy. Accordingly, this small difference is not expected to result in unanticipated adverse impacts on humans, livestock or NTOs (Hammond *et al.*, 2013).

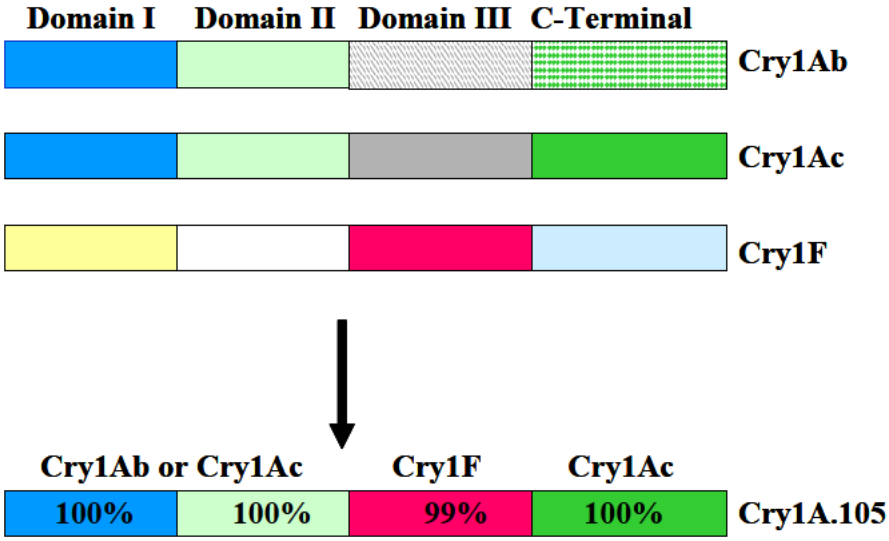


Figure 27. Schematic Representation of the Origin of Cry1A.105 Protein Domains

Different colors and patterns are used to differentiate the origin of domains. For simplicity, the lengths of domains in this illustration are not in proportion to the lengths of amino acid sequence of the respective domains.

B3(a)(ii) Mode-of-action of the Cry2Ab2 protein from MON 87751

Cry2Ab2 protein in MON 87751 is a protein consisting of a single polypeptide of 619 amino acids with an apparent molecular weight of approximately 62 kDa. Like other Cry proteins, it is synthesized as a prototoxin and is likely cleaved by digestive enzymes in the midgut of target organisms to an approximately 60 kDa activated protein (Bravo *et al.*, 2007).

The protein coding region of the *cry2Ab2* coding sequence present in MON 87751 is identical to the *cry2Ab2* coding sequence present in MON 89034 and both are a slight variant of the wild-type *cry2Ab2* coding sequence from *B. thuringiensis*. Accumulation of the Cry2Ab2 protein in MON 87751 is targeted to the chloroplasts due to the addition of a CTP coding sequence at the 5′ end of the coding sequence (refer to Section III). Experimental analysis of the N-terminus of MON 87751-produced Cry2Ab2 protein (described below) indicated that the conjunction of the CTP sequence with the Cry2Ab2 sequence resulted in CTP cleavage at a position 15 amino acids within the Cry2Ab2 protein sequence, likely as a result of the processing of the N-terminal CTP by a general stromal processing peptidase (SPP) (Richter and Lamppa, 1998). This deletion results in an amino acid sequence for the MON 87751-derived Cry2Ab2 protein that is 18 amino acids shorter than the MON 89034-derived Cry2Ab2 protein (Figure 28) because the MON 89034-derived protein is three amino acids longer than wild type Cry2Ab2 (U.S. EPA, 2010). The deletion of these amino acids at the N-terminus of the MON 87751-produced Cry2Ab2 protein are unlikely to impact protein specificity because they are not within the trypsin-resistant core that is responsible for target organism specificity and efficacy. Accordingly, this small difference is not expected to result in unanticipated adverse impacts on humans or livestock.

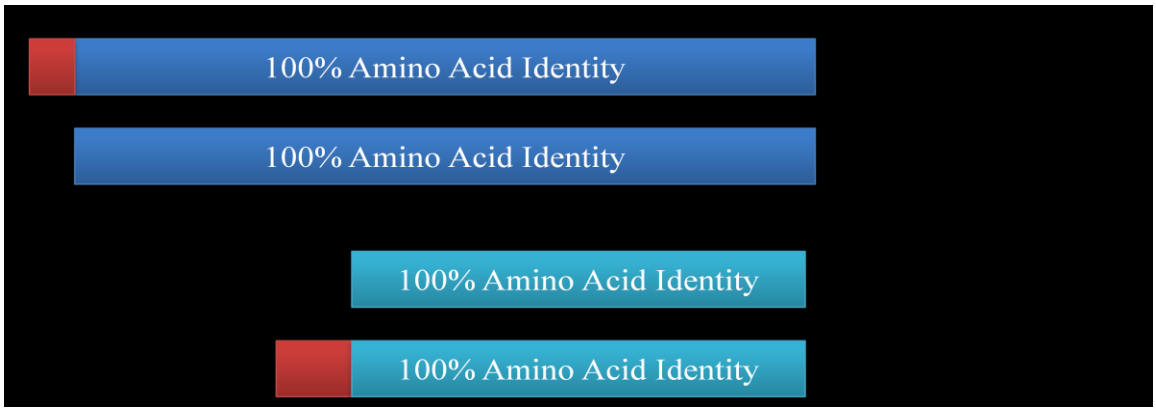


Figure 28. Schematic Representation of the Cry1A.105 and Cry2Ab2 Amino Acid Sequences in MON 87751 and MON 89034

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

Cry1A.105 and Cry2Ab2 protein levels in various tissues of MON 87751 relevant to the risk assessment were determined by a validated enzyme-linked immunosorbent assay (ELISA). Leaves from four growth stages (Over season leaf (OSL) 1 through OSL4), forage, root, and seed tissue samples of MON 87751 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 U.S.: Jackson County, Arkansas (site code ARNE); Jefferson County, Iowa (site code IARL); Pawnee County, Kansas (site code KSLA); Perquimans County, North Carolina (site code NCBD); and Lehigh County, Pennsylvania (site code PAGR). The field sites were representative of soybean producing regions suitable for commercial production. Flowers for the collection of pollen/anther tissue of MON 87751 were also

harvested during the 2012 U.S. growing season from a field site in Champaign County, Illinois (site code ILTH). At this site, tissue was collected from one non-randomized plot.

B3(b)(i) Expression levels of Cry1A.105 protein in MON 87751

Cry1A.105 protein levels were determined in all eight tissue types collected. The ELISA results obtained for each sample were averaged across the five sites, except for pollen/anther where only one site was analyzed because of the limited amount of tissue available, and are summarized in Table 16. Moisture content was not determined for pollen/anther because of a lack of tissue, therefore, pollen/anther results were only reported in fresh weight (fw).

[redacted]

Table 16. Summary of Cry1A.105 Protein Levels in Tissues from MON 87751 Grown in 2012 United States Field Trials

Tissue Type[1]	Development Stage[2]	Mean(SD) Range (µg/g fw)[3]	Mean(SD) Range (µg/g dw)[4]	LOQ/ LOD (µg/g fw)[5]
OSL1	V3 – V4	130 (50) 61 – 220	580 (250) 260 – 1100	1.500/0.406
OSL2	V5 – V7	120 (54) 13 – 220	590 (270) 68 – 1100	1.500/0.406
OSL3	R2 – R3	79 (45) 8.5 – 160	400 (220) 50 – 780	1.500/0.406
OSL4	R6	230 (82) 120 – 480	790 (280) 430 – 1600	1.500/0.406
Root	R6	<LOD (N/A) N/A – N/A	N/A (N/A) N/A – N/A	0.563/0.322
Forage	R6	62 (21) 31 – 110	230 (91) 110 – 440	1.500/0.524
Seed	R8	2.1 (0.46) 1.5 – 2.9	2.4 (0.50) 1.7 – 3.2	0.900/0.226
Pollen/Anther	R2	11 (N/A) N/A – N/A	N/A (N/A) N/A – N/A	1.500/N.D.[6]

[1] OSL= over season leaf
[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 fresh weight basis (fw). The means, SD, and ranges (minimum and maximum values) were calculated for each tissue across all sites (n=20 except for OSL1 where n=19 due to one sample expressing <LOD and pollen/anther where n=1). N/A: Not Applicable
[4] 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 dry weight values were calculated by dividing the µg/g fw by the dry weight conversion factor obtained from moisture analysis data. Moisture content was not determined for pollen/anther due to a lack of tissue, therefore, pollen/anther results were only reported in fresh weight.
[5] LOQ=limit of quantitation; LOD=limit of detection.
[6] N.D. = Not determined. Pollen/anther LOD was not determined due to an insufficient amount of tissue.

B3(b)(ii) Expression levels of Cry2Ab2 protein in MON 87751

Cry2Ab2 protein levels were determined in all eight tissue types collected. The ELISA results obtained for each sample were averaged across the five sites, except for pollen/anther where only one site was analyzed because of the limited amount of tissue available, and are summarized in Table 17. Moisture content was not determined for pollen/anther due to a lack of tissue, therefore, pollen/anther results were only reported in fresh weight.

Table 17. Summary of Cry2Ab2 Protein Levels in Tissues from MON 87751 Grown in 2012 United States Field Trials

Tissue Type[1]	Development Stage[2]	Mean(SD) Range (µg/g fw)[3]	Mean(SD) Range (µg/g dw)[4]	LOQ/LOD (µg/g fw)[5]
OSL1	V3 – V4	5.4 (0.74) 4.4 – 6.8	24 (5.9) 17 – 37	0.625/0.034
OSL2	V5 – V7	5.2 (0.70) 4.0 – 6.6	26 (3.1) 20 – 33	0.625/0.034
OSL3	R2 – R3	6.3 (0.80) 5.2 – 8.0	32 (5.2) 25 – 43	0.625/0.034
OSL4	R6	6.9 (0.79) 5.5 – 8.5	24 (2.7) 18 – 29	0.625/0.034
Root	R6	4.6 (1.0) 3.1 – 7.1	15 (2.7) 11 – 22	1.250/1.241
Forage	R6	3.9 (0.60) 3.0 – 5.1	14 (2.2) 11 – 18	0.313/0.060
Seed	R8	3.6 (0.71) 2.3 – 4.7	4.0 (0.77) 2.6 – 5.1	0.313/0.094
Pollen/Anther	R2	7.7 (N/A) N/A – N/A	N/A (N/A) N/A – N/A	0.313/N.D.[6]

[1] OSL= over season leaf
[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 fresh weight basis (fw). The means, SD, and ranges (minimum and maximum values) were calculated for each tissue across all sites (n=20 except for OSL1 where n=19 due to one sample expressing <LOD and pollen/anther where n=1). N/A: Not Applicable
[4] 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 dry weight values were calculated by dividing the µg/g fw by the dry weight conversion factor obtained from moisture analysis data. Moisture content was not determined for pollen/anther due to a lack of tissue, therefore, pollen/anther results were only reported in fresh weight.
[5] LOQ=limit of quantitation; LOD=limit of detection.
[6] N.D. = Not determined. Pollen/anther LOD was not determined due to an insufficient amount of tissue

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

History of safe use of the introduced proteins is a key consideration in the potential for allergenicity and toxicity and for assessment of dietary safety. The history of safe use of Cry1A.105 and Cry2Ab2 proteins are addressed below.

History of Safe Use of Cry1A.105 Protein

Microbial pesticides that contain Cry1Ac/Cry1F chimeric protein have been used for control of lepidopteran pests since 1997 (Baum, 1998; Baum *et al.*, 1999). Cry1A.105 itself is expressed in MON 89034 maize, which was reviewed and approved by regulatory agencies around the world and has been commercially available in the United States since 2009. The MON 87751 Cry1A.105 protein has 99% amino acid identity with the MON 89034-produced Cry1A.105 protein overall and 100% homology with the protease-resistant core which is responsible for target organism specificity and efficacy. U.S. EPA has approved a tolerance exemption for Cry1A.105 in maize (40 CFR 174.502) (U.S. EPA, 2010). In addition, a biotechnology-derived cotton expressing another chimeric protein consisting of domains or sequences from Cry1F, Cry1C, and Cry1Ab was reviewed and approved by regulatory agencies and has been commercialized (Gao *et al.*, 2006). The large scale cultivation of these crops without any indication of harmful impact on the environment, non-target insects, or mammals provides additional evidence for the safety of the Cry1A.105 protein. Taken together, these data demonstrate that the Cry1A.105 protein has a history of safe use and does not pose any adverse effects to human and animal health.

History of Safe Use of Cry2Ab2 Protein

Cry2Ab2 is derived from *B. thuringiensis* subsp. *kurstaki*. *B. thuringiensis* subsp. *kurstaki* has been used as an active ingredient in many commercial microbial pesticide products such as DiPel® and Cutlass®, and Cry2Ab protein is identified as one of the proteins in Cutlass® (Betz *et al.*, 2000). Cry2Ab2 is also expressed in commercially available MON 89034 maize. The MON 87751-produced Cry2Ab2 protein has 97% amino acid identity with the

® DiPel is a registered trademark of Abbott Laboratories. Cutlass is a registered trademark of Ecogen, Inc.

MON 89034-produced Cry2Ab2 protein overall and 100% homology with the protease-resistant core which is responsible for target organism specificity and efficacy. U.S. EPA has approved a tolerance exemption for Cry2Ab2 as expressed in maize and cotton (40 CFR 174.519) (U.S. EPA, 2010). The large scale cultivation of crops expressing Cry2Ab2 without any indication of harmful impact on the environment, non-target insects, or mammals provides additional evidence for the safety of the Cry2Ab2 protein. Taken together, these data demonstrate that the Cry2Ab2 protein has a history of safe use and does not pose any adverse effects to human and animal health.

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 soybean 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 use (Discussed in Section B3); 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 effects of a toxicological concern in an acute oral mammalian 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. The Cry1A.105 and Cry2Ab2 proteins in MON 87751 have been assessed for their potential toxicity as described below.

B4(a) Bioinformatics 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 the Cry1A.105 and Cry2Ab2 proteins 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 Cry1A.105 and Cry2Ab2 amino acid sequences 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[7], is a subset of sequences derived from the PRT_2013[8] database, that was selected using a keyword search and filtered to remove likely non-toxin proteins. The TOX_2013 database contains 8,881 sequences.

---

[7] The toxin database is a subset of sequences derived from the PRT_2013 database that was selected using a keyword search then filtered to remove likely non-toxin proteins. It is referred to as the TOX_2013 database and contains 8,881 sequences and was used for the evaluation of sequence similarities.

[8] A comprehensive collection of protein sequences that comprises GenBank release 193 was downloaded from NCBI and was used for the evaluation of sequence similarities. It is referred to as the PRT_2013 database and contains 27,998,271 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 Cry1A.105 and Cry2Ab2 proteins. As described, 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 Cry1A.105 and Cry2Ab2 proteins and any sequence in the TOX_2013 database, as no alignments displaying an *E*-score < 1×10^{-5} were observed.

[redacted]

B4(b) Stability to heat or processing and/or degradation in gastric model

B4(b)(i) Digestibility of the Cry1A.105 and Cry2Ab2 proteins

Proteins introduced into commercial food crops using biotechnology are evaluated for their safety for human and animal consumption. One aspect of this assessment is an evaluation of a protein's susceptibility to proteolytic digestion with enzymes of the gastrointestinal tract. Certain proteins can withstand proteolytic digestion by enzymes present in the gastrointestinal tract (Astwood *et al.*, 1996; Vassilopoulou *et al.*, 2006; Moreno *et al.*, 2005; Vieths *et al.*, 1999). A correlation between protein digestibility with pepsin in simulated gastric fluid (SGF) and the likelihood of the protein being an allergen has been previously reported (Astwood *et al.*, 1996), although this correlation is not complete (Fu *et al.*, 2002). The SGF assay serves as a tool to compare the relative susceptibility of proteins to digestion by pepsin. The SGF assay protocol has been standardized based on results obtained from an international, multi-laboratory study (Thomas *et al.*, 2004). This study showed that the standardized protocol provides reproducibility and consistency for determining relative susceptibility of proteins to digestion by pepsin. Using this standardized protocol, the digestive stability of Cry1A.105 and Cry2Ab2 proteins were analyzed.

Simulated intestinal fluid (SIF) has also been used as a stand-alone independent test system to assess the susceptibility of proteins to digestion by proteases characteristic of intestinal fluid (Okunuki *et al.*, 2002; Yagami *et al.*, 2000). The relationship between protein allergenicity and protein stability in the *in vitro* stand-alone SIF study is limited, because the protein has not been first exposed to the acidic, denaturing conditions simulating the stomach, as would be the case with *in vivo* digestion (FAO-WHO, 2001). Nonetheless, the susceptibility of protein to digestion in SIF allows for a better understanding of the likelihood that a protein can withstand proteolytic digestion by enzymes present in the gastrointestinal tract. Therefore the susceptibility,of Cry1A.105 and Cry2Ab2 proteins to digestion in SIF was also assessed.

Digestibility of Cry1A.105 Protein in SGF

The digestibility of *E. coli*-produced Cry1A.105 protein in SGF was assessed using two methods: analysis of a Brilliant Blue G Colloidal stained SDS-PAGE gel and visual analysis of a western blot probed with a goat anti-Cry1A.105 polyclonal antibody. Digestibility of Cry1A.105 protein in SGF was evaluated over time by analyzing digestion mixtures incubated for targeted time intervals. For SDS-PAGE analysis, approximately 1.0 µg of Cry1A.105 protein was analyzed for each time point (Figure 29). The no protein control at digestion time 0 and no protein control at digestion time 60 (Figure 29, Panel A, Lanes 2 and 13), which evaluate the stability of the pepsin in the test system (SGF) lacking the Cry1A.105 protein, demonstrated that the pepsin was observed as a stained protein band at ~38 kDa throughout the experimental phase. In addition, no change in the Cry1A.105 protein band intensity was observed over time in the absence of pepsin (compare no pepsin control at digestion time 0 sec to no pepsin control at digestion time 60 min; Figure 29, Panel A, Lanes 3 and 12) indicating that the digestion of the Cry1A.105 protein was due to the proteolytic activity of pepsin present in SGF and not due to instability of the protein while incubated at pH ~1.2 at ~37 °C for 60 min.

SDS-PAGE data showed that the intact Cry1A.105 protein was completely digested within 0.5 minutes of incubation in SGF (Figure 29, Panel A, Lane 5). Colloidal Brilliant Blue G stained SDS-PAGE analysis showed that a transient peptide fragment of ~60 kDa was present at 0.5 min but gone by 2 min (Figure 29, Panel A, Lanes 5 and 6), and a transiently-stable fragment under 5 kDa was observed after 20 minutes but was gone by 30 min in the SGF digestion (Figure 29 Panel A, Lanes 9-10). For the SDS-PAGE analysis, the limit of detection (LOD) of the Cry1A.105 protein was visually estimated to be 13 ng, or 0.013 µg (Figure 29, Panel B, Lane 6). This LOD was used to calculate the maximum amount of intact Cry1A.105 protein that could remain visually undetected after digestion, which corresponded to approximately 1.3% of the Cry1A.105 protein loaded. Based on that LOD, more than 98% (100.0% - 1.3% = 98.7%) of the intact Cry1A.105 protein was digested within 0.5 min of incubation in SGF.

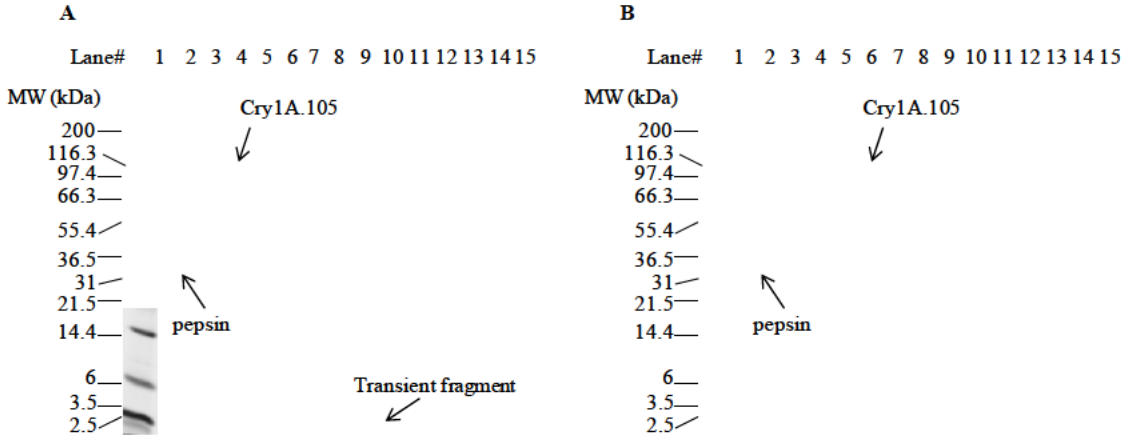


Figure 29. Colloidal Blue Stained SDS-PAGE Gel Showing the Digestion of Purified *E. coli*-Produced Cry1A.105 Protein in Simulated Gastric Fluid

Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded. In each gel, Cry1A.105 protein migrated to approximately 130 kDa and pepsin to approximately 38 kDa.

A: Cry1A.105 protein digestion in SGF. Based on pre-digestion protein concentrations, 1 µg of test protein was loaded in each lane containing Cry1A.105 protein.

B: LOD determination. Indicated amounts of the test protein from the SGF T0 sample were loaded to estimate the LOD of the Cry1A.105 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	Blank	-
15	Mark 12 MWM	-

Lane	Sample	Amount (ng)
1	Mark 12 MWM	-
2	SGF T0	208.7
3	SGF T0	104.4
4	SGF T0	52.2
5	SGF T0	26.1
6	SGF T0	13.0
7	SGF T0	6.5
8	SGF T0	3.3
9	SGF T0	1.6
10	SGF T0	0.8
11	Blank	-
12	Blank	-
13	Blank	-
14	Blank	-
15	Mark 12 MWM	-

The digestibility of the Cry1A.105 protein in SGF was also evaluated using western blot analysis. The Cry1A.105 protein was loaded at 20 ng per lane (based on pre-digestion concentration and purity values). Examination of the SGF western blot (Figure 30, Panel A) confirmed the results of the Colloidal Brilliant Blue G stained polyacrylamide gel. The full-length Cry1A.105 protein (~130 kDa) was digested below the LOD within 0.5 minutes of incubation in SGF (Figure 30, Panel A, Lane 5). No proteolytic fragments were observed for any of the time points. No bands were observed when Cry1A.105 protein was absent (Figure 30, Panel A, Lanes 2 and 13), demonstrating that the immunoreactive signal observed on this blot was due solely to Cry1A.105 protein. No change in the Cry1A.105 protein band intensity was observed in the no pepsin controls (Figure 30, Panel A, Lanes 3 and 12). This result reaffirms that the Cry1A.105 protein was stable in the assay system without pepsin and indicates that degradation of the Cry1A.105 protein was due to digestion by pepsin.

The LOD of the Cry1A.105 protein in the western blot analysis was estimated to be 0.31 ng (Figure 30, Panel B, Lane 8). This LOD was used to calculate the maximum amount of intact Cry1A.105 protein that could remain visually undetected after digestion. This corresponded to approximately 1.6 % of the total Cry1A.105 protein loaded. Based on the western blot LOD, it can be concluded that within 0.5 min more than 98% (100% - 1.6% = 98.4%) of the intact Cry1A.105 protein was digested.

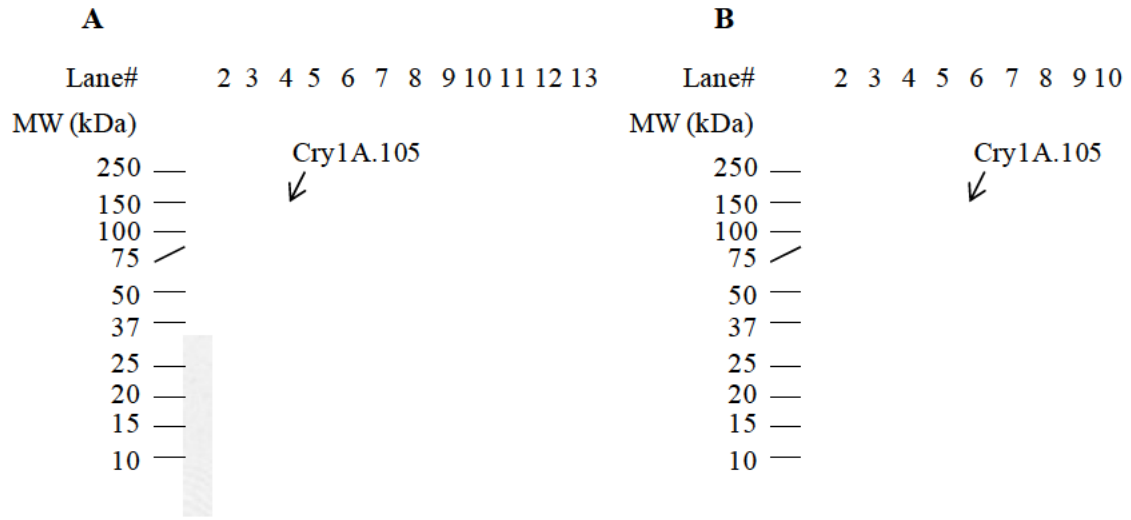


Figure 30. Western Blot Analysis of Purified *E. coli*-Produced Cry1A.105 Protein in Simulated Gastric Fluid

Western blots probed with an anti-Cry1A.105 antibody were used to assess the digestibility of Cry1A.105 in SGF. Molecular weights (kDa) are shown on the left of each blot, and correspond to the markers loaded (cropped from images). On each blot, Cry1A.105 protein migrated to approximately 130 kDa. Blank lanes were cropped from the images. A 30 sec exposure is shown.

A: Cry1A.105 protein digestion in SGF. Based on pre-digestion protein concentrations, 20 ng of test protein was loaded in each lane containing Cry1A.105 protein.

B: LOD determination. Indicated amounts of the test protein from the SGF T0 sample were loaded to estimate the LOD of the Cry1A.105 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	Blank	-
15	Blank	-

Lane	Sample	Amount (ng)
1	Precision Plus MWM	-
2	SGF T0	20
3	SGF T0	10
4	SGF T0	5
5	SGF T0	2.5
6	SGF T0	1.25
7	SGF T0	0.63
8	SGF T0	0.31
9	SGF T0	0.16
10	SGF T0	0.08
11	Blank	-
12	Blank	-
13	Blank	-
14	Blank	-
15	Blank	-

Digestive Fate of the Cry1A.105 Protein in SGF Followed by SIF

To better understand the fate of an observed <5 kDa transiently stable Cry1A.105 protein fragment during gastrointestinal digestion, sequential digestibility of the Cry1A.105 protein was conducted. This sequential digestibility was assessed both by visual analysis of a Colloidal Brilliant Blue G stained SDS-PAGE gel, and visual analysis of a western blot probed with an anti- Cry1A.105 polyclonal antibody.

For the sequential digestibility assay, the Cry1A.105 protein was incubated with SGF for 2 min, followed by incubation with SIF. For the Colloidal Brilliant Blue G stained SDS-PAGE assessment, lanes were loaded with 1 µg of Cry1A.105 test protein (based on pre-digestion concentrations) per incubation time point. Examination of SDS-PAGE data showed that the intact Cry1A.105 protein was digested within 2 min of incubation in SGF (Figure 31, Panel A, Lane 3) and the small transient fragment of ~5 kDa was completely digested within 0.5 min of SIF exposure (Figure 31, Panel A, Lane 7). No change in the ~5 kDa fragment band intensity was observed in the absence of pancreatin in the SEQ 0 min No Pancreatin Control and SEQ 2 hr No Pancreatin Control (Figure 31, Panel A, Lanes 5 and 14). This indicates that the digestion of the ~5 kDa fragment was due to the proteolytic activity of pancreatin present in SIF and not due to instability of the fragment while incubated at pH~7.5 at 37 ± 2ºC for 2 hr.

The SEQ 0 min No Test Protein Control and SEQ 2 hr No Test Protein Control (Figure 31, Panel A, Lanes 4 and 15) demonstrated the integrity of the SIF over the course of the experiment. The intensity of some pancreatin bands decreased somewhat during the course of the experiment, most likely due to autodigestion. This is not expected to adversely impact the SIF results, as the ~5 kDa transiently stable fragment was digested within 0.5 min of exposure to SIF.

The sequential digestion of Cry1A.105 protein was also assessed by western blot (Figure 31, Panel B), with 20 ng of the test protein (based on total protein pre-digestion concentrations) loaded per lane. No bands were detected after the SGF 0 min time point (Figure 31, Panel B, Lane 2).

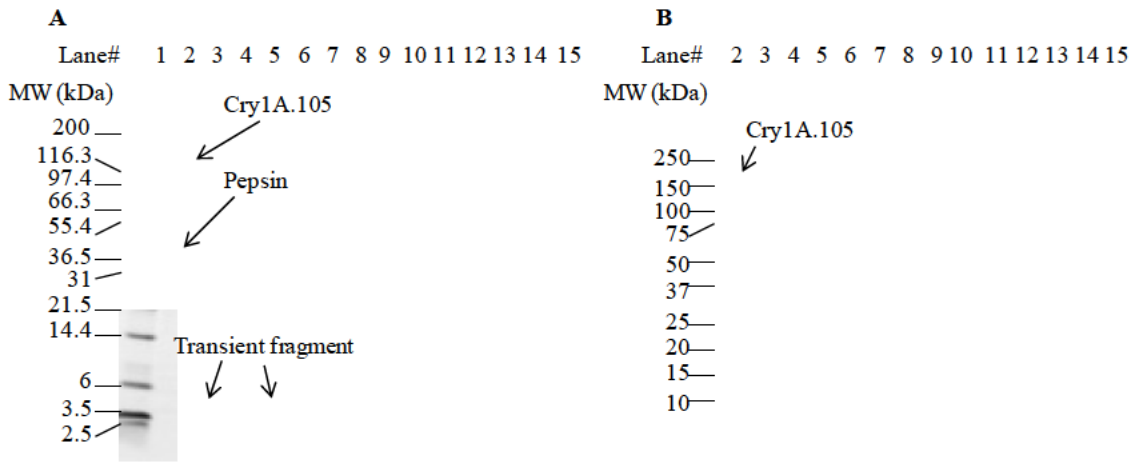


Figure 31. SDS-PAGE Gel and Western Blot Analysis Showing the Digestion of Purified *E. coli*-Produced Cry1A.105 Protein in Simulated Gastric Fluid Followed by Simulated Intestinal Fluid

SDS-PAGE and western blot analysis were used to assess the digestibility of Cry1A.105 in sequential digestion. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded (cropped from image B).

A: Colloidal Brilliant Blue G stained SDS-PAGE gel analysis of Cry1A.105 in sequential digestion. Based on pre-digestion protein concentrations, 1 µg of test protein was loaded in each lane containing Cry1A.105 protein.

B: Western blot analysis of Cry1A.105 in sequential digestion. Based on pre-digestion protein concentrations, 20 ng of test protein was loaded in each lane containing Cry1A.105 protein. A 30 sec exposure is shown.

Lane	Sample	Incubation Time
1	Mark 12 MWM	-
2	SGF 0 min	0
3	SGF 2 min	2
4	SEQ 0 min No Protein Control	0
5	SEQ 0 min No Pancreatin Control	0
6	SEQ T0	0
7	SEQ T1	0.5 min
8	SEQ T2	2 min
9	SEQ T3	5 min
10	SEQ T4	10 min
11	SEQ T5	30 min
12	SEQ T6	1 hr
13	SEQ T7	2 hr
14	SEQ 2 hr No Pancreatin Control	2 hr
15	SEQ 2 hr No Protein Control	2 hr

Lane	Sample	Incubation Time
1	Precision Plus MWM	-
2	SGF 0 min	0
3	SGF 2 min	2
4	SEQ 0 min No Protein Control	0
5	SEQ 0 min No Pancreatin Control	0
6	SEQ T0	0
7	SEQ T1	0.5 min
8	SEQ T2	2 min
9	SEQ T3	5 min
10	SEQ T4	10 min
11	SEQ T5	30 min
12	SEQ T6	1 hr
13	SEQ T7	2 hr
14	SEQ 2 hr No Pancreatin Control	2 hr
15	SEQ 2 hr No Protein Control	2 hr

Digestibility of Cry1A.105 in SIF

Pancreatin contains trypsin and other proteases, which are known to digest Cry proteins. Proteolytic conversion of the full-length protoxin to an active toxin has been described for all members of the Cryl, Cry2, Cry3, and Cry4 class proteins (Rukmini *et al.*, 2000). When exposed to trypsin or trypsin-like insect midgut proteases, Cry proteins are degraded to a stable "tryptic core" (Höfte and Whiteley, 1989). The digestibility of the *E. coli*-produced Cry1A.105 protein in SIF was assessed by western blot (Figure 32). The gel used to assess the digestibility of the Cry1A.105 protein in SIF by western blot was loaded with 20 ng of the Cry1A.105 protein (based on pre-digestion concentration and purity values) for each of the incubation time points. Western blot analysis indicated that the full-length Cry1A.105 protein (~130 kDa) was digested to a level below the LOD within 5 min in SIF (Figure 32, Panel A, Lane 5). As expected, an immunoreactive band corresponding to the size of the trypsin-resistant core (~55 kDa) was observed throughout the course of the SIF digestion (Schnepf *et al.*, 1998).

No bands were observed when Cry1A.105 protein was absent (Figure 32, Panel A, Lanes 2 and 14), demonstrating that the immunoreactive signal observed on this blot was due solely to Cry1A.105 protein. The no pancreatin controls that contained the Cry1A.105 protein but lacked pancreatin were loaded in Lanes 3 and 13. No change in the Cry1A.105 band intensities was observed when exposed to SIF without pancreatin for 24 hours. This demonstrated that degradation of the Cry1A.105 protein was due to digestion by pancreatin and not instability of the test substance at pH 7.5 and incubation at 37 °C.

To determine the LOD for Cry1A.105 protein in SIF by western blot analysis, dilutions of the SIF0 sample (loaded based on pre-digestion concentration and purity values) were analyzed on the same western blot (Figure 32, Panel B, Lanes 2-9). The lowest amount of Cry1A.105 protein observed on the western blot was 0.31 ng (Figure 32, Panel B, Lane 8). Therefore the LOD for Cry1A.105 protein was established to be 0.31 ng. Based on the LOD, it was determined that the full-length protein was at least 98.4% digested within 5 minute of incubation in SIF.

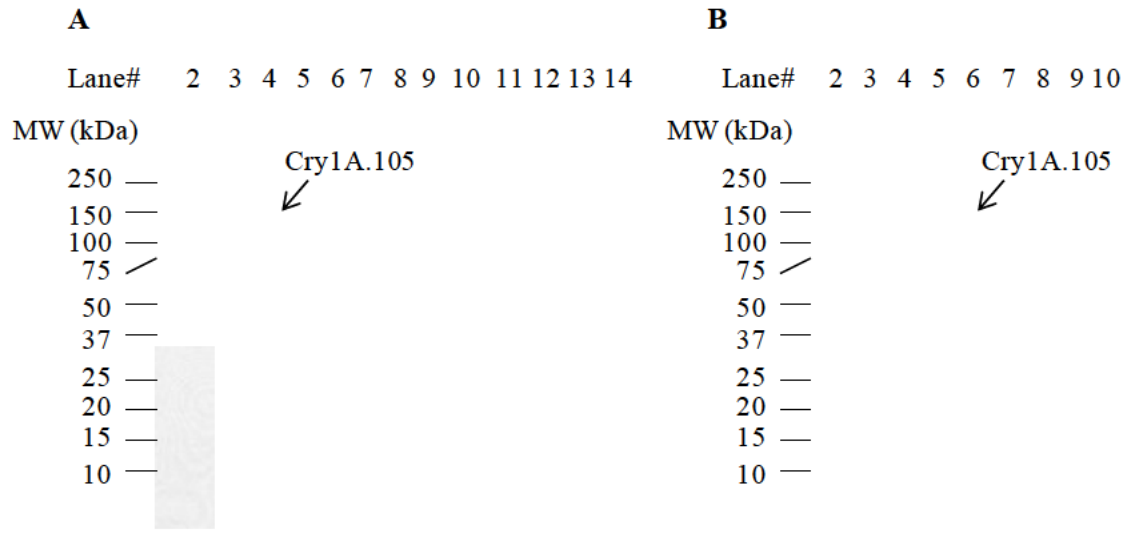


Figure 32. Western Blot Analysis of Purified *E. coli*-Produced Cry1A.105 Protein in Simulated Intestinal Fluid

Western blots were used to assess the digestibility of Cry1A.105 in SIF. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded (cropped from images). On each blot, Cry1A.105 protein migrated to approximately 130 kDa. Blank lanes were cropped from the images. A 30 sec exposure is shown.
A: Cry1A.105 protein digestion in SIF. Based on pre-digestion protein concentrations, 20 ng of test protein was loaded in each lane containing Cry1A.105 protein.
B: LOD determination. Indicated amounts of the test protein from the SIF T0 sample were loaded to estimate the LOD of the Cry1A.105 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	Blank	-

Lane	Sample	Amount (ng)
1	Precision Plus MWM	-
2	SIF T0	20
3	SIF T0	10
4	SIF T0	5
5	SIF T0	2.5
6	SIF T0	1.25
7	SIF T0	0.63
8	SIF T0	0.31
9	SIF T0	0.16
10	SIF T0	0.08
11	Blank	-
12	Blank	-
13	Blank	-
14	Blank	-
15	Blank	-

Digestibility of Cry1A.105 protein – Conclusions

Digestibility of the Cry1A.105 protein was evaluated in SGF, sequential digestion, and SIF. The results of the study demonstrate that greater than 98.7% of the full-length Cry1A.105 protein was digested in SGF within 0.5 min when analyzed by Colloidal Brilliant Blue G stained SDS-PAGE and at least 98.4% of the full-length Cry1A.105 protein was digested within 0.5 min when analyzed by western blot using a Cry1A.105 specific antibody. A transiently-stable protein fragment migrating at ~5 kDa was observed after 20 minutes of digestion in SGF but was gone by 30 min when analyzed by a Colloidal Brilliant Blue G stained polyacrylamide gel; neither fragment nor any other immunoreactive peptides were detected using western blot analysis. When the Cry1A.105 protein was subjected to a sequential digestion, *i.e.* brief digestion in SGF followed by digestion in SIF, the fragment that was transiently stable during SGF digestion was digested within 0.5 min by SIF. At least 98.4% of the full-length Cry1A.105 protein was digested within 5 min during incubation in SIF when analyzed by western blot. As expected, the trypsin-resistant core (~55 kDa) was observed throughout the course of the SIF stand-alone digestion. Rapid digestion of the full-length Cry1A.105 protein in SGF and SIF, together with rapid digestion in sequential digestion of the transiently stable ~5 kDa fragment indicates that the Cry1A.105 protein and its fragment are highly unlikely to pose any safety concern to human or animal health.

[redacted]

Digestibility of Cry2Ab2 Protein in SGF

The digestibility of *E. coli*-produced Cry2Ab2 protein in SGF was assessed using two methods: analysis of a Brilliant Blue G Colloidal stained SDS-PAGE gel and visual analysis of a western blot probed with a goat anti-Cry2Ab2 polyclonal antibody.

Digestibility of Cry2Ab2 protein in SGF was evaluated over time by analyzing digestion mixtures incubated for targeted time intervals. For SDS-PAGE analysis, approximately 1.0 µg of Cry2Ab2 protein was analyzed for each time point (Figure 33, Panel A). The no protein control at digestion time 0 and no protein control at digestion time 60 (Figure 33, Panel A, Lanes 2 and 13), which evaluate the stability of the pepsin in the test system (SGF) lacking the Cry2Ab2 protein, demonstrated that the pepsin was observed as a stained protein band at ~38 kDa throughout the experimental phase. In addition, no change in the Cry2Ab2 protein band intensity was observed over time in the absence of pepsin (compare no pepsin control at digestion time 0 sec to no pepsin control at digestion time 60 min; Figure 33, Panel A, Lanes 3 and 12) indicating that the digestion of the Cry2Ab2 protein was due to the proteolytic activity of pepsin present in SGF and not due to instability of the protein while incubated at pH ~1.2 at ~37 °C for 60 min.

SDS-PAGE data showed that the intact Cry2Ab2 protein was completely digested within 0.5 min of incubation in SGF (Figure 33, Panel A, Lane 5). A protein fragment (~4-5 kDa) was observed at 0.5 min of incubation, but was completely digested to a level below the limit of detection after two minutes in SGF (Figure 33, Panel A, Lanes 7-11). For the SDS-PAGE analysis, the LOD of the Cry2Ab2 protein was visually estimated to be 6.3 ng, or 0.0063 µg (Figure 33, Panel B, Lane 6). This LOD used to calculate the maximum amount of intact

Cry2Ab2 protein that could remain visually undetected after digestion, which corresponded to approximately 0.6% of the Cry2Ab2 protein loaded. Based on that LOD, more than 99.0% (100.0% - 0.6% = 99.4%) of the intact Cry2Ab2 protein was digested within 0.5 min of incubation in SGF.

The digestibility of the Cry2Ab2 protein in SGF was also evaluated using western blot analysis. The Cry2Ab2 protein was loaded at 20 ng per lane (based on pre-digestion concentration and purity values). Examination of the SGF western blot (Figure 34) confirmed the results of the Colloidal Brilliant Blue G stained polyacrylamide gel. The full-length Cry2Ab2 protein (~60 kDa) was digested below the LOD within 0.5 min of incubation in SGF (Figure 34, Panel A, Lane 5). No proteolytic fragments were observed for any of the time points. No bands were observed in the no protein controls when Cry2Ab2 protein was absent (Figure 34, Panel A, Lanes 2 and 13, respectively), demonstrating that the immunoreactive signal observed on this blot was due solely to Cry2Ab2 protein. No change in the Cry2Ab2 protein band intensity was observed in the no pepsin controls (Figure 34, Panel A, Lanes 3 and 12). This result reaffirms that the Cry2Ab2 protein was stable in the assay system without pepsin and indicates that degradation of the Cry2Ab2 protein was due to digestion by pepsin.

The LOD of the Cry2Ab2 protein in the western blot analysis was estimated to be 0.31 ng (Figure 34, Panel B, Lane 8). The LOD estimated was used to calculate the maximum amount of intact Cry2Ab2 protein that could remain visually undetected after digestion. This corresponded to approximately 1.6% of the total Cry2Ab2 protein loaded. Based on the western blot LOD, it can be concluded that more than 98% (100% - 1.6% = 98.4%) of the intact Cry2Ab2 protein was digested within 0.5 min.

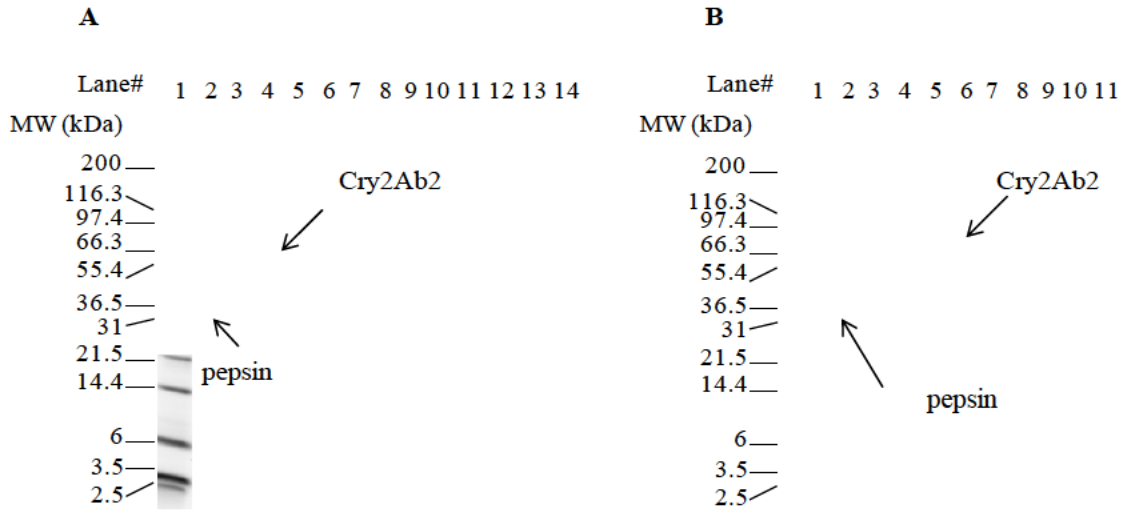


Figure 33. Colloidal Blue Stained SDS-PAGE Gel Showing the Digestion of Purified *E. coli*-Produced Cry2Ab2 Protein in Simulated Gastric Fluid

Colloidal Brilliant Blue G stained SDS-PAGE gels were used to assess the digestibility of Cry2Ab2 in SGF. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded. In each gel, Cry2Ab2 protein migrated to approximately 60 kDa and pepsin to approximately 38 kDa. Blank lanes were cropped from the images.
A: Cry2Ab2 protein digestion in SGF. Based on pre-digestion protein concentrations, 1 µg of test protein was loaded in each lane containing Cry2Ab2 protein.
B: LOD determination. Indicated amounts of the test protein from the SGF T0 sample were loaded to estimate the LOD of the Cry2Ab2 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	-
15	Blank	-

Lane	Sample	Amount (ng)
1	Mark 12 MWM	-
2	SGF T0	101.0
3	SGF T0	50.5
4	SGF T0	25.3
5	SGF T0	12.6
6	SGF T0	6.3
7	SGF T0	3.2
8	SGF T0	1.6
9	SGF T0	0.8
10	SGF T0	0.4
11	Mark 12 MWM	-
12	Blank	-
13	Blank	-
14	Blank	-
15	Blank	-

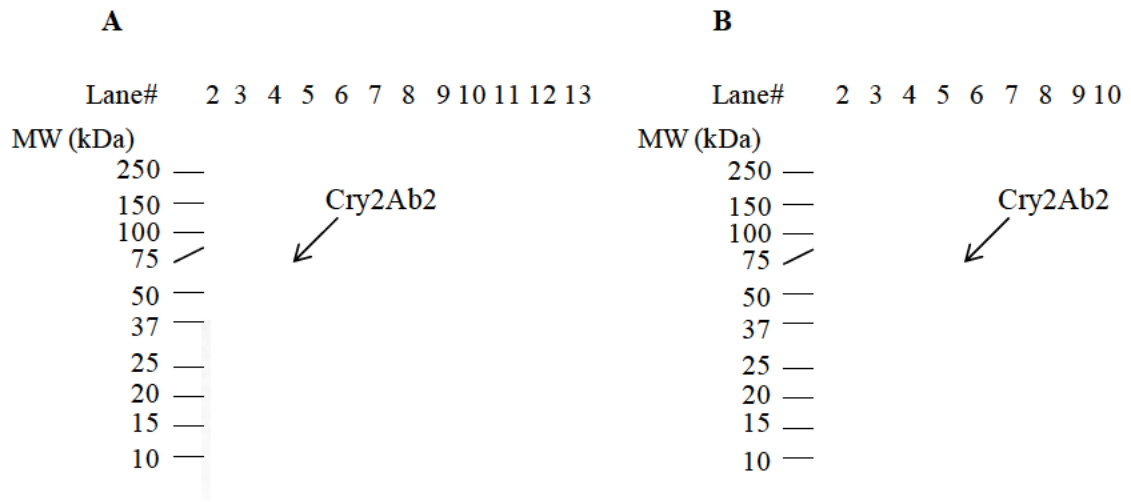


Figure 34. Western Blot Analysis of Purified *E. coli*-Produced Cry2Ab2 Protein in Simulated Gastric Fluid

Western blots probed with an anti-Cry2Ab2 antibody were used to assess the digestibility of Cry2Ab2 in SGF. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded (cropped from images). In each gel, Cry2Ab2 protein migrated to approximately 60 kDa. Blank lanes were cropped from the images. A 1 min exposure is shown.

A: Cry2Ab2 protein digestion in SGF. Based on pre-digestion protein concentrations, 20 ng of test protein was loaded in each lane containing Cry2Ab2 protein.

B: LOD determination. Indicated amounts of the test protein from the SGF T0 sample were loaded to estimate the LOD of the Cry2Ab2 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	Blank	-
15	Blank	-

Lane	Sample	Amount (ng)
1	Precision Plus MWM	-
2	SGF T0	19.9
3	SGF T0	10
4	SGF T0	5
5	SGF T0	2.5
6	SGF T0	1.25
7	SGF T0	0.62
8	SGF T0	0.31
9	SGF T0	0.16
10	SGF T0	0.08
11	Blank	-
12	Blank	-
13	Blank	-
14	Blank	-
15	Blank	-

Digestibility of Cry2Ab2 in SIF

Pancreatin contains trypsin and other proteases, which are known to digest Cry proteins. Proteolytic conversion of the full-length protoxin to an active toxin has been described for all members of the Cryl, Cry2, Cry3 and Cry4 class proteins (Rukmini *et al.*, 2000). When exposed to trypsin or trypsin-like insect midgut proteases, Cry proteins are degraded to a stable "tryptic core" (Höfte and Whiteley, 1989). The digestibility of the *E. coli*-produced Cry2Ab2 protein in SIF was assessed by western blot (Figure 35). The gel used to assess the digestibility of the Cry2Ab2 protein in SIF by western blot was loaded with 20 ng of the Cry2Ab2 protein (based on pre-digestion concentration and purity values) for each of the incubation time points.

Western blot analysis indicated that the full-length Cry2Ab2 protein (~60 kDa) was digested within 5 minute in SIF (Figure 35, Panel A, Lane 5). As expected, the trypsin-resistant core (~50 kDa) was observed throughout the 24 hour incubation in SIF. In addition, an immunoreactive band of ~15 kDa was present after 5 min of SIF digestion but was gone by 15 min (Figure 35, Panel A, Lanes 5 and 6). Immunoreactive bands of ~57 kDa and of just under 10 kDa were transiently present during SIF digestion but both were gone by 4 hr. An immunoreactive band of ~180 kDa was observed intermittently throughout SIF digestion (most obviously at 5 and 15 min). This ~180 kDa band was likely due to aggregation of the intact Cry2Ab2 and/or the tryptic core under the conditions necessary to conduct the SIF digestion assay.

No bands were observed in the no test protein controls when Cry2Ab2 protein was absent (Figure 35, Panel A, Lanes 2 and 14, respectively), demonstrating that the immunoreactive signal observed on this blot was due solely to Cry2Ab2 protein. The no pancreatin controls that contained the Cry2Ab2 protein but lacked pancreatin were loaded in Lanes 3 and 13. No change in the Cry2Ab2 band intensities was observed when exposed to SIF without pancreatin for 24 hours. This demonstrated that degradation of the Cry2Ab2 protein was due to digestion by pancreatin and not instability of the test substance at pH 7.5 and incubation at 37 °C.

To determine the LOD for Cry2Ab2 protein in SIF by western blot analysis, dilutions of the SIF0 sample (loaded based on pre-digestion concentration and purity values) were analyzed on the same western blot (Figure 35, Panel B , Lanes 2-10). The lowest amount of Cry2Ab2 protein observed on the western blot was 0.63 ng (Figure 35, Lane 7). Therefore the LOD for Cry2Ab2 protein was established to be 0.63 ng. Based on the LOD, it was determined that the full-length protein was at least 96.8 % digested within 5 minute of incubation in SIF.

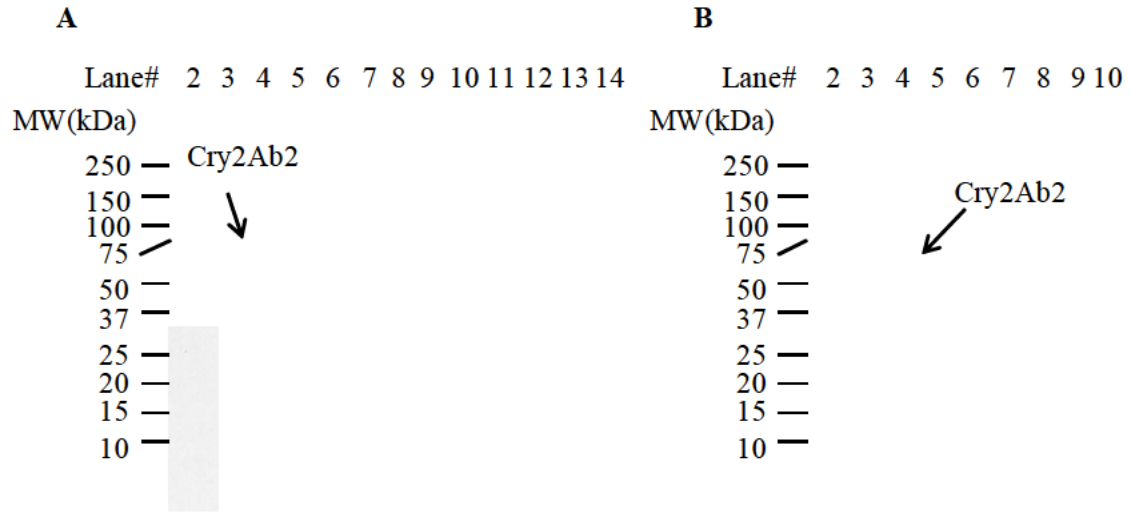


Figure 35. Western Blot Analysis of Purified *E. coli*-Produced Cry2Ab2 Protein in Simulated Intestinal Fluid

Western blots probed with an anti-Cry2Ab2 antibody were used to assess the digestibility of Cry2Ab2 in SIF. Molecular weights (kDa) are shown on the left of each gel, and correspond to the markers loaded (cropped from images). In each gel, intact Cry2Ab2 protein migrated to approximately 60 kDa. Blank lanes were cropped from the images. A 0.5 min exposure is shown.

A: Cry2Ab2 protein digestion in SIF. Based on pre-digestion protein concentrations, 20 ng of test protein was loaded in each lane containing Cry2Ab2 protein.

B: LOD determination. Indicated amounts of the test protein from the SIF T0 sample were loaded to estimate the LOD of the Cry2Ab2 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	Blank	-

Lane	Sample	Amount (ng)
1	Precision Plus MWM	-
2	SIF T0	20
3	SIF T0	10
4	SIF T0	5
5	SIF T0	2.5
6	SIF T0	1.25
7	SIF T0	0.63
8	SIF T0	0.31
9	SIF T0	0.16
10	SIF T0	0.08
11	Blank	-
12	Blank	-
13	Blank	-
14	Blank	-
15	Blank	-

Digestibility of Cry2Ab2 protein - Conclusions

Experiments designed to test the digestibility of the Cry2Ab2 protein in simulated digestive fluids were performed. Results indicate that the Cry2Ab2 protein was rapidly digested when incubated in SGF. At least 99% of the full length (~60 kDa) Cry2Ab2 protein was digested within 0.5 min in SGF when analyzed using Colloidal Brilliant Blue G stained SDS polyacrylamide gels. A proteolytic fragment (~4-5 kDa) was observed in the initial time points, but was absent after two minutes of incubation in SGF. At least 98% of the full length Cry2Ab2 protein was digested within 0.5 min in SGF when evaluated using western blot analysis. No proteolytic fragments were observed for samples evaluated using western blot analysis. The full length Cry2Ab2 protein was also rapidly digested when incubated in SIF. At least 96% of the full length Cry2Ab2 protein was digested within 5 minute in SIF when evaluated using western blot analysis. As expected, a protease resistant fragment was observed throughout the 24 hour time course of the SIF digestion, since proteolytic (trypsin) conversion of the full length protoxin to an active toxin has been described for all members of the Cryl, Cry2, Cry3, and Cry4 families. Taken together, these results suggest that Cry2Ab2 protein will be readily digestible in the mammalian digestive tract.

[redacted]

B4(b)(ii) Heat stability of Cry1A.105 and Cry2Ab2 proteins

Heat Stability of Purified Cry1A.105 Protein

Heat treatment is widely used in soybean processing and in the preparation of foods containing components derived from soybean (Hammond and Jez, 2011). The effect of heat treatment on the Cry1A.105 and Cry2Ab2 proteins was evaluated using a functional assay to assess the impact of temperature on activity and using SDS-PAGE to assess the impact of temperature on protein integrity.

The effect of heating at various temperatures on the functional activity of Cry1A.105 protein for 15 min and 30 min is presented in Table 18 and Table 19, respectively. Treatment of Cry1A.105 for 15 min and 30 min at 75 ºC or above resulted in ≥ 97% loss of detectable functional activity. SDS-PAGE analysis demonstrated that heating at 25, 37, 55, 75 and 95 ºC for 15 minutes (Figure 36) or 25, 37, 55 and 75 ºC for 30 minutes (Figure 37) did not change the apparent intensity of the full length Cry1A.105 protein (~130 kDa). When incubated at 95 ºC for 30 minutes, a slight reduction in band intensity of ~130 kDa was observed (Figure 37, Lane 6). In addition, the appearance of several faint bands that are higher molecular weight than the 130 kDa Cry1A.105 were observed at heat treatments of 75 and 95 ºC, these are most likely heat-denatured protein aggregates. These data demonstrate that the Cry1A.105 protein behaves with a predictable tendency toward loss of functional activity at elevated temperatures.

[redacted]

Table 18. Activity of Cry1A.105 Protein after 15 Minutes at Elevated Temperatures

Temperature	EC_{50}[1] (µg Cry1A.105/ml diet)	Relative Activity (% of control substance)[3,4]
0 ºC (control substance)	0.0043	100%
25 ºC	0.0033	130.3%
37 ºC	0.0040	107.5%
55 ºC	0.0041	104.9%
75 ºC	>0.20[2]	<2.2%
95 ºC	>0.20[2]	<2.2%

[1]Determined from concentration response curves that included eight Cry1A.105 concentrations ranging from 0.0016 – 0.20 µg Cry1A.105 protein/ml diet. Each Cry1A.105 concentration contained one replicate with a target number of sixteen Corn earworm (CEW)larvae.
[2]0.20 µg Cry1A.105/ml diet represents the highest concentration tested.
[3]The control sample (not heated) was assigned as 100 % activity, and others are rounded to one decimal number.
[4]Relative Activity = (EC_{50} of Control substance µg /ml diet ÷ EC_{50} of heat treated sample µg /ml diet) × 100%

Table 19. Activity of Cry1A.105 Protein after 30 Minutes at Elevated Temperatures

Temperature	EC_{50}[1] (µg Cry1A.105/ml diet)	Relative Activity (% of control substance)[3,]
0 ºC (control substance)	0.0043	100%
25 ºC	0.0037	116.2%
37 ºC	0.0065	67.2%
55 ºC	0.0058	74.1%
75 ºC	>0.20[2]	<2.2%
95 ºC	>0.20[2]	<2.2%

[1]Determined from concentration response curves that included eight Cry1A.105 concentrations ranging from 0.0016 – 0.20 µg Cry1A.105 protein/ml diet. Each Cry1A.105 concentration contained one replicate with a target number of sixteen Corn earworm (CEW)larvae.
[2]0.20 µg Cry1A.105/ml diet represents the highest concentration tested.
[3]The control sample (not heated) was assigned as 100 % activity, and others are rounded to one decimal number.
[4]Relative Activity = (EC_{50} of Control substance µg /ml diet ÷ EC_{50} of heat treated sample µg /ml diet) × 100%.

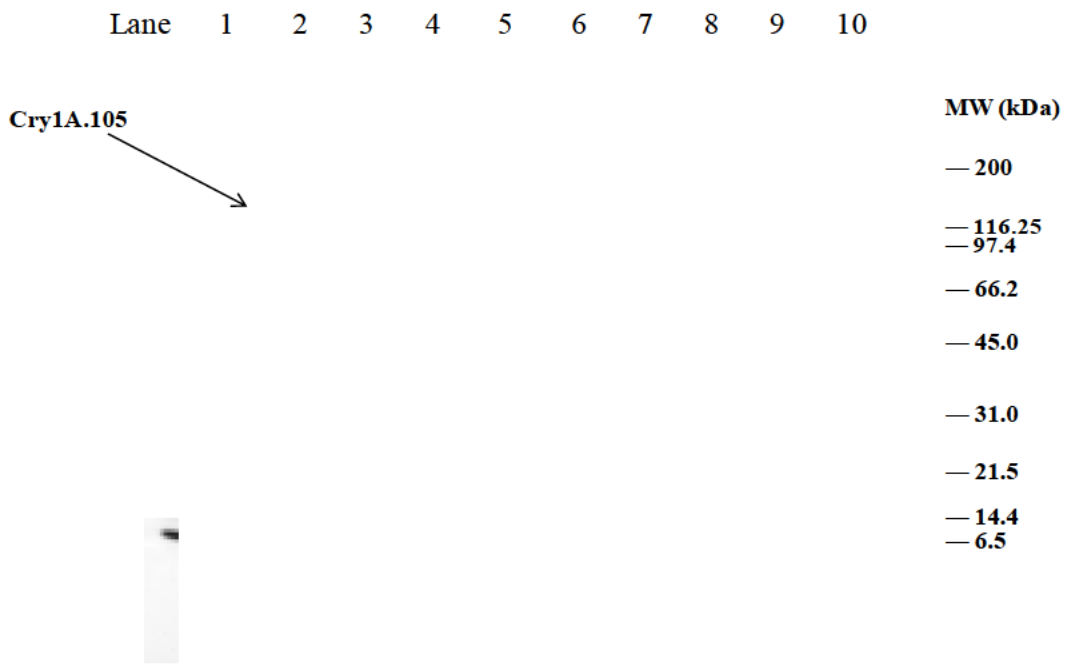


Figure 36. SDS-PAGE of Cry1A.105 Protein Follwing Heat Treatment for 15 Minutes

Heat-treated samples of Cry1A.105 (3.0 µg total protein) were subjected to SDS-PAGE using a Tris-glycine 4-20 % polyacrylamide gel 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	Description	Total Amount
1	Broad Range Molecular Weight Standards	4.5 µg
2	Cry1A.105 Protein 25 ºC for 15 min	3.0 µg
3	Cry1A.105 Protein 37 ºC for 15 min	3.0 µg
4	Cry1A.105 Protein 55 ºC for 15 min	3.0 µg
5	Cry1A.105 Protein 75 ºC for 15 min	3.0 µg
6	Cry1A.105 Protein 95 ºC for 15 min	3.0 µg
7	Cry1A.105 Protein Control Substance	3.0 µg
8	Cry1A.105 Protein Reference 100 % Equivalence	3.0 µg
9	Cry1A.105 Protein Reference 10 % Equivalence	0.3 µg
10	Broad Range Molecular Weight Standards	4.5 µg

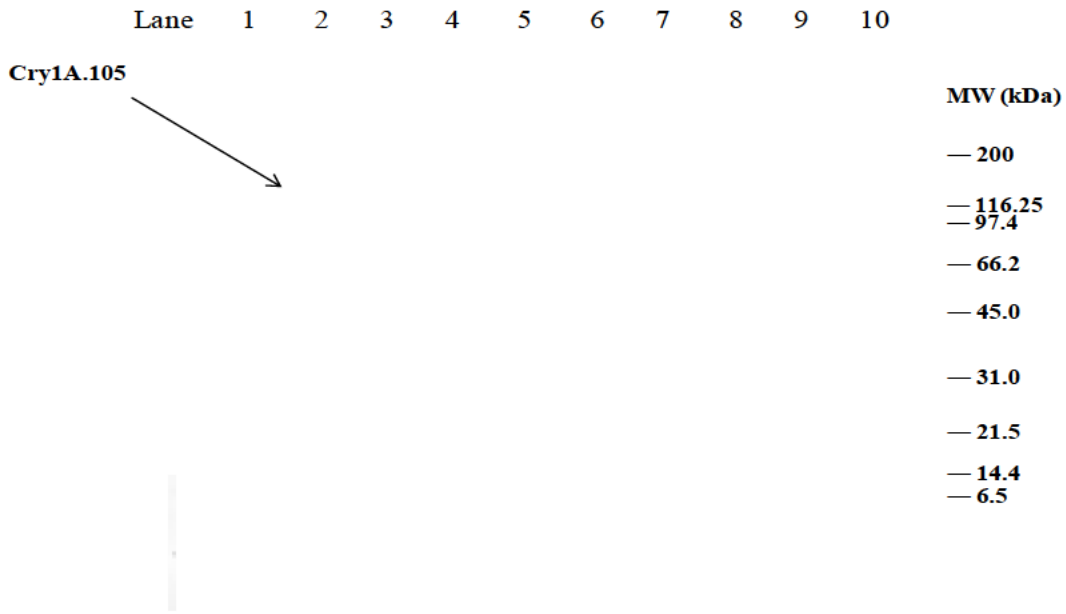


Figure 37. SDS-PAGE of Cry1A.105 Protein Following Heat Treatement for 30 Minutes

Heat-treated samples of Cry1A.105 (3.0 µg total protein) were subjected to SDS-PAGE using a Tris-glycine 4-20 % polyacrylamide gel 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	Description	Total Amount
1	Broad Range Molecular Weight Standards	4.5 µg
2	Cry1A.105 Protein 25 ºC for 30 minutes	3.0 µg
3	Cry1A.105 Protein 37 ºC for 30 minutes	3.0 µg
4	Cry1A.105 Protein 55 ºC for 30 minutes	3.0 µg
5	Cry1A.105 Protein 75 ºC for 30 minutes	3.0 µg
6	Cry1A.105 Protein 95 ºC for 30 minutes	3.0 µg
7	Cry1A.105 Protein Control Substance	3.0 µg
8	Cry1A.105 Protein Reference 100 % Equivalence	3.0 µg
9	Cry1A.105 Protein Reference 10 % Equivalence	0.3 µg
10	Broad Range Molecular Weight Standards	4.5 µg

Heat Stability of Purified Cry2Ab2 Protein

The effect of heating at various temperatures on the functional activity of Cry2Ab2 protein for 15 min and 30 min is presented in Table 20 and Table 21, respectively. Temperatures of Cry2Ab2 for 15 min and 30 min at 55 ºC or above resulted in ≥ 96.4 % of detectable functional activity. SDS-PAGE analysis demonstrated that heating at 25, 37 and 55 ºC at 15 (Figure 38) or heating at 25, 37 and 55 ºC at 30 minutes (Figure 39), had no effect on the apparent band intensity of the Cry2Ab2 protein. When incubated at 75 ºC for 15 minutes a reduction in band intensity of the Cry2Ab2 protein was observed. The reduction in the band intensity became more pronounced for the Cry2Ab2 protein heated at 75 ºC for 30 minutes and 95 ºC for 15 minutes, and the band intensity was reduced further to only slightly higher than a 10% reference standard when heated at 95 ºC for 30 minutes. These data demonstrate that Cry2Ab2 behaves with a predictable tendency toward loss of functional activity at elevated temperatures.

Table 20. Activity of Cry2Ab2 Protein after 15 Minutes at Elevated Temperatures

Temperature	EC_{50}[1] (µg Cry2Ab2/ml diet)	Relative Activity (% of control substance)[3,4]
0 ºC (control substance)	0.068	100 %
25 ºC	0.077	88.3 %
37 ºC	0.11	61.8 %
55 ºC	1.87	3.6 %
75 ºC	>10.0[2]	< 0.7 %
95 ºC	>10.0[2]	< 0.7 %

[1]Determined from concentration response curves that included nine Cry2Ab2 concentrations ranging from 0.0391 – 10.0 µg Cry2Ab2 protein/ml diet. Each Cry2Ab2 concentration contained one replicate with a target number of sixteen Corn earworm (CEW) larvae.
[2]10 µg Cry2Ab2/ml diet represents the highest concentration tested.
[3]The control sample (not heated) was assigned as 100 % activity, and others are rounded to one decimal number.
[4]Relative Activity = (EC_{50} of Control substance µg/ml diet ÷ EC_{50} of heat treated sample µg /ml diet) × 100%

Table 21. Activity of Cry2Ab2 Protein after 30 Minutes at Elevated Temperatures

Temperature	EC_{50}[1] (µg Cry2Ab2/ml diet)	Relative Activity (% of control substance)[3,4]
0 ºC (control substance)	0.068 ± 0.0069	100 %
25 ºC	0.090 ± 0.0075	75.6 %
37 ºC	0.076 ± 0.0064	89.5 %
55 ºC	8.58	0.8 %
75 ºC	>10.0[2]	< 0.7 %
95 ºC	>10.0[2]	< 0.7 %

[1]Determined from concentration response curves that included nine Cry2Ab2 concentrations ranging from 0.0391 – 10.0 µg Cry2Ab2 protein/ml diet. Each Cry2Ab2 concentration contained one replicate with a target number of sixteen Corn earworm (CEW) larvae.
[2]10 µg Cry2Ab2/ml diet represents the highest concentration tested.
[3]The control sample (not heated) was assigned as 100 % activity, and others are rounded to one decimal number.
[4]Relative Activity = (EC_{50} of Control substance µg/ml diet ÷ EC_{50} of heat treated sample µg /ml diet) × 100%

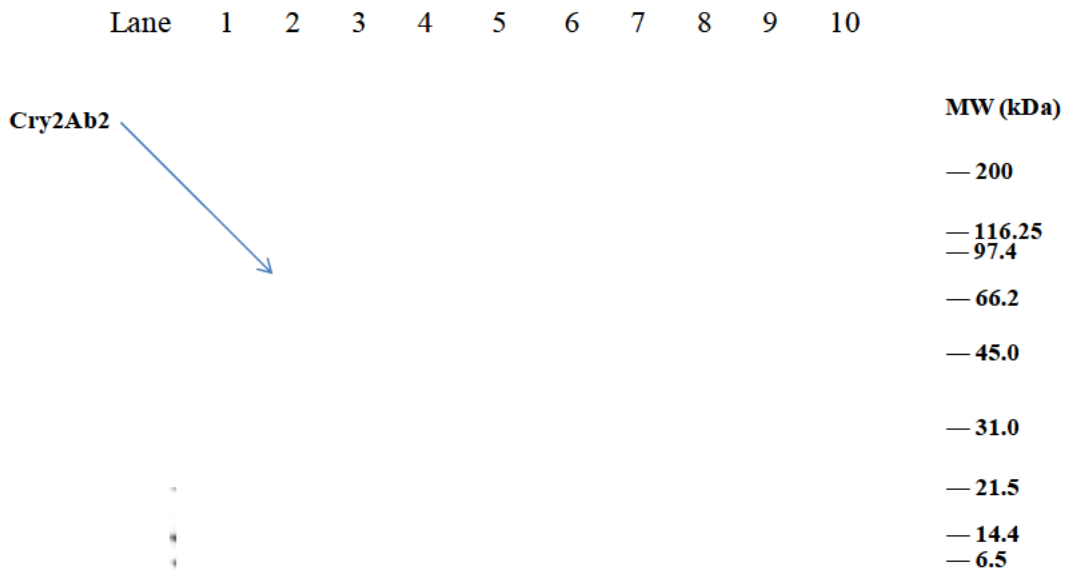


Figure 38. SDS-PAGE of Cry2Ab2 Protein Following Heat Treatment for 15 Minutes

Heat treated samples of Cry2Ab2 (3.0 µg total protein) were subjected to SDS-PAGE using a Tris-glycine 4-20 % polyacrylamide gel under denaturing and reducing conditions. The 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	Description	Total Amount
1	Broad Range Molecular Weight Standards	4.5 µg
2	Cry2Ab2 25 ºC for 15 minutes	3.0 µg
3	Cry2Ab2 37 ºC for 15 minutes	3.0 µg
4	Cry2Ab2 55 ºC for 15 minutes	3.0 µg
5	Cry2Ab2 75 ºC for 15 minutes	3.0 µg
6	Cry2Ab2 95 ºC for 15 minutes	3.0 µg
7	Cry2Ab2 Control	3.0 µg
8	Cry2Ab2 Reference Substance, 100 % Equivalence	3.0 µg
9	Cry2Ab2 Reference Substance, 10 % Equivalence	0.3 µg
10	Broad Range Molecular Weight Standards	4.5 µg

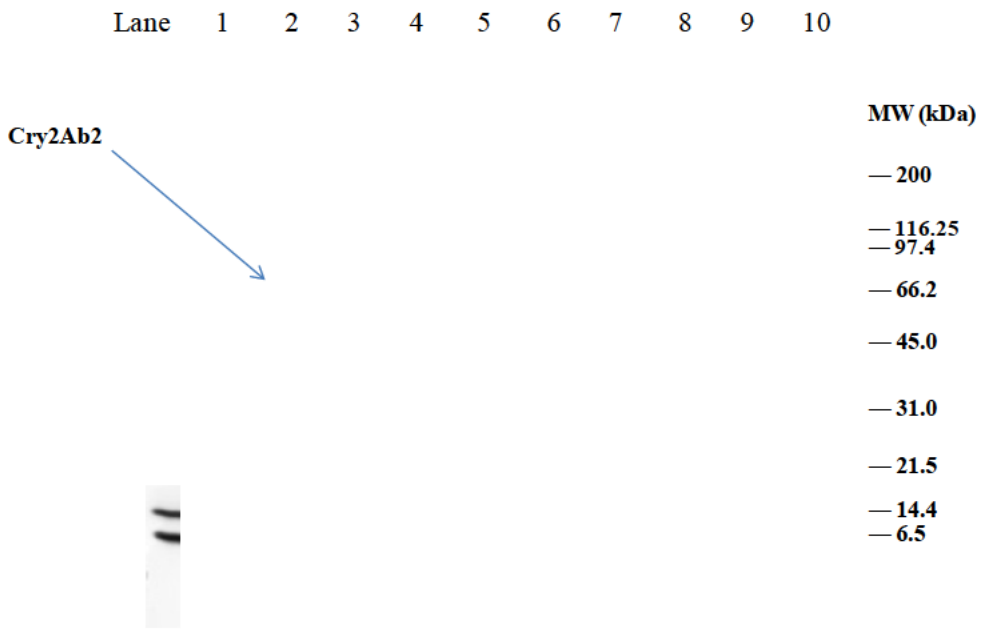


Figure 39. SDS-PAGE of Cry2Ab2 Protein Following Heat Treatement for 30 Minutes

Heat treated samples of Cry2Ab2 (3.0 µg total protein) were subjected to SDS-PAGE using a Tris-glycine 4-20 % polyacrylamide gel 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	Description	Total Amount
1	Broad Range Molecular Weight Standards	4.5 µg
2	Cry2Ab2 25 ºC for 30 minutes	3.0 µg
3	Cry2Ab2 37 ºC for 30 minutes	3.0 µg
4	Cry2Ab2 55 ºC for 30 minutes	3.0 µg
5	Cry2Ab2 75 ºC for 30 minutes	3.0 µg
6	Cry2Ab2 95 ºC for 30 minutes	3.0 µg
7	Cry2Ab2 Control	3.0 µg
8	Cry2Ab2 Reference Substance, 100 % Equivalence	3.0 µg
9	Cry2Ab2 Reference Substance, 10 % Equivalence	0.3 µg
10	Broad Range Molecular Weight Standards	4.5 µg

B4(c) Acute or short-term oral toxicity on novel protein(s)

Acute Oral Toxicity Study with the Cry1A.105 and Cry2Ab2 Proteins

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. (Delaney *et al.*, 2008; Codex Alimentarius, 2009; Hammond *et al.*, 2013). Based on the safety assessment of the Cry1A.105 and Cry2Ab2 proteins expressed in MON 87751 (Sections B4) toxicology studies are not considered necessary. However, acute toxicology studies, using the versions of the Cry1A.105 and Cry2Ab2 proteins expressed in MON 89034, are included to provide further confirmation of the safety of Cry1A.105 and Cry2Ab2. These toxicology studies resulted in very high NOAELs for each protein. While there are minor N-terminus sequence differences between the Cry1A.105 and Cry2Ab2 proteins used in the acute toxicity studies compared to the MON 87751-expressed proteins (Section B3(a)), the protease-resistant core domains of the Cry1A.105 and Cry2Ab2 proteins expressed in MON 87751 share 100% deduced amino acid identity to the core domains of the Cry1A.105 and Cry2Ab2 proteins expressed in MON 89034. These protease-resistant core domains are responsible for insecticidal toxicity and specificity (Gill *et al.*, 1992; Widner and Whiteley, 1989). Since the MON 87751 Cry1A.105 and Cry2Ab2 proteins are functionally equivalent to the MON 89034 Cry1A.105 and Cry2Ab2 proteins used in the acute toxicology assays, the conclusions from the toxicity testing are applicable to MON 87751 (Hammond *et al.*, 2013). Furthermore, neither Cry1A.105 nor Cry2Ab2 protein shares similarity to proteins that are harmful to human or animal health, and each version of Cry1A.105 and Cry2Ab2 has been shown to be readily inactivated by heat treatment and degraded by digestion in SGF. This demonstration that the activity and stability characteristics are consistent between the slightly different versions of Cry1A.105 and Cry2Ab2 proteins justifies the use of the previously generated acute toxicology data to support the safety assessment of MON 87751-produced Cry1A.105 and Cry2Ab2 proteins.

Acute Oral Toxicity Study with the Cry1A.105 Protein

Cry1A.105 protein was administered by oral gavage (as two doses about 4 hours apart) to 10 male and 10 female CD-1 mice at a total dose level of 2072 mg/kg body weight. There were no treatment-related effects on survival, clinical observations, body weight gain, food consumption or gross pathology. Therefore, the No Observable Adverse Effect Levels (NOAEL) for Cry1A.105 was considered to be 2072 mg/kg body weight, the highest dose tested.

Acute Oral Toxicity Study with the Cry2Ab2 Protein

Cry2Ab2 protein was administered by oral gavage (as two doses about 4 hours apart) to 10 male and 10 female CD-1 mice at a total dose level of 2198 mg/kg body weight. There were no treatment-related effects on survival, clinical observations, body weight gain, food consumption or gross pathology. Therefore, the NOAEL for Cry2Ab2 was considered to be 2198 mg/kg body weight, the highest dose tested.

Acute Oral Toxicity Studies Effectively Assess Potential Protein Hazards

Most known protein toxins act through acute mechanisms to exert toxicity (Hammond and Fuchs, 1998; Pariza and Johnson, 2001; Sjoblad *et al.*, 1992; Delaney *et al.*, 2008; Hammond *et al.*, 2013). The primary exceptions to this rule consist of certain anti-nutritional proteins such as lectins, which manifest toxicity in a short term (few weeks) feeding study (Liener, 1994). Neither Cry1A.105 nor Cry2Ab2 are similar to any anti-nutritional proteins or any other known protein harmful to human or animal health. Thus, acute studies were considered sufficient to evaluate the toxicity of the Cry1A.105 and Cry2Ab2 proteins.

For details please refer to Bonnette, 2006 (CRO-2005-049) and Bonnette, 2006 (CRO-2005-050).

B5 Assessment of Potential Allergenicity

History of safe use of the introduced protein is a key consideration in the potential for allergenicity and toxicity and for assessment of dietary safety. The history of safe use of Cry1A.105 and Cry2Ab2 proteins have been previously addressed in Section B3(f).

Additionally, following the guidelines adopted by the Codex Alimentarius Commission, an assessment of potential allergenicity of introduced proteins has been conducted by comparing the characteristics of the introduced protein to characteristics of known allergens (Codex Alimentarius, 2009). A protein is not likely to be associated with allergenicity if: 1) the protein is from a nonallergenic source; 2) the protein represents only a very small portion of the total plant protein; 3) the protein does not share structural similarities to known allergens based on the amino acid sequence; 4) the protein does not show resistance to pepsin digestion; or 5) the protein is not stable to heat treatment.

The Cry1A.105 and Cry2Ab2 proteins in MON 87751 have been assessed for their potential allergenicity according to these safety assessment guidelines.

B5(a) Source of introduced protein

The donor organism for *cry1A.105* and *cry2Ab2*, *B. thuringiensis,* has been used commercially in the United States to produce microbial-derived products with insecticidal activity. Applications of sporulated *B. thuringiensis* have a long history of safe use for pest control in agriculture, especially in organic farming (Cannon, 1993; WHO, 1999; U.S. EPA, 1988). Microbial pesticides containing *B. thuringiensis* Cry proteins have been subjected to extensive toxicity testing showing no adverse effects to human health (U.S. EPA, 2005; U.S. EPA, 2001; Baum *et al.*, 1999; Betz *et al.*, 2000; Mendelsohn *et al.*, 2003; McClintock *et al.*, 1995). Additionally, there are no confirmed cases of allergic reactions to Cry proteins in microbial-derived *B. thuringiensis* products during more than 50 years of use.

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

Structural Similarity of Cry1A.105 and Cry2Ab 2 to Known Allergens

B5(b)(i) Structural similarity of Cry1A.105 in MON 87751 to known 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 recommends that a sliding window search with a scientifically justified peptide size could be used to identify immunologically relevant peptides in otherwise unrelated proteins. Therefore, the extent of sequence similarities between the Cry1A.105 protein sequence and known allergens, gliadins, and glutenins was assessed using the FASTA sequence alignment tool and an eight-amino acid sliding window search (Codex Alimentarius, 2009; Thomas *et al.*, 2005). The data generated from these analyses confirm that the Cry1A.105 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[9]) was obtained from Food Allergy Research and Resource Program Database (FARRP, 2011) and was used for the evaluation of sequence similarities shared between the Cry1A.105 protein and all proteins. The AD_2013 database contains 1,491 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 that have an *E*-score of less than or equal to 1×10^{-5} are considered to have meaningful homology. Results indicate that the Cry1A.105 protein sequence does not share meaningful similarity with sequences in the allergen database. No alignment met nor 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 are not useful to the allergy assessment process (Thomas *et al.*, 2005). No eight contiguous amino acid identities were detected when the Cry1A.105 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 Cry1A.105 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 Cry1A.105 protein sequence and proteins in the allergen database. These data show that the Cry1A.105 protein sequence lacks both structurally and immunologically relevant similarities to known allergens, gliadins, and glutenins.

---

[9] The allergen, gliadin, and glutenin sequence database (AD_2013) was obtained from Food Allergy Research and Resource Program Database (FARRP_2013) (http://www.allergenonline.org) and was used for the evaluation of sequence similarities. It is referred to as the AD_2013 database and contains 1,630 sequences.

[redacted]

B5(b)(ii) Structural similarity of Cry2Ab2 in MON 87751 to known allergens

The FASTA analysis (as described in Section B5(b)(i)) indicates that the Cry2Ab2 protein sequence does not share significant 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}. Furthermore, no short (eight amino acid) polypeptide matches were shared between the Cry2Ab2 protein sequence and proteins in the allergen database. These data show that the Cry2Ab2 protein sequence lacks both structurally and immunologically relevant sequence similarities to known allergens, gliadins, and glutenins.

[redacted]

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

Digestibility and Heat Stability of the Cry1A.105 and Cry2Ab2 Proteins

Both the Cry1A.105 and Cry2Ab2 proteins were readily digestible in simulated gastric fluid (SGF) that demonstrate the susceptibility of proteins to digestion by pepsin. Digestion in simulated intestinal fluid (SIF) was also assessed. As expected, Cry1A.105 and Cry2Ab2 subjected to SIF were each processed to a trypsin-resistant core that was stable throughout the SIF digestion period. These results are consistent with observations for these and other Cry proteins subjected to SIF digestion. Rapid degradation of the Cry1A.105 and Cry2Ab2 proteins in SGF makes it highly unlikely that either protein would be absorbed in a form other than component nutritional amino acids in the small intestine and have any adverse effects on human or animal health.

Both Cry1A.105 and Cry2Ab2 proteins when heated to 75ºC or above showed a greater than 96% loss in detectable functional activity. Thus, both Cry1A.105 and Cry2Ab2 behave with a predictable tendency toward loss of functional activity at elevated temperatures. Therefore, it is anticipated that exposure to functionally active Cry1A.105 and Cry2Ab2 protein from the consumption of MON 87751 or foods derived from MON 87751 will be negligible.

For further details please refer to Section B4(b).

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

Not applicable.

B5(e) Protein as a proportion of total protein

The Cry1A.105 and Cry2Ab2 Proteins as a Proportion of Total Protein

The Cry1A.105 protein was detected in all plant tissues assayed at a number of time points during the growing season (Table 16), with the exception of root tissues which were below the limit of detection. Among tested tissues of MON 87751, grain is the most relevant to the assessment of food allergenicity since grain is a food source. The mean level of Cry1A.105 protein in grain of MON 87751 is 2.4 µg/g dw. The mean percent dry weight of

total protein in grain of MON 87751 is 40.58% (or 405,800 µg/g; Table 16). The percentage of Cry1A.105 protein as a proportion of total protein in MON 87751 grain is approximately 0.0006%, and is calculated as follows:

$$(2.4\ \mu g/g \div 405{,}800\ \mu g/g) \times 100\% \approx 0.0006\%\ \text{of total protein}$$

Therefore, the Cry1A.105 protein represents a very small portion of the total protein in the harvested grain of MON 87751.

The Cry2Ab2 protein was detected in all plant tissues assayed, at a number of time points during the growing season (Table 17). Among tested tissues of MON 87751, grain is the most relevant to the assessment of food allergenicity since grain is a food source. The mean level of Cry2Ab2 protein in grain of MON 87751 is 4.0 µg/g dw. The mean percent dry weight of total protein in grain of MON 87751 is 40.58% (or 405,800 µg/g; Table 17). The percentage of Cry2Ab2 protein as a proportion of total protein in MON 87751 grain is approximately 0.001%, and is calculated as follows:

$$(4.0\ \mu g/g \div 405{,}800\ \mu g/g) \times 100\% \approx 0.001\%\ \text{of total protein}$$

Therefore, the Cry2Ab2 protein represents a very small portion of the total protein in the harvested grain of MON 87751.

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

Not applicable.

B7 Compositional Assessment

Seed and forage samples were collected from MON 87751, a conventional control and a total of 19 different reference varieties grown in the United States during the 2012 field season. The reference varieties were included in the compositional analysis 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 87751, the conventional control and reference soybean varieties were grown under normal agronomic field conditions for their respective geographic regions.

The evaluation of MON 87751 followed considerations relevant to the compositional quality of soybean as defined by the OECD consensus document (OECD, 2012). Seed samples were assessed for levels of nutrients including proximates, carbohydrates by calculation, fiber, amino acids, fatty acids, vitamins, and minerals. The anti-nutrients assessed in seed included lectin, trypsin inhibitors, phytic acid, raffinose and stachyose. Other components assessed included isoflavones. Forage samples were assessed for proximates, carbohydrates by calculation and fiber. In all, 66 different components were analyzed.

Of the 66 measured components, 14 fatty acids (caprylic, capric, lauric, myristic, myristoleic, pentadecanoic, pentadecenoic, palmitoleic, heptadecanoic, heptadecenoic, gamma linolenic, eicosadienoic, eicosatrienoic, and arachidonic acids) had more than 50% of the observations below the assay limit of quantitation (LOQ) and as a result, were excluded from the statistical analyses. Moisture values were measured for conversion of components to dry weight in seed and forage, but were not statistically analyzed. Therefore, statistics were provided for 50 components for all samples (six in forage and 44 in seed).

The statistical comparison of MON 87751 and the conventional control was based on compositional data combined across all field sites (the combined-site analysis). Statistically significant differences were identified at the 5% level ($\alpha = 0.05$). The compositional data from the reference varieties were combined across all field sites to calculate a 99% tolerance interval for each component to estimate the natural variability of each component in soybean.

A statistically significant difference between MON 87751 and the conventional control does not necessarily imply biological relevance from a food and feed perspective. Therefore, statistically significant differences observed in the combined-site analysis between MON 87751 and the conventional control were evaluated further to determine whether the detected difference indicated a biologically relevant compositional change or supported a conclusion of compositional equivalence, as follows:

1) Determination of the mean difference between MON 87751 and the conventional control to be used in steps two and three. For protein and amino acids only[10], the relative magnitude of the difference (percent change relative to the control) between MON 87751 and the conventional control was determined to allow an assessment of the difference in amino acids in relation to a difference in protein;

2) Assessment of the relative impact of MON 87751 in the context of variation within the conventional control germplasm grown across multiple sites (*i.e.* variation due to environmental influence). This assessment evaluates the mean difference between MON 87751 and the conventional control in the context of the range of values for the conventional control (maximum value minus the minimum value) derived from the combined-site analysis. When a mean difference is less than ranges seen due to natural environmental variation within even a single, closely related germplasm, the difference is typically not a food or feed safety concern; and

3) Assessment of the relative impact of MON 87751 compared to natural variation due to multiple sources (*e.g.* environmental and germplasm influences). This assessment compares the mean difference between MON 87751 and the conventional control to variation in conventional soybean as estimated by in-study reference variety values and assessing whether the mean component value of MON 87751 was within the reference variety 99% tolerance interval, literature values and/or the ILSI Crop Composition Database values (ILSI-CCDB) (ILSI, 2011).

These evaluations of natural variation within the conventional control and conventional soybean references are important as crop composition is known to be greatly influenced by environment and variety (Harrigan *et al.*, 2009). Only if mean differences between MON 87751 and the conventional control are large relative to natural variation inherent to conventional soybean would further assessments be required to establish whether the change in composition would have an impact from a food and feed safety or nutritional perspective. The steps reviewed in this assessment therefore describe whether the differences between MON 87751 and the conventional control are meaningful from a food/feed perspective or whether they support a conclusion of compositional equivalence.

The compositional analysis provided a comprehensive comparative assessment of the levels of key nutrients, anti-nutrients, and other components in soybean seed and forage of MON 87751 and the conventional control. Of the 50 components statistically assessed for MON 87751 there were no statistically significant differences in 42 components. Only eight

[10] Since total amino acids measured in a seed analysis are predominately derived from hydrolysis of protein, changes in protein levels may have a corresponding impact on levels of individual amino acids. However, the mean difference for individual amino acid levels will be less than the mean difference for protein because each amino acid constitutes only a part of the protein. For this reason, the relative magnitudes of difference (percent change relative to the control) for amino acids and protein were determined to allow an assessment of the difference in amino acids in relation to a difference in protein. When the relative magnitudes of difference for amino acids were related to the relative magnitude of difference for protein, then steps 2 and 3 are not discussed for amino acids.

components (protein, glycine, proline, phosphorus, vitamin E, and raffinose in seed, and total fat and NDF in forage) showed a significant difference ($p<0.05$) between MON 87751 and the conventional control. For these components, the mean difference in component values between MON 87751 and the conventional control was less than the range of the conventional control values. Additionally, the mean difference in component values between MON 87751 and the conventional control was less than the range of reference soybean variety values. Finally, the MON 87751 mean component values were within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB values. These results support the overall conclusion that MON 87751 was not a major contributor to variation in component levels in soybean seed and forage and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components. A detailed description of the assessment of statistically significant differences observed between MON 87751 and the conventional control is provided in the following section. These data indicated that the components with significant differences were not compositionally meaningful from a food and feed safety perspective.

[redacted]

B7(a) Levels of key nutrients, toxicants and anti-nutrients

Nutrient Levels in Soybean Seed

Seed samples were analyzed for levels of nutrients including proximates (four components), carbohydrates by calculation, fiber (two components), amino acids (18 components), fatty acids (22 components), vitamin E (α-tocopherol), vitamin K1 (phylloquinone) and minerals (2 components). Moisture was measured for conversion of components to dry weight, but was not statistically analyzed.

Proteins and Amino Acids

Protein levels in soybean seed generally average ~40% dry weight (dwt), with values reported in the USDA soybean germplasm collection, for example, ranging from 34.1 to 56.8% dwt (Wilson, 2004). Protein content is a quantitative trait controlled by many genetic loci (Panthee *et al.*, 2005; Akond *et al.*, 2012) and is influenced by both genotype and environment (Rotundo and Westgate, 2009).

A statistically significant difference ($p<0.05$) between MON 87751 and the conventional control was observed for protein (Table 22). The mean protein value was 40.58% dwt for MON 87751 and 40.12% dwt for the conventional control, a difference of 0.46% dwt. This difference was evaluated in the context of the range of the conventional control values; 5.09% dwt, calculated from the minimum (37.88% dwt) and maximum (42.97% dwt) protein values. The mean difference in protein values between MON 87751 and the conventional control was less than the range of the conventional control values, indicating that MON 87751 does not impact levels of protein more than natural variation within the conventional control grown at multiple locations. The mean difference in protein values between MON 87751 and the conventional control was also less than the variation seen in the reference variety values (ranged 34.68 to 45.22% dwt, a magnitude of 10.54% dwt), and the

MON 87751 mean value was within the 99% tolerance interval, the values observed in the literature, and the ILSI-CCDB (Table 29).

Since total amino acids measured in a seed analysis are predominantly derived from hydrolysis of protein, differences in amino acid levels between MON 87751 and the conventional control were assessed relative to the difference in protein levels. The relative magnitude of the difference in mean protein values for MON 87751 and the conventional control was 1.14% (Table 22). Correspondingly, relative magnitudes of difference for the 18 amino acids were all ≤ 2.17%. The difference between MON 87751 and the conventional control were significant for two of the amino acids (glycine and proline) (Table 22), and reflected small relative magnitudes of differences between MON 87751 and the conventional control, as would be expected based on the small relative magnitude of difference in protein.

The data demonstrated that MON 87751 was not a major contributor to variation in protein and amino acid levels in soybean seed and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components. Also, the mean values of these components were within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB values (Table 29). These data confirmed that the significant differences in mean values of protein, glycine and proline were not compositionally meaningful from a food and feed safety perspective.

Total Fat and Fatty Acids

Fat levels in soybean seed generally average ~20% dwt, with values reported in the USDA soybean germplasm collection, for example, ranging from 8.3 to 27.9% (Wilson, 2004). Total fat content is a quantitative trait controlled by many genetic loci (Panthee *et al.*, 2005; Akond *et al.*, 2012) and is influenced by both genotype and environment (Rotundo and Westgate, 2009). There were no significant differences in seed total fat and fatty acid content between MON 87751 and the conventional control (Table 23). The data demonstrated that MON 87751 was not a major contributor to variation in total fat and fatty acid levels in soybean seed and confirmed the similarity of MON 87751 to the conventional control in levels of these components.

A total of fourteen fatty acids (caprylic, capric, lauric, myristic, myristoleic, pentadecanoic, pentadecenoic, palmitoleic, heptadecanoic, heptadecenoic, gamma linolenic, eicosadienoic, eicosatrienoic, and arachidonic acids) with more than 50% of observations below the assay LOQ were excluded from statistical analysis. These fatty acids are present in only low amounts in soybean seed, if present at all (Berman *et al.*, 2009; Lundry *et al.*, 2008). This study confirmed that this observation extended to MON 87751.

Carbohydrates by Calculation and Fiber

In addition to protein and fat, the major biomass components assessed in soybean seed included carbohydrates by calculation and fiber (ADF and NDF). There were no significant differences in carbohydrate and fiber content between MON 87751 and the conventional control (Table 24). The data demonstrated that MON 87751 was not a major contributor to

variation in carbohydrate and fiber levels in soybean seed and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components.

Ash and Minerals

Ash was assessed in soybean seed, in addition to protein, fat, carbohydrates by calculation, and fiber. Mineral components (calcium and phosphorus), a constituent of ash, are discussed in this section.

No statistically significant differences were observed for ash or calcium (Table 25). A significant difference ($p<0.05$) between MON 87751 and the conventional control was observed for phosphorus (Table 25).

For phosphorus, the mean value was 0.54 g/100g dwt for MON 87751 and 0.53 g/100g dwt for the conventional control, a difference of 0.010 g/100g dwt. This difference was evaluated in the context of the range of the conventional control values, 0.23 g/100g dwt, calculated from the minimum (0.44 g/100g dwt) and maximum (0.67 g/100g dwt) phosphorus values. The mean difference was less than the range of the conventional control values, indicating that MON 87751 does not impact levels of phosphorus more than natural variation within the conventional control grown at multiple locations. The mean difference in phosphorus values between MON 87751 and the conventional control was also less than the variability seen in the respective reference variety values (ranged 0.42 to 0.71 g/100 g dwt, a magnitude of 0.29 g/100 g dwt), and the MON 87751 mean phosphorus value was within the reference variety 99% tolerance interval, the values observed in the literature and the ILSI-CCDB (Table 29).

The data demonstrated that MON 87751 was not a major contributor to variation in ash and mineral levels in soybean seed and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components. Also, the mean values of these components were within the 99% tolerance interval, the values observed in the literature, and the ILSI-CCDB values. These data confirmed that the significant difference in the mean value of phosphorus was not compositionally meaningful from a food and feed safety perspective.

Vitamins

Soybean oil is considered a source of biologically available vitamin K1 with levels ranging from 102.5 to 250 μg/100 g of oil reported in the published literature (OECD, 2012). Vitamin E (α-tocopherol) is an important nutrient and maintains oxidative stability of soybean oil. Vitamin E levels in soybean seed are known to be affected by environment and germplasm (Seguin *et al.*, 2010); levels in soybean seed harvested from six different locations in Eastern Canada over a single year ranged from 0.87 to 3.32 mg/100 g dwt (Seguin *et al.*, 2009).

No statistically significant difference was observed for vitamin K1 (Table 26). A statistically significant difference ($p<0.05$) was observed for vitamin E.

For vitamin E the mean value was 2.59 mg/100 g dwt for MON 87751 and 2.78 mg/100 g dwt for the conventional control, a difference of -0.19 mg/100 g dwt. This difference was evaluated in the context of the range of the conventional control values, 3.04 mg/100 g dwt, calculated from the minimum (1.36 mg/100 g dwt) and maximum (4.39 mg/100 g dwt) vitamin E values. The mean difference was less than the range of the conventional control values, indicating that the presence of MON 87751 does not impact levels of vitamin E more than natural variation within the conventional control grown at multiple locations. The mean difference in vitamin E values between MON 87751 and the conventional control was also less than the variation seen in the respective reference variety values (ranged 1.04 to 4.99 mg/100g dwt, a magnitude of 3.95 mg/100 g dwt). The MON 87751 mean vitamin E value was within the 99% tolerance interval, the values observed in the literature and the ILSI-CCDB (Table 29).

The data demonstrated that MON 87751 was not a major contributor to variation in vitamin levels in soybean seed and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components. Also, the mean values of these components were within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB values. These data confirmed that the significant difference in the mean value of vitamin E was not compositionally meaningful from a food and feed safety perspective.

Anti-Nutrient Levels in Soybean Seed

Anti-nutrients assessed include lectin, trypsin inhibitors, phytic acid, raffinose and stachyose. Lectins and trypsin inhibitors are proteins; lectins have carbohydrate-binding properties whereas trypsin inhibitors can inhibit protein digestion. Both proteins can negatively impact animal growth and performance and as a result, soybeans are processed to denature these proteins prior to consumption (Qin *et al.*, 1996). Variation of nearly five-fold in lectin levels was reported based on a screen of over 100 varieties (Pull *et al.*, 1978). A two-fold variation in trypsin inhibitor activity attributed to genetic and environmental factors has also been observed in soybean (Vollmann *et al.*, 2003; Kumar *et al.*, 2003).

Phytic acid is considered an anti-nutrient because of its mineral-chelating properties. Levels in soybean seed are influenced by factors such as genotype, environment, soil type and agronomic practices (Raboy *et al.*, 1984; Raboy and Dickinson, 1993). The oligosaccharides, stachyose and raffinose, are considered anti-nutrients because of their contribution to gas production and resulting flatulence following consumption (OECD, 2012). Raffinose and stachyose levels in soybean are quantitative traits and influenced by genotype and environment (Kumar *et al.*, 2010).

No statistically significant differences were observed for lectin, trypsin inhibitors, phytic acid, and stachyose (Table 27). A statistically significant difference ($p<0.05$) was observed for raffinose.

For raffinose, the mean value was 0.88% dwt for MON 87751 and 0.95% dwt for the conventional control, a difference of -0.065% dwt. This difference was evaluated in the context of the range of the conventional control values (0.88% dwt), calculated from the

minimum (0.62% dwt) and maximum (1.50% dwt) raffinose values. The mean difference in raffinose values between MON 87751 and the conventional control was less than the range of the conventional control values, indicating that MON 87751 does not impact levels of raffinose more than natural variation within the conventional control grown at multiple locations. The mean difference in raffinose values between MON 87751 and the conventional control was also less than the variation seen in the respective reference values (ranged 0.54 to 1.45% dwt, a magnitude of 0.91% dwt) and the MON 87751 mean value was within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB (Table 29).

The data demonstrated that MON 87751 was not a major contributor to variation in anti-nutrient levels in soybean seed and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components. Also the mean values of these components were within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB values (Table 29). These data confirmed that the significant difference in the mean value of raffinose was not compositionally meaningful from a food and feed safety perspective.

Other Components in Soybean Seed

In addition to the nutrients and anti-nutrients analyzed in soybean seed the levels of isoflavones were also assessed. Isoflavones have been reported to have biological activity as estrogenic and anti-estrogenic compounds and they may impact animal reproduction when consumed in large quantities (OECD, 2012). Soybean isoflavones demonstrate considerable variation in levels in mature seed. Eldridge and Kwolek (1983) observed that the concentration of isoflavones varies from variety to variety, and there are also differences when the same variety is grown in different locations; a seven-fold variation was observed for levels of daidzein, while glycitein and genistein presented 2.5 and 76-fold variation respectively. Other studies further demonstrate the extensive variability observed in isoflavone levels (Gutierrez-Gonzalez *et al.*, 2009; Hoeck *et al.*, 2000; Morrison *et al.*, 2008). No statistically significant differences were observed for daidzein, genistein and glycitein (Table 27). The data provided here demonstrated that MON 87751 is not a major contributor to variation in isoflavone levels in soybean seed and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components.

Nutrient Levels in Soybean Forage

Forage samples were assessed for proximates, carbohydrates by calculation, and fiber. No statistically significant differences were observed for ash, carbohydrates by calculation, protein, and ADF (Table 28). Statistically significant differences ($p<0.05$) were observed for total fat and NDF.

For total fat, the mean value was 6.03% dwt for MON 87751 and 6.43% dwt for the conventional control, a difference of -0.40% dwt. This difference was evaluated in the context of the range of the conventional control values (5.02% dwt), calculated from the minimum (4.04% dwt) and maximum (9.06% dwt) values. The mean difference in total fat values between MON 87751 and the conventional control was less than the range of the

conventional control values, indicating that MON 87751 does not impact levels more than natural variation within the conventional control grown at multiple locations. The mean difference in total fat values between MON 87751 and the conventional control was also less than the variation seen in the reference variety values (ranged 2.74 to 9.74% dwt, a magnitude of 7.00% dwt), and the MON 87751 mean value was within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB (Table 29).

For NDF, the mean value was 36.77% dwt for MON 87751 and 34.08% dwt for the conventional control, a difference of 2.69% dwt. This difference was evaluated in the context of the range of the conventional control values (14.36% dwt), calculated from the minimum (27.69% dwt) and maximum (42.05% dwt) NDF values. The mean difference in NDF values between MON 87751 and the conventional control was less than the range of the conventional control values, indicating that MON 87751 does not impact levels more than natural variation within the conventional control grown at multiple locations. The mean difference in NDF values between MON 87751 and the conventional control was also less than the variation seen in the reference variety values (ranged 25.71 to 52.96% dwt, a magnitude of 27.25% dwt), and the MON 87751 mean value was within the 99% tolerance interval and the values observed in the literature (Table 29).

The data demonstrated that MON 87751 was not a major contributor to variation in proximates, carbohydrates by calculation, and fiber levels in soybean forage and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components. Also, the mean values of these components were within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB values. These data confirmed that the significant difference in the mean values of total fat and NDF were not compositionally meaningful from a food and feed safety perspective.

Table 22. Summary of Soybean Seed Protein and Amino Acids for MON 87751, the Conventional Control and References

Component (% dwt)[1]	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	p-Value	% Relative[6]
Protein	40.58 (0.50) 38.01 - 43.87	40.12 (0.50) 37.88 - 42.97	(34.68 - 45.22) 34.33, 45.17	5.09	0.46 (0.16)	0.023	1.14
Alanine	1.75 (0.018) 1.66 - 1.89	1.75 (0.018) 1.67 - 1.92	(1.58 - 1.95) 1.53, 1.93	0.25	0.0039 (0.0077)	0.611	0.22
Arginine	3.03 (0.047) 2.75 - 3.28	3.00 (0.047) 2.72 - 3.36	(2.48 - 3.62) 2.42, 3.52	0.64	0.028 (0.026)	0.317	0.93
Aspartic Acid	4.51 (0.048) 4.23 - 4.86	4.48 (0.048) 4.17 - 4.89	(3.82 - 5.18) 3.84, 5.07	0.72	0.032 (0.022)	0.185	0.72
Cystine	0.56 (0.017) 0.46 - 0.63	0.57 (0.017) 0.42 - 0.66	(0.42 - 0.67) 0.43, 0.67	0.25	-0.0010 (0.0081)	0.903	-0.18
Glutamic Acid	6.89 (0.085) 6.37 - 7.38	6.85 (0.085) 6.17 - 7.51	(5.64 - 8.18) 5.60, 7.96	1.34	0.044 (0.047)	0.360	0.64

Table 22. Summary of Soybean Seed Protein and Amino Acids for MON 87751, the Conventional Control and References (continued)

Component (% dwt)[1]	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	p-Value	% Relative[6]
Glycine	1.73 (0.017) 1.64 - 1.86	1.71 (0.017) 1.63 - 1.83	(1.51 - 1.90) 1.47, 1.89	0.19	0.018 (0.0069)	0.014	1.04
Histidine	1.01 (0.011) 0.89 - 1.09	1.01 (0.011) 0.94 - 1.11	(0.87 - 1.15) 0.90, 1.11	0.17	0.0076 (0.010)	0.469	0.76
Isoleucine	1.92 (0.022) 1.79 - 2.09	1.90 (0.022) 1.78 - 2.06	(1.66 - 2.23) 1.63, 2.13	0.28	0.015 (0.013)	0.296	0.80
Leucine	3.04 (0.032) 2.82 - 3.28	3.02 (0.032) 2.86 - 3.23	(2.70 - 3.43) 2.68, 3.36	0.37	0.023 (0.016)	0.186	0.76
Lysine	2.51 (0.025) 2.35 - 2.77	2.49 (0.025) 2.32 - 2.71	(2.18 - 2.84) 2.16, 2.77	0.40	0.024 (0.014)	0.096	0.97
Methionine	0.56 (0.0092) 0.51 - 0.61	0.55 (0.0092) 0.47 - 0.63	(0.47 - 0.63) 0.45, 0.64	0.16	0.012 (0.0071)	0.135	2.17

Table 22. Summary of Soybean Seed Protein and Amino Acids for MON 87751, the Conventional Control and References (continued)

Component (% dwt)[1]	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	p-Value	% Relative[6]
Phenylalanine	2.02 (0.024) 1.87 - 2.17	2.01 (0.024) 1.83 - 2.20	(1.74 - 2.36) 1.75, 2.27	0.37	0.013 (0.011)	0.290	0.65
Proline	2.07 (0.025) 1.91 - 2.24	2.03 (0.025) 1.90 - 2.18	(1.76 - 2.32) 1.72, 2.32	0.28	0.032 (0.011)	0.025	1.59
Serine	1.90 (0.022) 1.74 - 2.10	1.89 (0.022) 1.72 - 2.09	(1.59 - 2.23) 1.60, 2.17	0.37	0.0067 (0.020)	0.734	0.36
Threonine	1.58 (0.011) 1.50 - 1.66	1.57 (0.011) 1.50 - 1.66	(1.42 - 1.74) 1.39, 1.72	0.17	0.012 (0.0076)	0.127	0.75
Tryptophan	0.55 (0.0085) 0.49 - 0.61	0.54 (0.0085) 0.49 - 0.61	(0.46 - 0.61) 0.47, 0.60	0.12	0.0081 (0.0053)	0.131	1.49
Tyrosine	1.55 (0.013) 1.48 - 1.66	1.54 (0.013) 1.47 - 1.62	(1.38 - 1.72) 1.35, 1.71	0.14	0.014 (0.0062)	0.066	0.88

Table 22. Summary of Soybean Seed Protein and Amino Acids for MON 87751, the Conventional Control and References (continued)

Component (% dwt)[1]	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	p-Value	% Relative[6]
Valine	1.87 (0.017) 1.73 - 1.99	1.87 (0.017) 1.74 - 1.96	(1.64 - 2.16) 1.66, 2.05	0.22	0.0073 (0.0098)	0.460	0.39

[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 soybean varieties.
[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference soybean varieties. Negative limits set to zero.
[5]Maximum value minus minimum value for the control soybean variety.
[6]The relative magnitude of the difference in mean values between MON 87751 and the control, expressed as a percent of the control.

Table 23. Summary of Soybean Seed Total Fat and Fatty Acids for MON 87751, the Conventional Control and References

Component	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
Total Fat (% dwt)[1]	19.21 (0.43) 16.15 - 21.21	19.48 (0.43) 16.84 - 21.08	(18.10 - 22.97) 17.12, 24.20	4.24	-0.27 (0.14)	0.088
16:0 Palmitic[6]	11.69 (0.19) 11.13 - 12.26	11.59 (0.19) 10.03 - 12.36	(9.37 - 12.56) 8.39, 13.35	2.33	0.10 (0.18)	0.603
18:0 Stearic	4.22 (0.12) 3.61 - 4.95	4.29 (0.12) 3.65 - 4.98	(3.27 - 6.11) 2.12, 6.45	1.32	-0.066 (0.036)	0.106
18:1 Oleic	20.81 (0.94) 17.08 - 26.89	21.50 (0.94) 16.98 - 27.41	(17.21 - 34.03) 13.27, 32.25	10.42	-0.69 (0.50)	0.211
18:2 Linoleic	54.57 (0.73) 50.33 - 56.50	53.93 (0.73) 47.96 - 56.99	(45.98 - 58.27) 44.95, 62.64	9.03	0.63 (0.30)	0.074
18:3 Linolenic	7.89 (0.38) 6.09 - 9.97	7.86 (0.38) 6.19 - 10.23	(5.37 - 10.34) 4.29, 10.55	4.04	0.035 (0.10)	0.740

Table 23. Summary of Soybean Seed Total Fat and Fatty Acids for MON 87751, the Conventional Control and References (continued)

Component	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
20:0 Arachidic	0.32 (0.0087) 0.28 - 0.37	0.33 (0.0087) 0.28 - 0.39	(0.26 - 0.50) 0.17, 0.50	0.11	-0.0066 (0.0039)	0.133
20:1 Eicosenoic	0.19 (0.0084) 0.14 - 0.24	0.19 (0.0084) 0.14 - 0.25	(0.13 - 0.25) 0.12, 0.26	0.11	-0.00023 (0.0037)	0.953
22:0 Behenic	0.31 (0.0090) 0.28 - 0.36	0.32 (0.0090) 0.27 - 0.40	(0.28 - 0.49) 0.20, 0.48	0.13	-0.0075 (0.0075)	0.354

[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 soybean varieties.
[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference soybean varieties. Negative limits set to zero.
[5]Maximum value minus minimum value for the control soybean variety.
[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 represents 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: 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, arachidonic acid, and caprylic acid.

Table 24. Summary of Soybean Seed Carbohydrates by Calculation and Fiber for MON 87751, the Conventional Control and References

Component (% dwt)[1]	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
Carbohydrates by	35.30 (0.58)	35.63 (0.58)	(29.42 - 38.18)	6.63	-0.33 (0.19)	0.131
Calculation	31.27 - 37.69	31.95 - 38.57	29.43, 39.69			
Acid Detergent Fiber	13.99 (0.44)	13.98 (0.44)	(10.02 - 17.59)	4.91	0.010 (0.60)	0.986
	10.75 - 17.44	12.15 - 17.06	10.71, 17.25			
Neutral Detergent Fiber	15.56 (0.33)	15.57 (0.33)	(12.25 - 19.29)	4.71	-0.0088 (0.40)	0.982
	12.56 - 17.65	13.97 - 18.68	12.66, 18.44			

[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 soybean varieties.
[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference soybean varieties. Negative limits set to zero.
[5]Maximum value minus minimum value for the control soybean variety.

Table 25. Summary of Soybean Seed Ash and Minerals for MON 87751, the Conventional Control and References

Component	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
Ash (% dwt)[1]	4.91 (0.10) 4.46 - 5.56	4.89 (0.10) 4.34 - 5.70	(4.16 - 5.69) 4.32, 5.74	1.36	0.013 (0.046)	0.789
Calcium (g/100g dwt)	0.29 (0.010) 0.24 - 0.34	0.29 (0.010) 0.25 - 0.34	(0.21 - 0.40) 0.20, 0.41	0.10	-0.0083 (0.0088)	0.376
Phosphorus (g/100g dwt)	0.54 (0.020) 0.45 - 0.68	0.53 (0.020) 0.44 - 0.67	(0.42 - 0.71) 0.40, 0.70	0.23	0.010 (0.0041)	0.020

[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 soybean varieties.
[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference soybean varieties. Negative limits set to zero.
[5]Maximum value minus minimum value for the control soybean variety.

Table 26. Summary of Soybean Seed Vitamins for MON 87751, the Conventional Control and References

Component	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
Vitamin E (α-Tocopherol)	2.59 (0.29)	2.78 (0.29)	(1.04 - 4.99)	3.04	-0.19 (0.042)	<0.001
(mg/100g dwt)[1]	1.21 - 4.05	1.36 - 4.39	0, 5.12			
Vitamin K (phylloquinone)	0.64 (0.067)	0.67 (0.067)	(0.28 - 0.98)	0.59	-0.029 (0.016)	0.111
(µg/g dwt)	0.40 - 0.96	0.41 - 1.00	0.036, 1.10			

[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 soybean varieties.
[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference soybean varieties. Negative limits set to zero.
[5]Maximum value minus minimum value for the control soybean variety.

Table 27. Summary of Soybean Seed Anti-nutrients and Isoflavones for MON 87751, the Conventional Control and References

Component	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
Anti-nutrient						
Lectin (H.U./mg dwt)[1]	3.42 (0.23) 1.78 - 5.22	3.18 (0.23) 1.71 - 5.61	(1.63 - 7.46) 0.59, 6.30	3.90	0.24 (0.24)	0.353
Phytic Acid (% dwt)	1.22 (0.067) 0.85 - 1.78	1.19 (0.067) 0.72 - 1.68	(0.88 - 1.82) 0.80, 1.68	0.96	0.024 (0.019)	0.234
Raffinose (% dwt)	0.88 (0.087) 0.59 - 1.47	0.95 (0.087) 0.62 - 1.50	(0.54 - 1.45) 0.44, 1.27	0.88	-0.065 (0.018)	0.007
Stachyose (% dwt)	4.09 (0.079) 3.66 - 4.37	4.06 (0.079) 3.60 - 4.47	(3.39 - 4.62) 3.15, 4.80	0.87	0.030 (0.060)	0.628
Trypsin Inhibitor (TIU/mg dwt)	26.21 (1.97) 16.32 - 40.79	26.82 (1.97) 14.90 - 44.39	(14.79 - 42.42) 12.17, 38.77	29.49	-0.61 (1.48)	0.693
Isoflavone (µg/g dwt)						
Daidzein	901.33 (98.11) 223.46 - 1313.65	893.90 (98.11) 288.55 - 1243.53	(164.64 - 1312.77) 0, 1494.46	954.99	7.43 (42.93)	0.867

Table 27. Summary of Soybean Seed Anti-nutrients and Isoflavones for MON 87751, the Conventional Control and References (continued)

Component	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
Isoflavone (µg/g dwt)						
Genistein	756.78 (77.73) 243.18 - 993.17	755.11 (77.73) 273.13 - 1101.04	(245.14 - 1318.18) 150.41, 1437.69	827.91	1.67 (17.78)	0.925
Glycitein	77.67 (5.37) 37.29 - 117.25	84.03 (5.37) 47.91 - 132.77	(37.32 - 210.17) 8.04, 211.36	84.86	-6.36 (4.13)	0.129

[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 soybean varieties.
[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference soybean varieties. Negative limits set to zero.
[5]Maximum value minus minimum value for the control soybean variety.

Table 28. Summary of Soybean Forage Proximates and Fiber for MON 87751, the Conventional Control and References

Component (% dwt)[1]	MON 87751 Mean (S.E.)[2] Range	Control Mean (S.E.) Range	Conventional Reference (Range)[3] Tolerance Interval[4]	Control Range Value[5]	Difference (MON 87751 - Control) Mean (S.E.)	Difference (MON 87751 - Control) p-Value
Ash	6.23 (0.22) 4.97 - 7.59	6.11 (0.22) 4.88 - 7.50	(4.44 - 7.84) 4.54, 7.78	2.62	0.12 (0.064)	0.061
Carbohydrates by Calculation	68.30 (1.29) 61.33 - 75.86	67.62 (1.29) 61.92 - 75.27	(58.91 - 76.76) 59.21, 80.18	13.35	0.68 (0.31)	0.065
Protein	19.47 (0.96) 14.07 - 25.35	19.85 (0.96) 14.50 - 24.73	(13.50 - 26.62) 11.64, 25.13	10.23	-0.37 (0.24)	0.164
Total Fat	6.03 (0.50) 3.35 - 9.65	6.43 (0.50) 4.04 - 9.06	(2.74 - 9.74) 0.43, 11.11	5.02	-0.40 (0.14)	0.024
Acid Detergent Fiber	29.52 (0.73) 23.66 - 35.95	28.72 (0.73) 21.31 - 36.67	(19.97 - 44.27) 18.50, 40.91	15.35	0.79 (0.73)	0.282
Neutral Detergent Fiber	36.77 (1.18) 26.19 - 43.63	34.08 (1.18) 27.69 - 42.05	(25.71 - 52.96) 22.42, 50.76	14.36	2.69 (0.88)	0.018

[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 soybean varieties.
[4]With 95% confidence, interval contains 99% of the values expressed in the population of conventional reference soybean varieties. Negative limits set to zero.
[5]Maximum value minus minimum value for the control soybean variety.

Table 29. Literature and ILSI Database Ranges for Components in Soybean Forage and Seed

Seed Tissue Components[1]	Literature Range[2]	ILSI Range[3]
Seed Nutrients		
Proximates (% dw)		
Ash	4.61 – 6.32[a]; 4.32 – 5.88[b]	3.89 – 6.99
Carbohydrates by calculation	32.75 – 40.98[a]; 29.88 – 43.48[b]	29.6 – 50.2
Moisture (% fw)	6.24 – 12.10[a]; 5.44 – 11.70[b]	4.7 – 34.4
Protein	34.78 – 43.35[a]; 32.29 – 42.66[b]	33.19 – 45.48
Total Fat	14.40 – 20.91[a]; 15.10 – 23.56[b]	8.10 – 23.56
Fiber (% dw)		
Acid Detergent Fiber	9.22 – 26.26[a]; 11.81 – 19.45[b]	7.81 – 18.61
Neutral Detergent Fiber	10.79 – 23.90[a]; 13.32 – 23.57[b]	8.53 – 21.25
Amino Acids (% dw)		
Alanine	1.62 – 1.89[a]; 1.43 – 1.93[b]	1.51 – 2.10
Arginine	2.57 – 3.34[a]; 2.15 – 3.05[b]	2.29 – 3.40
Aspartic acid	4.16 – 5.02[a]; 4.01 – 5.72[b]	3.81 – 5.12
Cystine/Cysteine	0.52 – 0.69[a]; 0.41 – 0.71[b]	0.37 – 0.81
Glutamic acid	6.52 – 8.19[a]; 5.49 – 8.72[b]	5.84 – 8.20
Glycine	1.59 – 1.90[a]; 1.41 – 1.99[b]	1.46 – 2.00
Histidine	0.96 – 1.13[a]; 0.86 – 1.24[b]	0.88 – 1.18
Isoleucine	1.59 – 2.00[a]; 1.41 – 2.02[b]	1.54 – 2.08
Leucine	2.79 – 3.42[a]; 2.39 – 3.32[b]	2.59 – 3.62
Lysine	2.36 – 2.77[a]; 2.19 – 3.15[b]	2.29 – 2.84
Methionine	0.45 – 0.63[a]; 0.39 – 0.65[b]	0.43 – 0.68
Phenylalanine	1.82 – 2.29[a]; 1.62 – 2.44[b]	1.63 – 2.35
Proline	1.83 – 2.23[a]; 1.63 – 2.25[b]	1.69 – 2.28
Serine	1.95 – 2.42[a]; 1.51 – 2.30[b]	1.11 – 2.48
Threonine	1.44 – 1.71[a]; 1.23 – 1.74[b]	1.14 – 1.86
Tryptophan	0.30 – 0.48[a]; 0.41 – 0.56[b]	0.36 – 0.50
Tyrosine	1.27 – 1.53[a]; 0.74 – 1.31[b]	1.02 – 1.61
Valine	1.68 – 2.11[a]; 1.50 – 2.13[b]	1.60 – 2.20
Fatty Acids (% total FA)		
8:0 Caprylic	not available	0.148 – 0.148
10:0 Capric	0.15 – 0.27[b]	not available
12:0 Lauric	not available	0.082 – 0.132
14:0 Myristic	0.063 – 0.11[b]	0.071 – 0.238
14:1 Myristoleic	not available	0.121 – 0.125
15:0 Pentadecanoic	not available	not available
15:1 Pentadecenoic	not available	not available
16:0 Palmitic	9.80 – 12.63[b]	9.55 – 15.77
16:1 Palmitoleic	0.055 – 0.14[b]	0.086 – 0.194
17:0 Heptadecanoic	0.076 – 0.13[b]	0.085 – 0.146
17:1 Heptadecenoic	0.019 – 0.064[b]	0.073 – 0.087
18:0 Stearic	3.21 – 5.63[b]	2.70 – 5.88
18:1 Oleic	16.69 – 35.16[b]	14.3 – 32.2
18:2 Linoleic	44.17 – 57.72[b]	42.3 – 58.8

Table 29. Literature and ILSI Database Ranges for Components in Soybean Forage and Seed (continued)

Seed Tissue Components[1]	Literature Range[2]	ILSI Range[3]
18:3 Gamma Linolenic	not available	not available
18:3 Linolenic	4.27 – 9.90[b]	3.00 – 12.52
20:0 Arachidic	0.35 – 0.57[b]	0.163 – 0.482
20:1 Eicosenoic	0.13 – 0.30[b]	0.140 – 0.350
20:2 Eicosadienoic	0.016 – 0.071[b]	0.077 – 0.245
20:3 Eicosatrienoic	not available	not available
20:4 Arachidonic	not available	not available
22:0 Behenic	0.35 – 0.65[b]	0.277 – 0.595
Vitamins		
Vitamin E (mg/100g dw)	1.29 – 4.80[a]; 1.12 – 8.08[b]	0.19 – 6.17
Vitamin K (µg/g fw)	0.35-0.47[c]	not available
Minerals[4]		
Calcium	0.20-0.22[c], 0.24-0.41[d]	0.12-0.31
Phosphorus	0.48-0.64[c], 0.40-0.61[d]	0.50-0.94
Seed Anti-Nutrients		
Lectin (H.U./mg fw)	0.45 – 10.87[a]; 0.090 – 11.18[b]	0.105-9.038
Trypsin Inhibitor (TIU/mg dw)	20.79 – 59.03[a]; 18.14 – 42.51[b]	19.59 – 118.68
Phytic Acid (% dw)	0.41 – 1.92[a]; 0.81 – 2.66[b]	0.63 – 1.96
Raffinose (% dw)	0.26 – 0.84[a]; 0.43 – 1.85[b]	0.21 – 0.66
Stachyose (% dw)	1.53 – 3.04[a]; 1.97 – 6.65[b]	1.21 – 3.50
Isoflavones	**(µg/g dw)**	**(mg/kg dw)**
Daidzein	224.03 – 1571.91[a]; 198.95 – 1458.24[b]	60.0 – 2453.5
Genistein	338.24 – 1488.89[a]; 148.06 – 1095.57[b]	144.3 – 2837.2
Glycitein	52.72 – 298.57[a]; 32.42 – 255.94[b]	15.3 – 310.0
Forage Tissue Components[1]	**Literature Range[2]**	**ILSI Range[3]**
Forage Nutrients		
Proximate (% dw)		
Ash	5.28 – 9.24[a]; 4.77 – 8.54[b]	6.72 – 10.78
Carbohydrates by calculation	62.25 – 72.30[a]; 60.61 – 77.26[b]	59.8 – 74.7
Moisture (% fw)	68.50 – 78.40[a]; 62.76 – 80.20[b]	73.5 – 81.6
Protein	16.48 – 24.29[a]; 12.68 – 23.76[b]	14.38 – 24.71
Total Fat	2.65 – 9.87[a]; 2.96 – 7.88[b]	1.30 – 5.13
Fiber (% dw)		
Acid Detergent Fiber	23.86 – 50.89[a]; 25.49 – 47.33[b]	not available
Neutral Detergent Fiber	19.61 – 43.70[a]; 30.96 – 64.19[b]	not available

[1]fw=fresh weight; dw=dry weight; H.U. = hemagglutinating unit; TIU = trypsin inhibitor unit.
[2]Literature range references; [a]Lundry et al. (2008); [b]Berman et al. (2009), [c]Souci et al. (2008) (in fresh weight) (Accessed July 19, 2013), [d]Bellaloui et al. (2011).
[3]ILSI range is from ILSI Crop Composition Database, 2011 (Accessed August 14, 2013) (ILSI, 2011).
[4]Units for minerals are g/100g fw for data obtained from Souci et al. (2008) and in g/100g dw for data obtained from Bellaloui et al. (2011).

Compositional Assessment of MON 87751: Summary and Conclusion

Compositional analysis was conducted on seed and forage of MON 87751 and a conventional control grown at eight sites in the United States during a 2012 field production. The compositional analysis, based on the OECD consensus document for soybean, also included measurement of nutrients, anti-nutrients and other components in all varieties, including the conventional reference soybean varieties, to provide data on the natural variability of each compositional component analyzed.

Of the 50 components statistically assessed for MON 87751 only eight components (protein, glycine, proline, phosphorus, vitamin E, and raffinose in seed, and total fat and NDF in forage) showed a significant difference between MON 87751 and the conventional control. For these eight components, the mean difference in component values between MON 87751 and the conventional control was less than the range of the conventional control values and the reference variety values. The MON 87751 mean component values were within the 99% tolerance interval, the values observed in the literature, or the ILSI-CCDB values.

These results support the overall conclusion that MON 87751 was not a major contributor to variation in component levels in soybean seed and forage and confirmed the compositional equivalence of MON 87751 to the conventional control in levels of these components. These data indicated that the components with statistically significant differences were not compositionally meaningful from a food and feed safety perspective.

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

There is no compositional change shown in MON 87751 as described in section B5(a).

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 87751 are as safe and nutritious as those derived from commercially-available, conventional soybean for which there is an established history of safe consumption. Therefore, animal feeding studies do not add value to the safety of MON 87751.

Part 3 STATUTORY DECLARATION – AUSTRALIA

I, [redacted], 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 at Monsanto Australia, Level 12 / 600 St Kilda Road, Melbourne VIC 3004 on ..... of February 2015.

Declared before me ……

[Signature]

.....

.....

.....

[Full name, qualification and address of person before whom the declaration is made]

Part 4 REFERENCES

Akond, G.M., B. Ragin, R. Bazzelle, S.K. Kantartzi, K. Meksem and M.A. Kassem. 2012. Quantitative trait loci associated with moisture, protein, and oil content in soybean [*Glycine max* (L.) Merr.]. Journal of Agricultural Science 4: 16-25.

An, Y.-Q., J.M. McDowell, S. Huang, E.C. McKinney, S. Chambliss and R.B. Meagher. 1996. Strong, constitutive expression of the *Arabidopsis ACT2/ACT8* actin subclass in vegetative tissues. The Plant Journal 10: 107-121.

ASA. 2014. SoyStats 2014. American Soybean Association, Chesterfield, Missouri.

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

Axelos, M., C. Bardet, T. Liboz, A. Le Van Thai, C. Curie and B. Lescure. 1989. The gene family encoding the *Arabidopsis thaliana* translation elongation factor EF-1α molecular cloning characterization and expression. Molecular and General Genetics 219: 106-112.

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.

Baum, J.A. 1998. Transgenic *Bacillus thuringiensis*. Phytoprotection 79: 127-130.

Baum, J.A., T.B. Johnson and B.C. Carlton. 1999. *Bacillus thuringiensis*: Natural and recombinant bioinsecticide products. Pages 189-209 in Methods in Biotechnology: Biopesticides: Use and Delivery. Volume 5. F.R. Hall and J.J. Menn (eds.). Humana Press Inc., Totowa, New Jersey.

Bäumlein, H., W. Boerjan, I. Nagy, R. Bassüner, M. Van Montagu, D. Inzé and U. Wobus. 1991. A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and *Arabidopsis* plants. Molecular and General Genetics 225: 459-467.

Bautista-Zapanta, J.-n., K. Suzuki and K. Yoshida. 2002. Characterization of four ribosomal RNA operons in the genome of *Agrobacterium tumefaciens* MAFF301001. Nucleic Acids Research S2: 91-92.

Beck, E., G. Ludwig, E.A. Auerswald, B. Reiss and H. Schaller. 1982. Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn*5*. Gene 19: 327-336.

Bellaloui, N., J.R. Smith, A.M. Gillen and J.D. Ray. 2011. Effects of maturity, genotypic background, and temperature on seed mineral composition in near-isogenic soybean lines in the early soybean production system. Crop Science 51: 1161-1171.

Berman, K.H., G.G. Harrigan, S.G. Riordan, M.A. Nemeth, C. Hanson, M. Smith, R. Sorbet, E. Zhu and W.P. Ridley. 2009. Compositions of seed, forage, and processed fractions from

insect-protected soybean MON 87701 are equivalent to those of conventional soybean. Journal of Agricultural and Food Chemistry 57: 11360-11369.

Betz, F.S., B.G. Hammond and R.L. Fuchs. 2000. Safety and advantages of *Bacillus thuringiensis*-protected plants to control insect pests. Regulatory Toxicology and Pharmacology 32: 156-173.

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.

Bravo, A., S.S. Gill and M. Soberón. 2007. Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. Toxicon 49: 423-435.

Brookes, G. and P. Barfoot. 2012. GM crops: Global socio-economic and environmental impacts 1996-2010. PG Economics Ltd, Dorchester, United Kingdom.

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.

Cannon, R.J.C. 1993. Prospects and progress for *Bacillus thuringiensis*-based pesticides. Pesticide Science 37: 331-335.

Caviness, C.E. 1966. Estimates of natural cross-pollination in Jackson soybeans in Arkansas. Crop Science 6: 211-212.

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

Coruzzi, G., R. Broglie, C. Edwards and N.-H. Chua. 1984. Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1, 5-bisphosphate carboxylase. EMBO Journal 3: 1671-1679.

Crickmore, N. 2012. List of *Bacillus thuringiensis* holotype toxins. University of Sussex, Sussex, United Kingdom. http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/holo2.html [Accessed June 26, 2013].

De Almeida, E.R.P., V. Gossele, C.G. Muller, J. Dockx, A. Reynaerts, J. Botterman, E. Krebbers and M.P. Timko. 1989. Transgenic expression of two marker genes under the control of an *Arabidopsis rbcS* promoter: Sequences encoding the Rubisco transit peptide increase expression levels. Molecular and General Genetics 218: 78-86.

de Maagd, R.A., A. Bravo, C. Berry, N. Crickmore and H.E. Schnepf. 2003. Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. Annual Review of Genetics 37: 409-433.

de Maagd, R.A., A. Bravo and N. Crickmore. 2001. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends in Genetics 17: 193-199.

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.

Donovan, W.P. 1991. CryIIB crystal protein gene from *Bacillus thuringiensis*. Patent 5,073,632, U.S. Patent Office, Washington, D.C.

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.

EFSA. 2007. EFSA statement on the fate of recombinant DNA or proteins in meat, milk and eggs from animals fed with GM feed. European Food Safety Authority, Parma, Italy.

Eldridge, A.C. and W.F. Kwolek. 1983. Soybean isoflavones: Effect of environment and variety on composition. Journal of Agricultural and Food Chemistry 31: 394-396.

English, L. and S.L. Slatin. 1992. Mode of action of delta-endotoxins from *Bacillus thuringiensis*: A comparison with other bacterial toxins. Insect Biochemistry and Molecular Biology 22: 1-7.

FAO-WHO. 2001. Evaluation of allergenicity of genetically modified foods. Report of a joint FAO/WHO expert consultation on allergenicity of foods derived from biotechnology. Food and Agriculture Organization of the United Nations, Rome, Italy.

FARRP. 2011. Allergen database, version 11. University of Nebraska, Food Allergy Research and Resource Program, Lincoln, Nebraska. http://www.allergenonline.org/databasebrowse.shtml.

Fehily, A.M. 2003. Soy (soya) beans: Dietary importance. Pages 5392-5398 in Encyclopedia of Food Sciences and Nutrition. Second Edition. B. Caballero, L.C. Trugo, and P.M. Finglas (eds.). Academic Press, Amsterdam.

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.

FSANZ. 2008. Australia New Zealand Food Standards Code – Amendment No. 104 – 2008. Food Standards Australia New Zealand, Canberra, Australia.

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.

Gao, Y., K.J. Fencil, X. Xu, D.A. Schwedler, J.R. Gilbert and R.A. Herman. 2006. Purification and characterization of a chimeric Cry1F δ–endotoxin expressed in transgenic cotton plants. Journal of Agricultural and Food Chemistry 54: 829-835.

Gill, S.S., E.A. Cowles and P.V. Pietrantonio. 1992. The mode of action of *Bacillus thuringiensis* endotoxins. Annual Review of Entomology 37: 615-636.

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.

Gutierrez-Gonzalez, J.J., X. Wu, J. Zhang, J.-D. Lee, M. Ellersieck, J.G. Shannon, O. Yu, H.T. Nguyen and D.A. Sleper. 2009. Genetic control of soybean seed isoflavone content: Importance of statistical model and epistasis in complex traits. Theoretical and Applied Genetics 119: 1069-1083.

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., 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.

Hensley, K., E.J. Benaksas, R. Bolli, P. Comp, P. Grammas, L. Hamdheydari, S. Mou, Q.N. Pye, M.F. Stoddard, G. Wallis, K.S. Williamson, M. West, W.J. Wechter and R.A. Floyd. 2004. New perspectives on Vitamin E: γ–tocopherol and carboxyethylhydroxy chroman metabolites in biology and medicine. Free Radical Biology & Medicine 36: 1-15.

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.

Hoeck, J.A., W.R. Fehr, P.A. Murphy and G.A. Welke. 2000. Influence of genotype and environment on isoflavone contents of soybean. Crop Science 40: 48-51.

Höfte, H. and H.R. Whiteley. 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiological Reviews 53: 242-255.

Hunt, A.G. 1994. Messenger RNA 3' end formation in plants. Annual Review of Plant Physiology and Plant Molecular Biology 45: 47-60.

Hymowitz, T. 2004. Speciation and cytogenetics. Pages 97-136 in Soybeans: Improvement, Production, and Uses. Third Edition. H.R. Boerma and J.E. Specht (eds.). American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc., Madison, Wisconsin.

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/.

Jensen, O.N., A.V. Podtelejnikov and M. Mann. 1997. Identification of the components of simple protein mixtures by high-accuracy peptide mass mapping and database searching. Analytical Chemistry 69: 4741-4750.

Jonas, D.A., I. Elmadfa, K.-H. Engel, K.J. Heller, G. Kozianowski, A. König, D. Müller, J.F. Narbonne, W. Wackernagel and J. Kleiner. 2001. Safety considerations of DNA in food. Annals of Nutrition and Metabolism 45: 235-254.

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 Journal 5: 149-163.

Krebbers, E., J. Seurinck, L. Herdies, A.R. Cashmore and M.P. Timko. 1988. Four genes in two diverged subfamilies encode the ribulose-1,5-biphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. Plant Molecular Biology 11: 745-759.

Kumar, V., A. Rani, L. Goyal, A.K. Dixit, J.G. Manjaya, J. Dev and M. Swamy. 2010. Sucrose and raffinose family oligosaccharides (RFOs) in soybean seeds as influenced by genotype and growing location. Journal of Agricultural and Food Chemistry 58: 5081-5085.

Kumar, V., A. Rani, C. Tindwani and M. Jain. 2003. Lipoxygenase isozymes and trypsin inhibitor activities in soybean as influenced by growing location. Food Chemistry 83: 79-83.

Ladics, G.S., G.A. Bannon, A. Silvanovich and R.F. Cressman. 2007. Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search

for the elucidation of potential identities to known allergens. Molecular Nutrition and Food Research 51: 985-998.

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

Liu, H., A.T. Trieu, L.A. Blaylock and M.J. Harrison. 1998. Cloning and characterization of two phosphate transporters from *Medicago truncatula* roots: Regulation in response to phosphate and colonization by arbuscular mycorrhizal (AM) fungi. Molecular Plant-Microbe Interactions 11: 14-22.

Liu, K.S. 2004. Soybeans as a powerhouse of nutrients and phyhtochemicals. Pages 1-22 in Soybeans as Functional Foods and Ingredients. AOCS Press, Champaign, Illinois.

Lu, B.-R. 2004. Conserving biodiversity of soybean gene pool in the biotechnology era. Plant Species Biology 19: 115-125.

Lundry, D.R., W.P. Ridley, J.J. Meyer, S.G. Riordan, M.A. Nemeth, W.A. Trujillo, M.L. Breeze and R. Sorbet. 2008. Composition of grain, forage, and processed fractions from second-generation glyphosate-tolerant soybean, MON 89788, is equivalent to that of conventional soybean (*Glycine max* L.). Journal of Agricultural and Food Chemistry 56: 4611-4622.

Martinell, B., M. Petersen, D. Somers, Y. Wan, E. Williams and X. Ye. 2011. Methods for plant transformation using spectinomycin selection. Patent 8,030,544 B2, U.S. Patent Office, Washington, D.C.

McClintock, J.T., C.R. Schaffer and R.D. Sjoblad. 1995. A comparative review of the mammalian toxicity of *Bacillus thuringiensis*-based pesticides. Pesticide Science 45: 95-105.

Mendelsohn, M., J. Kough, Z. Vaituzis and K. Matthews. 2003. Are *Bt* crops safe? Nature Biotechology 21: 1003-1009.

Messina, M.J. 1999. Legumes and soybeans: Overview of their nutritional profiles and health effects. American Journal of Clinical Nutrition 70: 439S-450S.

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.

Morrison, M.J., E.R. Cober, M.F. Saleem, N.B. McLaughlin, J. Frégeau-Reid, B.L. Ma, W. Yan and L. Woodrow. 2008. Changes in isoflavone concentration with 58 years of genetic improvement of short-season soybean cultivars in Canada. Crop Science 48: 2201-2208.

OECD. 1993. Traditional crop breeding practices: An historical review to serve as a baseline for assessing the role of modern biotechnology. Organisation for Economic Co-operation and Development, Paris, France.

OECD. 2000. Consensus document on the biology of *Glycine max* (L.) Merr. (Soybean). ENV/JM/MONO(2000)9. Series on Harmonization of Regulatory Oversight in Biotechnology No.15. Organisation for Economic Co-operation and Development, Paris, France.

OECD. 2001. Consensus document on compositional considerations for new varieties of soybean: Key food and feed nutrients and anti-nutrients. ENV/JM/MONO(2001)15. Series on the Safety of Novel Foods and Feeds No.2. Organisation for Economic Co-operation and Development, Paris, France.

OECD. 2012. Revised consensus document on compositional considerations for new varieties of soybean [*Glycine max* (L.) Merr.]: Key food and feed nutrients, anti-nutrients, toxicants and allergens. ENV/JM/MONO(2012)24. Series on the Safety of Novel Foods and Feeds No. 25. Organisation for Economic Co-operation and Development, Paris, France.

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.

Palmer, R.G., K.E. Newhouse, R.A. Graybosch and X. Delannay. 1987. Chromosome structure of the wild soybean. Journal of Heredity 78: 243-247.

Panthee, D.R., V.R. Pantalone, D.R. West, A.M. Saxton and C.E. Sams. 2005. Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. Crop Science 45: 2015-2022.

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.

Pedersen, P. 2004. Soybean growth and development. Iowa State University, Ames, Iowa.

Piper, K.R., S.B. von Bodman, I. Hwang and S.K. Farrand. 1999. Hierarchical gene regulatory systems arising from fortuitous gene associations: Controlling quorum sensing by the opine regulon in *Agrobacterium*. Molecular Microbiology 32: 1077-1089.

Pull, S.P., S.G. Pueppke, T. Hymowitz and J.H. Orf. 1978. Soybean lines lacking the 120,000-dalton seed lectin. Science 200: 1277-1279.

Qin, G., E.R. ter Elst, M.W. Bosch and A.F.B. van der Poel. 1996. Thermal processing of whole soya beans: Studies on the inactivation of antinutritional factors and effects on ileal digestibility in piglets. Animal Feed Science and Technology 57: 313-324.

Raboy, V. and D.B. Dickinson. 1993. Phytic acid levels in seeds of *Glycine max* and *G. soja* as influenced by phosphorus status. Crop Science 33: 1300-1305.

Raboy, V., D.B. Dickinson and F.E. Below. 1984. Variation in seed total phosphorus, phytic acid, zinc, calcium, magnesium, and protein among lines of *Glycine max* and *G. soja*. Crop Science 24: 431-434.

Rademacher, T.W., R.B. Parekh and R.A. Dwek. 1988. Glycobiology. Annual Review of Biochemistry 57: 785-838.

Ray, J.D., T.C. Kilen, C.A. Abel and R.L. Paris. 2003. Soybean natural cross-pollination rates under field conditions. Environmental Biosafety Research 2: 133-138.

Richins, R.D., H.B. Scholthof and R.J. Shepherd. 1987. Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Research 15: 8451-8466.

Richter, S. and G.K. Lamppa. 1998. A chloroplast processing enzyme functions as the general stromal processing peptidase. Proceedings of the National Academy of Sciences of the United States of America 95: 7463-7468.

Rogers, S.G. 2000. Promoter for transgenic plants. Patent 6,018,100, U.S. Patent Office, Washington, D.C.

Rotundo, J.L. and M.E. Westgate. 2009. Meta-analysis of environmental effects on soybean seed composition. Field Crops Research 110: 147-156.

Rukmini, V., C.Y. Reddy and G. Venkateswerlu. 2000. *Bacillus thuringiensis* crystal δ-endotoxin: Role of proteases in the conversion of protoxin to toxin. Biochimie 82: 109-116.

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.

Schafer, F.Q., E.E. Kelley and G.R. Buettner. 2003. Oxidative stress and antioxidant intervention. Pages 849-869 in Critical Reviews of Oxidative Stress and Aging: Advances in Basic Science, Diagnostics and Intervention. R.G. Cutler and H. Rodriguez (eds.). World Scientific, New Jersey.

Schnepf, E., N. Crickmore, J. van Rie, D. Lereclus, J. Baum, J. Feitelson, D.R. Zeigler and D.H. Dean. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiology and Molecular Biology Reviews 62: 775-806.

Seguin, P., G. Tremblay, D. Pageau and W. Liu. 2010. Soybean tocopherol concentrations are affected by crop management. Journal of Agricultural and Food Chemistry 58: 5495-5501.

Seguin, P., P. Turcotte, G. Tremblay, D. Pageau and W. Liu. 2009. Tocopherols concentration and stability in early maturing soybean genotypes. Agronomy Journal 101: 1153-1159.

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.

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.

Snyder, H.E. and L.A. Wilson. 2003. Soy (soya) beans: Processing for the food industry. Pages 5383-5389 in Encyclopedia of Food Sciences and Nutrition. B. Caballero, L.C. Trugo, and P.M. Finglas (eds.). Academic Press, Amsterdam.

Souci, S.W., W. Fachmann and H. Kraut. 2008. Food composition and nutrition tables. Medpharm online database. Result for soya bean *Glycine hyspida* Maxim. medpharm GmbH Scientific Publishers, Stuttgart, Germany. http://www.sfk-online.net/cgi-bin/sfkstart.mysql?language=english [Accessed July 19, 2013].

Speicher, K.D., N. Gorman and D.W. Speicher. 2009. N-terminal sequence analysis of proteins and peptides. Pages 11.10.11-11.10.31 in Current Protocols in Protein Science. John Wiley and Sons, Inc., Hoboken, New Jersey.

Sutcliffe, J.G. 1979. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symposia on Quantitative Biology 43: 77-90.

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.

U.S. EPA. 1988. Guidance for the reregistration of pesticide products containing *Bacillus thuringiensis* as the active ingredient. 540/RS-89-023. U.S. Environmental Protection Agency, Washington, D.C.

U.S. EPA. 2001. Biopesticides registration action document - *Bacillus thuringiensis* plant-incorporated protectants. U.S. Environmental Protection Agency, Washington, D.C. http://www.epa.gov/pesticides/biopesticides/pips/bt_brad.htm [Accessed July 24, 2013].

U.S. EPA. 2005. *Bacillus thuringiensi*s Cry3Bb1 protein and the genetic material necessary for its production (Vector ZMIR13L) in event MON 863 corn & *Bacillus thuringiensis* Cry1Ab delta-endotoxin and the genetic material necessary for its production in corn (006430, 006484) fact sheet. U.S. Environmental Protection Agency, Washington, D.C. http://www.epa.gov/pesticides/biopesticides/ingredients_keep/factsheets/factsheet_006430-006484.htm [Accessed September 23, 2013].

U.S. EPA. 2010. Biopesticides registration action document: *Bacillus thuringiensis* Cry1A.105 and Cry2Ab2 insecticidal proteins and the genetic material necessary for their production in corn [PC Codes 006515 (Cry2Ab2), 006514 (Cry1A.105)]. U.S. Environmental

Protection Agency, Office of Pesticide Programs, Biopesticides and Pollution Prevention Division, Washington D.C.

U.S. FDA. 1992. Statement of policy: Foods derived from new plant varieties. Federal Register 57: 22984-23005.

USDA-APHIS. 2008. Letter to Monsanto regarding determination of nonregulated status for MON 89034. U.S. Department of Agriculture, Animal and Plant Health Inspection Service, Washington, D.C.

USDA-APHIS. 2014. Letter to Monsanto regarding extended determination of nonregulated status for MON 87751 soy (*Glycine max*). U.S. Department of Agriculture, Animal and Plant Health Inspection Service, Washington, D.C.

Vachon, V., R. Laprade and J.-L. Schwartz. 2012. Current models of the mode of action of *Bacillus thuringiensis* insecticidal crystal proteins: A critical review. Journal of Invertebrate Pathology 111: 1-12.

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.

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.

Vollmann, J., H. Grausgruber, H. Wagentristl, H. Wohleser and P. Michele. 2003. Trypsin inhibitor activity of soybean as affected by genotype and fertilisation. Journal of the Science of Food and Agriculture 83: 1581-1586.

WHO. 1999. Microbial pest control agent: *Bacillus thuringiensis*. Environmental Health Criteria 217. World Health Organization, Geneva, Switzerland. http://www.inchem.org/documents/ehc/ehc/ehc217.htm [Accessed May 14, 2009].

Widner, W.R. and H.R. Whiteley. 1989. Two highly related insecticidal crystal proteins of *Bacillus thuringiensis* subsp. *kurstaki* possess different host range specificities. Journal of Bacteriology 171: 965-974.

Wilson, R.F. 2004. Seed composition. Pages 621-677 in Soybeans: Improvement, Production, and Uses. Third Edition. H.R. Boerma and J.E. Specht (eds.). American Society of Agronomy, Inc., Crop Science Society of America, Inc., Soil Science Society of America, Inc., Madison, Wisconsin.

Wong, E.Y., C.M. Hironaka and D.A. Fischhoff. 1992. *Arabidopsis thaliana* small subunit leader and transit peptide enhance the expression of *Bacillus thuringiensis* proteins in transgenic plants. Plant Molecular Biology 20: 81-93.

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.

Ye, X., E.J. Williams, J. Shen, S. Johnson, B. Lowe, S. Radke, S. Strickland, J.A. Esser, M.W. Petersen and L.A. Gilbertson. 2011. Enhanced production of single copy backbone-free transgenic plants in multiple crop species using binary vectors with a pRi replication origin in *Agrobacterium tumefaciens*. Transgenic Research 20: 773-786.

Yoshimura, Y., K. Matsuo and K. Yasuda. 2006. Gene flow from GM glyphosate-tolerant to conventional soybeans under field conditions in Japan. Environmental Biosafety Research 5: 169-173.

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.