FOOD STANDARDS
Australia New Zealand
Te Mana Kounga Kai – Ahitereiria me Aotearoa

21 June 2018
[50-18]

Supporting document 1

Safety assessment – Application A1154

Food derived from insect-protected cotton line MON88702

Executive summary

Background

A genetically modified (GM) cotton line with OECD Unique Identifier MON-887Ø2-4, hereafter referred to as MON88702, has been developed to be protected from piercing and sucking insects, in particular the Hemipteran cotton pests *Lygus hesperus* and *L. lineolaris*. This protection is achieved through expression of a modified *Bacillus thuringiensis* (Bt) gene *mCry51Aa2*, which encodes a novel Bt Cry protein.

In conducting a safety assessment of food derived from MON88702, a number of criteria have been addressed including: characterisation of the transferred genes including their origin, function and stability in the cotton genome; the nature of the introduced proteins and their potential to be either allergenic or toxic in humans; compositional analyses and any resultant changes in the whole food. This approach evaluates the intended and any unintended changes in the plant.

This safety assessment addresses only food safety and nutritional issues of the GM food. It therefore does not address:

- risks related to the environmental release of GM plants used in food production
- risks to animals that may consume feed derived from GM plants
- the safety of food derived from the non-GM (conventional) plant.

Food derived from the non-GM (conventional) plant with an accepted history of safe use is used as the benchmark for the comparative analysis.

History of use

The host organism is cultivated cotton (*Gossypium hirsutum* L.). Cotton is one of the oldest cultivated crops and is grown worldwide primarily as a fibre crop but also as a source of food products derived from the seed. Such products need to be highly processed because of the presence of natural toxicants (gossypol) and anti-nutrients (cyclopropenoid fatty acids) in unprocessed cottonseed. The main food products include cottonseed oil and linters.

Molecular characterisation

MON88702 contains one novel gene introduced via *Agrobacterium*-mediated transformation. The *mCry51Aa2* gene is a modified version of the endogenous gene derived from *B. thuringiensis* and encodes a truncated protein that also contains eight amino acid substitutions. The modified protein has been shown to be as effective an insecticide as the wild type protein. Detailed molecular analyses of MON88702 indicate that *mCry51Aa2* has been integrated at a single insertion site. Plasmid backbone analysis shows no extraneous sequences, including antibiotic resistance genes derived from the plasmid, were transferred to the MON88702 genome.

The introduced genetic elements and the expression of new proteins in MON88702 were shown by phenotypic analysis and molecular techniques to be stably inherited from one generation to the next across multiple generations. The pattern of inheritance supports the conclusion that the herbicide-tolerance traits occur within a single locus in the MON88702 genome and are inherited in accordance with Mendelian principles.

Characterisation and safety assessment of new substances

Newly expressed proteins

mCry51Aa2 was detected in all plant parts analysed from MON88702. Expression of the protein was highest in leaf tissue and lowest in pollen. The mean level of mCry51Aa2 in seed from MON88702 was approximately 130 µg/g dry weight, which corresponds to ~0.054% of total protein.

A range of characterisation studies confirmed the identity of the plant-expressed mCry51Aa2 and its equivalence with the corresponding protein produced in a bacterial expression system. The plant mCry51Aa2 protein had the expected molecular weight, immunoreactivity, lack of glycosylation, amino acid sequence and insecticidal activity.

A combination of bioinformatic and *in vitro* analyses of mCry51Aa2 confirmed the absence of significant similarity to known allergenic or toxin proteins and that it would be readily digested in the gut The expressed protein was also shown to be heat labile indicating that mCry51Aa2 would be readily degraded during the processing used to produce cottonseed oil and in cooking.

Compositional analyses

Detailed compositional analyses were done on seed from MON88702 and the control DP393 cultivar grown under normal agricultural conditions over five field-trial sites, located in traditional cotton growing regions of the USA. The analyses included proximates (protein, fat, ash, moisture, carbohydrates by calculation), fibre components, fatty acids, amino acids, minerals, vitamins and anti-nutrients. The levels of 56 of these key analytes in MON88702 were compared to those in the control and also to compositional data from a range of commercial non-GM cotton varieties available from the published literature, a publicly available database and from previous cotton applications reviewed by FSANZ.

Statistically significant differences were found between seeds from MON88702 and the control for nine of the analytes measured, however all differences were small in magnitude and were within the range established for existing commercial cotton varieties. Overall, the compositional data support the conclusion that there are no biologically significant differences in the levels of key constituents in seed from MON88702 compared to conventional cotton varieties available on the market.

Conclusion

No potential public health and safety concerns have been identified in the assessment of MON88702. On the basis of the data provided in the present Application, and other available information, food derived from MON88702 is considered to be as safe for human consumption as food derived from conventional cotton varieties.

Table of contents

Index of Figures

Index of Tables

List of Abbreviations

ABCA	Agricultural Biotechnology Council of Australia
APVMA	Australian Pesticides And Veterinary Medicines Authority
BC	backcrossed
BLOSUM	BLOcks SUbstitution Matrix
bp	base pairs
Bt	*Bacillus thuringiensis*
bw	body weight
CERA	Center for Environmental Risk Assessment
COMPARE	COMprehensive Protein Allergen REsource
Cry	crystalline protein
Da	dalton
DNA	deoxyribonucleic acid
DW	dry weight
ELISA	enzyme-linked immunosorbent assay
h	hour
ILSI	International Life Sciences Institute
INS	International numbering system for food additives
FASTA	Fast alignment search tool – all
FDA	Food and Drug Administration (USA)
FAO	Food and Agriculture Organization of the United Nations
FSANZ	Food Standards Australia New Zealand
g	gram

GM	genetically modified
Hsp	heat shock protein
kDa	kilodalton
LC	lethal concentration
LOD	limit of detection
MALDI-TOF MS	matrix assisted laser desorption ionization-time of flight mass spectrometry
mg	milligram
min	minute
ml	millilitre
NCBI	National Centre for Biotechnology Information
NCPA	National Cottonseed Products Association (USA)
ng	nanogram
NOAEL	no observed adverse effect level
OECD	Organisation for Economic Co-operation and Development
OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
RNA	ribose nucleic acid
SAS	Statistical Analysis Software
SDS	sodium dodecyl sulfate
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
T-DNA	transfer DNA
U	units
µg	microgram
USA	United States of America
USDA	United States Department of Agriculture
WHO	World Health Organisation

1 Introduction

FSANZ has received an Application from Monsanto Australia Limited to vary Schedule 26 in the *Australia New Zealand Food Standards Code* (the Code). The variation is to add food derived from the genetically modified (GM) insect-protected cotton line MON88702 (hereafter referred to as MON88702), with the OECD Unique Identifier MON-887Ø2-4. This line is protected from piercing and sucking insects belonging to the hemiptera and thysanoptera orders.

Protection from the piercing and sucking insect pests is achieved by expression of a modified Cry51Aa2 (designated mCry51Aa2) insecticidal crystalline (Cry) protein, encoded by the gene *mCry51Aa2* that was derived from the soil bacterium *Bacillus thuringiensis*. FSANZ has previously approved a large number of Applications where *B. thuringiensis* Cry proteins have been introduced into crops for insect-protection but this is the first Application based on the Cry51 protein.

Cotton lines containing the MON88702 transformation event have been granted a licence for a limited and controlled release in Australia by the OGTR in 2017 (DIR 147[1]) but initial commercial cultivation will be based in the United States of America (USA). Food from MON88702 cotton will therefore initially enter the Australian and New Zealand food supply via imported products.

2 History of use

2.1 Host organism

The information provided here has been summarised from more detailed reports published by the Organisation for Economic Cooperation and Development (OECD 2008), the Office of the Gene Technology Regulator (OGTR 2016), the Agricultural Biotechnology Council of Australia (ABCA 2012), the United States Department of Agriculture Foreign Agricultural Service (USDA 2001, 2016, 2017), and the National Cottonseed Products Association (NCPA 2002, NCPA unknown date). Numerical and statistical data have been sourced from the FAOSTAT website from the Food and Agriculture Organization of the United Nations (FAO 2017).

The host organism *Gossypium hirsutum* L. is the most highly cultivated cotton species, and currently accounts for 90% of global cotton production. Cotton is predominantly grown for the textile industry, where evidence of human use has been found in archaeological sites in Pakistan, Egypt and Mexico dating back 5000 years. After ginning to remove fibres for textile manufacturing, cottonseed is processed into four major products: cottonseed oil, meal, hulls, and linters. Processing of cottonseed typically yields (by weight): 16% oil, 45% meal, 26% hulls, and 9% linters, with 4% lost during processing. Only cottonseed oil and linters are utilised as food sources in Australia and New Zealand.

Cottonseed oil is commonly used for deep frying in restaurants and the fast food industry and can be used in margarines and salad oils. In Australia and New Zealand, cottonseed oil has been in the food supply for approximately 50-60 years and globally there is a long history of safe use, with commercialisation of the oil for human use beginning in the early 1900's. Cottonseed oil is imported into Australia and New Zealand to supplement domestic supply. In Australia, current domestic use of cottonseed oil is around 116 million tons per year.

Cotton linters are a source of pure cellulose that can be used for food purposes. In food,

[1] http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/dir147

cellulose from linters can be used to prepare the food additives microcrystalline cellulose (INS# 460) and sodium carboxylmethyl cellulose (INS# 466). These additives are used to perform different functions, for example as emulsifiers in low fat ice creams or as anticaking agents in packaged shredded cheese.

The cotton plant is known to produce toxic compounds. Gossypol is a terpenoid aldehyde found in the pigment glands of root, leaves, flower buds and seeds. It is highly toxic to non-ruminant mammals, birds, insects and microorganisms as it can interfere in lysine metabolism and mitochondrial function. Plants that produce low levels of gossypol or have no gland tissue are highly susceptible to disease and pest infestations thus this toxin is considered to play an important part in the plants natural defence system. Another group of compounds present in cotton seeds are the cyclopropenoid fatty acids, which are thought to deter insects. They are considered anti-nutrients because they can interfere in metabolism of saturated fats.

In cotton seed, the levels of gossypol range from 0.4-2.0% and the cyclopropenoid fatty acids range from 0.5-1%. During the processing of oil from the seeds, both compounds are inactivated resulting in minimal presence in human food products. Problems can arise when the partially processed cottonseed is used to make flour or meal for human foods, which is common in countries like Central America, India and Pakistan. Gland-less cotton that does not produce gossypol has been developed for these countries where cotton meal is used as a source of protein in the human diet.

The cotton cultivar ‘DP393’ was used as the parental variety for the genetic modification described in this Application, and thus is regarded as the near-isogenic line for the purposes of the comparative assessment. It is a non-transgenic, upland variety developed by Delta and Pineland Technology Holding Company (Bridge et al, 2005) and released for commercial cultivation in the USA in 2005.

2.2 Donor organisms

2.2.1 Bacillus thuringiensis

The DNA sequence encoding the Cry51Aa2 protein was originally isolated from the *B. thuringiensis* strain EG2934 (Baum et al, 2012; 2017). *B. thuringiensis* is not considered pathogenic to humans (WHO 1999; ABSA 2017; Raymond and Federici, 2017) although some strains have been co-isolated from foods associated with causing diarrhoea and an incidence has been reported of *B. thuringiensis* being isolated from a corneal ulcer, after use of a *B. thuringiensis* pesticide product (Samples and Buettner, 1983).

The presence of *B. thuringiensis* in foods is not considered unusual considering these bacteria are ubiquitous in the environment. Specific strains of these bacteria have been utilised as commercial microbial insecticides in agriculture and forestry since 1938 in France and 1961 in the USA; thus this organism has a long history of safe use (Nexter et al, 2002; CERA 2011). Since 2010, there have been approximately 180 biopesticide products registered in the USA and China, 120 registered in the European Union, 45 in Australia and 13 in New Zealand, with high usage in forestry and the organic farming industry (Kabaluk et al, 2010; APVMA 2017). The strains used in agricultural biopesticide preparations are unlikely to be the cause of foodborne diarrhoea, since they have been shown to not produce enterotoxins (Cho et al, 2015). In foods associated with diarrhoea, the *B. thuringiensis* bacteria were co-isolated with *B. cereus* strains, of which both were shown to contain genes for enterotoxins.

2.2.2 Other organisms

Genetic elements from several other organisms have been used in the genetic modification of MON88702 (refer to Table 1). These non-coding sequences are used to drive or enhance expression of the new gene. None of the sources of these genetic elements are associated with toxic or allergenic responses in humans. The genetic elements derived from plant pathogens (Figwort mosaic virus and Cassava vein mosaic virus) are not pathogenic in themselves and do not cause pathogenic symptoms in MON88702.

3 Molecular characterisation

Molecular characterisation is necessary to provide an understanding of the genetic material introduced into the host genome and helps to frame the subsequent parts of the safety assessment. The molecular characterisation addresses three main aspects:

- the transformation method together with a detailed description of the DNA sequences introduced to the host genome
- a characterisation of the inserted DNA, including any rearrangements that may have occurred as a consequence of the transformation
- the genetic stability of the inserted DNA and any accompanying expressed traits.

The Applicant has submitted the following unpublished studies for the molecular characterisation of MON88702.

Unpublished studies

Molecular characterization of insect protected cotton MON 88702 (2017) Report MSL0028391. Monsanto Company.

Bioinformatics evaluation of the DNA sequences flanking the 5′ and 3′ junctions of the MON 88702 insert: assessment of putative polypeptides utilizing the AD_2017, TOX_2017, and PRT_2017 Databases (2017) Report MSL0028798. Monsanto Company.

Bioinformatics evaluation of the transfer DNA insert in MON 88702 utilizing the AD_2017, TOX_2017, and PRT_2017 databases (2017) Report MSL0028694. Monsanto Company.

Segregation analysis of the T-DNA insert in insect-protected cotton MON 88702 across three generations (2016) Report MSL0027485. Monsanto Company.

Demonstration of the presence of Cry51Aa2.834_16 protein in Lygus cotton leaf samples across multiple generations of MON 88702 (2016) Report MSL0027352. Monsanto Company.

3.1 Transformation Method

In order to create MON88702, plasmid PV-GHIR508523 was transformed into the cotton variety DP393. Plasmid PV-GHIR508523 contains two T-DNA inserts and some backbone sequences (Figure 1). The first T-DNA (T-DNA I) encodes the mCry51Aa2 protein and the second (T-DNA II) contains a spectinomycin selectable marker gene (*aadA*).

The transformation method involved infection of meristem tissue, dissected from cotton embryos, with the disarmed *Agrobacterium* strain AB33 containing the PV-GHIR508523 plasmid (Chen et al, 2012). The meristem tissues were then placed onto selective media containing spectinomycin, carbencillin and cefotaxime. The spectinomycin enables selection of transformants, while the carbencillin and cefotaxime suppresses the growth of the *Agrobacterium*. The meristem tissues were then placed into media to encourage shoot development followed by growth on propagation plugs to encourage root development. Rooted plants (R0) with normal phenotype were transferred to soil and allowed to self-

pollinate and produce R1 seed. The production of the R1 seed allows for random segregation and independent assortment, to separate the two T-DNA regions. PCR screening was then used to identify seeds carrying only the T-DNA I region, allowing selection of plants that would not contain the antibiotic resistance gene located in the T-DNA II.

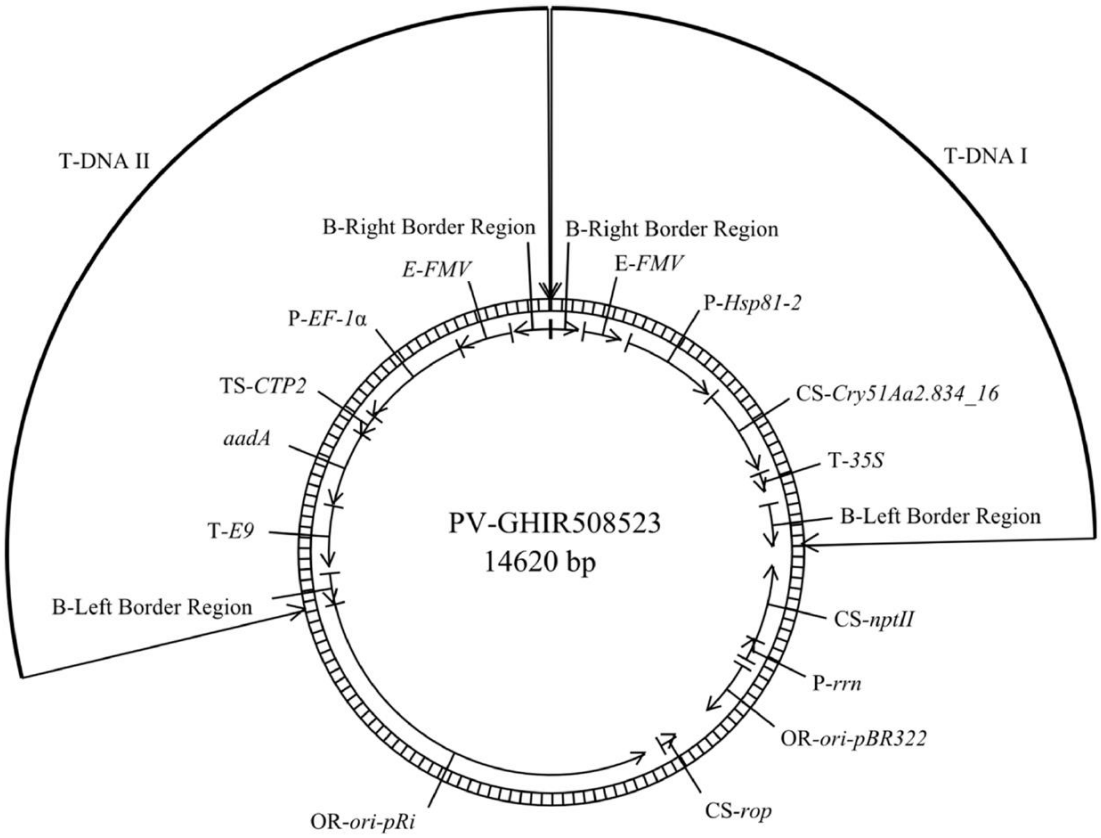

Figure 1: Plasmid map of PV-GHIR508523. The plasmid contains two T-DNA inserts. T-DNA I is carrying the mCry51Aa2 gene cassette and T-DNA II is carrying the selectable marker (aadA) gene cassette that is segregated out of the plant after selection of positive transformants. The vector backbone, which is not incorporated into the plant, is required for preparing the plasmid, passaging through standard laboratory Escherichia coli and into the Agrobacterium.

3.2 Detailed description of T-DNA I

The plasmid PV-GHIR508523 (Figure 1), used to generate MON88702, contains the *mCry51Aa2* gene cassette in T-DNA I (Figure 2). Expression of *mCry51Aa2* is under the control of the heat shock protein 81-2 promoter (P-*Hsp81-2*) from *Arabidopsis thaliana* and the 35S RNA terminator sequence (T) from Cauliflower mosaic virus. The protein coding sequence for the *mCry51Aa2* gene cassette encodes a single polypeptide of 306 amino acids. An enhancer sequence from the 35S RNA from Figwort mosaic virus (E-*FMV*) is located 5' to the promoter, to enhance transcription (Rogers, 2000).

Figure 2: The mCry51Aa2 gene cassette in PV-GHIR508523.

There are intervening sequences present in the T-DNA region as outlined in Table 2. These sequences assist with cloning, mapping and sequence analysis.

Table 2: The genetic elements contained in the T-DNA I region of pPV-GHIR508523, used to create MON88702.

Genetic element	Relative position	Source	Description, Function & Reference
Right border region	1 - 285	*Agrobacterium tumefaciens*	Right border sequence used to transfer the T-DNA region to the host (Depicker et al, 1982; Zambryski et al, 1988)
Intervening sequence	286 - 338	synthetic	
E-*FMV*	339 - 745	*Figwort mosaic virus*	Enhancer from the 35S RNA (Richens et al, 1987) that enhances transcription in most plant cells (Rogers, 2000)
Intervening sequence	746 - 820	synthetic	
P-*Hsp81-2*	821 - 1828	*Arabidopsis thaliana*	Promoter and 5′UTR leader sequence for the heat shock protein 81-2 that directs transcription in plant cells (Yabe et al, 1994)
Intervening sequence	1829 - 1865	synthetic	
Cry51Aa2.834_16	1866 - 2786	*Bacillus thuringiensis*	Coding sequence of the modified Cry51Aa2 protein that provides insect resistance (Baum et al, 2012; Anderson et al, 2015; Gowda et al, 2016)
Intervening sequence	2787 - 2818	synthetic	
T-*35S*	2819 - 3018	*Cauliflower mosaic virus*	3´ untranslated region of the 35S RNA gene containing the termination sequence (Mogen et al, 1990)
Intervening sequence	3019 - 3156	synthetic	
Left Border region	3157 - 3598	*Agrobacterium tumefaciens*	Left border sequence used to transfer the T-DNA region to the host (Barker et al, 1983)

3.3 Development of the cotton line from original transformation

A breeding programme was undertaken for the purposes of obtaining generations suitable for characterising MON88702 .The breeding pedigree for the various generations is given in Figure 3.

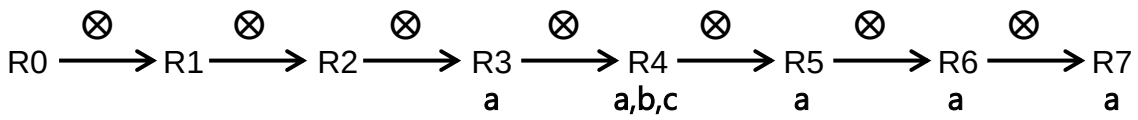

Figure 3: Breeding path used in the characterisation of the MON88702 line. R0 corresponds to the transformed plant, where PV-GHIR508523 was introduced into DP393. The R1 plant used to propagate the subsequent generations was identified by PCR to contain only the T-DNA I sequences. ⊗ designates self-pollination; (a) generations used to confirm insert stability; (b) generation used for molecular characterisation; (c) generation used for breeding commercial varieties containing the MON88702 event.

3.4 Characterisation of the inserted DNA and site(s) of insertion

A range of analyses were undertaken to characterise the genetic modification in MON88702. These analyses focussed on the nature and stability of the insertion and whether any unintended re-arrangements or products may have occurred as a consequence of the transformation procedure. When characterising MON88702, different generations of plants were analysed and these are outlined in Figure 3.

3.4.1 Identifying the number of integration sites

Next-generation sequence analysis was performed on seed-derived genomic DNA from MON88702 and the parental DP393 cultivar. A reference sequence was generated using plasmid PV-GHIR508523 and the parental DNA spiked with PV-GHIR508523. After preparation of a paired-end library using ~400 bp lengths of sheared genomic DNA, the samples were sequenced using Illumina HiSeq technology. Sufficient sequence fragments were obtained to cover the entire genomes of MON88702 and DP393, with a depth of coverage > 80x. Comparison of the sequence between the DP393 and MON88702 showed that a single integration event has occurred, with only two junction sites detected.

3.4.2 Detection of backbone sequence

Next-generation sequencing of the seed-derived DNA from MON88702 and the parental DP393 cultivar resulted in zero reads that mapped to the backbone sequences: OR-*ori-pRi*, CS-*nptII* or P-*rrn*; shown in Figure 1. A low level of detection was seen for the OR-*ori-pBR322* and CS-*rop* sequences in both the parental control and MON88702. As these sequences appeared in both samples at a similar level, their presence was most likely caused by contamination from environmental bacteria as has been previously reported by Yang et al (2013) and Zastrow-Hayes et al (2015). Furthermore, no reads from MON88702 or DP393 mapped to the promoter (P-*EF-1*α), transit peptide (TS-*CTP2*), aadA gene or terminator (T-*E9*) sequences contained in the T-DNA II region. As there is duplication of the Figwort mosaic virus enhancer (E-*FMV*) in both T-DNA I and T-DNA II, this sequence was detected in MON88702, but the surrounding sequences indicate that it was only the T-DNA I enhancer sequence present in MON88702.

3.4.3 Inheritance and genetic stability of the inserted DNA

A variety of methods were used to show genetic stability and inheritance of the inserted DNA in MON88702. First, next-generation sequencing of the seed-derived DNA from five generations of MON88702 (R3-R7) showed there was a single identical pair of junction sites present across all generations. This provided evidence of inheritance and genetic stability.

Secondly, immunoblotting was performed using leaf-derived protein from the five generations analysed by sequencing (generations R3-R7) and probed with a monoclonal antibody against mCry51Aa2. A *B. thuringiensis*-produced mCry51Aa2 protein was run as a positive control and leaf-derived protein from the parental control DP393 was run as a negative control. The results showed that all five generations expressed the mCry51Aa2 protein, confirming that the phenotype was inherited across generations.

Finally, chi-square (X^2) analysis was performed after breeding the R4 MON88702 with a null proprietary recurrent parental variety (as outlined in Figure 4). The inheritance pattern was assessed in the BC1F1, BC3F1 and BC2F2 generations by an endpoint PCR assay, using PCR primers targeting the T-DNA I region introduced into MON88702.

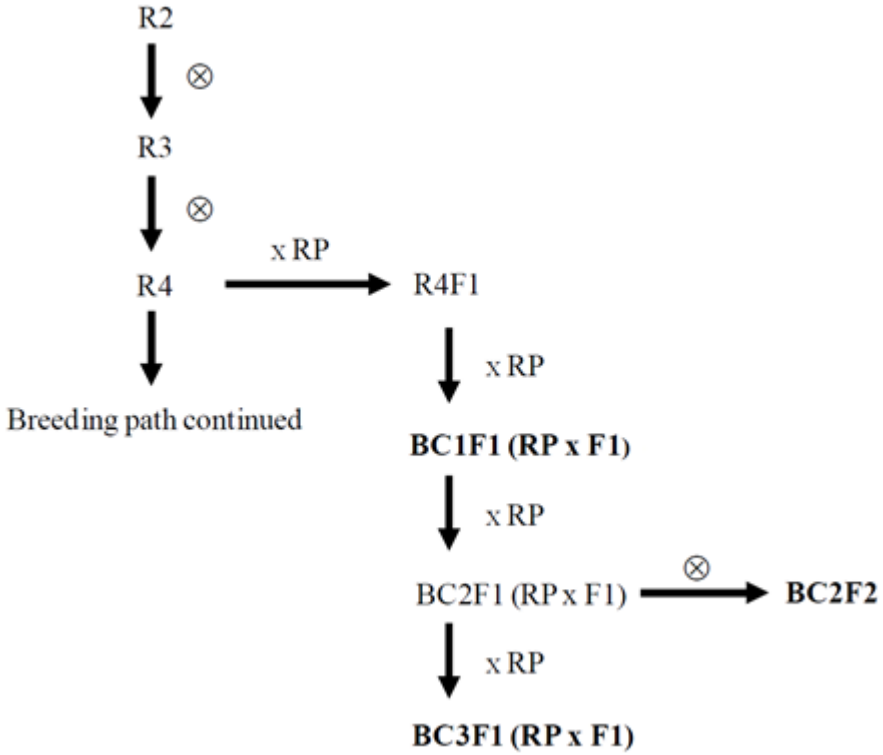

Figure 4: Breeding path used to assess the inheritance and genetic stability of MON88702.

The expected segregation ratio of 1:1 was observed, after crossing the hemizygous R4F1 plant with a null recurrent parent (Table 3). The expected segregation ratio of 1:2:1 was also observed after allowing the hemizygous plants at the BC2F1 generation to self-pollinate (Table 4). These data support the conclusion that the T-DNA is present at a single locus in MON88702 and was inherited predictably according to Mendelian principles in subsequent generations.

Table 3: Segregation results for MON88702 in BC1F1 and BC3F1

	BC1F1		BC3F1	
	Observed	**Expected**	**Observed**	**Expected**
Positive	52%	50%	49%	50%
Negative	48%	50%	51%	50%
X^2	0.30		0.09	
P	0.582		0.763	

Table 4: Segregation results for MON88702 in BC2F2

	BC2F2	
	Observed	**Expected**
homozygous	25%	25%
hemizygous	48%	50%
null	27%	25%
X^2	0.37	
P	0.832	

3.4.4 Insert integrity and site of integration

Next-generation sequencing of the seed-derived DNA from MON88702, using the plasmid DNA as a reference sequence, showed that a single copy of the DNA was integrated into the host genome. No deletions, insertions, mutations or rearrangements of the inserted DNA were detected. There was some truncation of the left and right border regions of the inserted DNA but this would not affect the expression of the *mCry51Aa2* gene.

Next-generation sequencing of both MON88702 and DP393 genomic DNA was also used to identify the site of integration. Based on the identified host genomic DNA sequence flanking the insert, a PCR was designed to produce two overlapping fragments, with primer pairs that bound to the flanking sequences and within the insert. Sequence analysis of the PCR fragments confirmed the insertion site and sequence of the inserted DNA.

3.4.5 Open reading frame (ORF) analysis

The DNA sequences of the insert and the two junctions between the host's genomic DNA and the insert, were translated into putative polypeptides across all six reading frames. ORFs of 8 or more amino acids were initially captured, identifying a total of 16 ORFs.

3.4.5.1 Bioinformatic analysis for potential allergenicity

The Applicant has provided the results of *in silico* analyses comparing the 16 amino acid sequences identified as potential proteins to known allergenic proteins listed in the Comprehensive Protein Allergen Resource (COMPARE[2]) database, from the Health and Environmental Science Institute. At the date of the search, there were 1,970 sequences in the allergen database. Three types of analyses were done:

(a) Full length sequence search – a FASTA alignment was performed comparing the whole sequence to the database entries. Significant homology was determined when there was more than 50% similarity between the query protein and database entry (BLOSUM50), with the E-value threshold set at 1×10^{-5} (1e-5).

(b) 80-mer sliding window search – a FASTA alignment was performed comparing all contiguous 80 amino acids within the ORF to the database entries. Matches were identified if there was greater than 35% homology.

(c) 8-mer exact match search – A FASTA alignment was performed comparing contiguous 8 amino acids within the ORF to the database entries. Matches were identified if there was 100% homology.

Of the 16 potential ORFs used to query the allergen database, no matches that met the criteria outlined above were found to any of the known allergenic proteins.

3.4.5.2 Bioinformatic analysis for potential toxicity

The Applicant provided results from *in silico* analyses comparing the 16 potential ORFs to known protein toxins identified in the NCBI protein database. A FASTA algorithm was used with a BLOSUM50 scoring matrix and the E-value threshold conservatively set to 1×10^{-5} (1e-5). Of the 16 ORFs identified, only one sequence that was in-frame with the mCry51Aa2 protein resulted in a match to a potential toxin, which is further discussed in Section 4.4.1. No other matches were found to any of the known proteins toxins.

[2] http://comparedatabase.org/database/

3.4.6 Conclusion

The data provided by the Applicant showed that a single integration event has occurred at a specific locus. The complete *mCry51Aa2* gene cassette has been inserted, without rearrangement, deletions or insertions. Furthermore, no backbone or T-DNA II sequences from the transforming plasmid were present, including antibiotic resistance genes. The introduced DNA was shown to be stably inherited from one generation to the next. Several ORFs encoding potential proteins were identified and were shown to have no homology to known allergens or unexpected homology to known toxins. Except for the mCry51Aa2 protein, it would be unlikely that the identified ORFs would be expressed due to the absence of required regulatory elements.

4 Characterisation and safety assessment of novel substances

In considering the safety of novel proteins it is important to consider that a large and diverse range of proteins are ingested as part of the normal human diet without any adverse effects. Only a small number of dietary proteins have the potential to impair health, because they have anti-nutrient properties or they can cause allergies in some consumers (Delaney et al. 2008). As proteins perform a wide variety of functions, different possible effects have to be considered during the safety assessment including potential toxic, anti-nutrient or allergenic effects.

To effectively identify any potential hazards, knowledge of the characteristics, concentration and localisation of all newly expressed proteins in the organism as well as a detailed understanding of their biochemical function and phenotypic effects is required. It is also important to determine if the newly expressed protein is expressed in the plant as expected, including whether any post-translational modifications have occurred.

The Applicant has submitted the following unpublished studies regarding the molecular characterisation of the novel proteins expressed in the MON88702.

Unpublished studies

- Characterization of the Cry51Aa2.834_16 protein purified from the cotton seed of MON 88702 and comparison of the physicochemical and functional properties of the plant-produced and Bt-produced Cry51Aa2.834_16 proteins (2016) Report MSL0027791. Monsanto Company.
- Assessment of Cry51Aa2.834_16 protein levels in cotton tissues collected from MON 88702 produced in United States field trials during 2015 (2016) Report MSL0027766. Monsanto Company.
- Bioinformatics evaluation of the mCry51Aa2 protein in MON 88702 utilizing the AD_2017, TOX_2017, and PRT_2017 databases (2017) Report MSL0028423. Monsanto Company.
- Assessment of the *in vitro* digestibility of Cry51Aa2.834_16 protein by pepsin and pancreatin (2017) Report MSL0028885. Monsanto Company.
- The effect of heat treatment on the functional activity of Cry51Aa2.834_16 protein (2016) Monsanto Company.
- An acute oral gavage toxicity study of *Bacillus thuringiensis* (Bt)-produced Cry51Aa2.834_16 protein in CD-1 mice (2017) Charles River Laboratories and Monsanto Company.

4.1 Description of the mCry51Aa2 protein

Cry51Aa2 is a member of the crystal (Cry) family of insecticidal proteins produced by *B. thuringiensis.* More specifically, it belongs to a group of aerolysin β-pore-forming toxins (Jerga et al, 2016). Cry proteins are contact insecticides, requiring ingestion by the target insect and passage into the midgut in order to function (Jurat-Fuentes and Crickmore, 2017). In the midgut, the Cry proteins are activated after cleavage by proteases, followed by binding of the active toxin to midgut epithelial cell receptors, allowing the formation of membrane pores (Bravo et al, 2004). Once the pores have formed, the mode of action leading to death is thought to follow one of two paths. As has been shown for members of the Cry proteins such as Cry1Ab, the pores enable movement of cations into cells, generating an osmotic imbalance and an associated influx of water, which causes the cells to swell and burst. Death of the insect is caused in part by the destruction of the midgut but also due to movement of bacteria through the damaged midgut lining into the hemocoel, resulting in septicaemia (Adang et al, 2014). An alternate mode of action associated with Cry34/35Ab1, Cry3Aa1 and Cry6Aa1 is that the cell lysis caused by the activated Cry proteins leads to occlusion of the midgut, which prevents the animal from feeding, leading to starvation (Bowling et al, 2017). The pathway to which Cry51Aa2 acts has not yet been determined.

The target specificity of the toxin is dependent on several factors, such as the proteases that activate the toxin and the epithelial cell receptor interaction that allows pore formation, which is unique for each organism (Jurat-Fuentes and Crickmore, 2017). Evidence shows that it would be highly unlikely Cry proteins that target a specific insect species would be active in humans and other non-target animals (Shimada et al, 2003; Bachman et al, 2017; Farmer et al, 2017; reviewed in van Frankenhuyzen, 2013).

The modified Cry protein expressed in MON88702 contains eight amino acid substitutions (F46S, Y54H, S95A, F147S, Q149E, S167R, P219R, R273W) and a deletion of three amino acids (Δ196-198) as outlined in Gowda et al (2016). The purpose of these changes are to improve the effectiveness of the protein to the major Hemipteran cotton pests *Lygus hesperus* and *L. lineolaris*, which have developed resistance to several classes of insecticides in the USA. Even with these changes the protein is 96% homologous to the wild type protein, found in *B. thuringiensis*. Although there are changes to the protein sequence, humans have been exposed to over a hundred different Cry proteins that have been identified in *B. thuringiensis*, each with their own primary structure but high similarities in secondary and tertiary structure, without adverse effects.

The *mCry51Aa2* gene prepared by the Applicant encodes a protein of 306 amino acids with a theoretical mass of 33594.5 Da.

4.2 Expression of the mCry51Aa2 protein in cotton tissues

Protein expression in plant tissues was determined by an enzyme-linked immunosorbent assay (ELISA). The ELISA system used was a validated in-house assay system, specific to the mCry51Aa2 protein being expressed in cotton tissues. A standard curve was generated using microbially expressed mCry51Aa2, of which the characterisation is described in section 4.3.

In order to determine the sites of accumulation of the mCry51Aa2 protein, samples were collected from MON88702 (generation R4) grown across five field-trial sites[3] during the 2015 growing season in the USA. For each tissue sample analysed, four samples were processed from each field-trial site. Specific tissues were collected at different growth stages. The

[3] Graham County, AZ; Rapides County, LA; Washington County, MS; Perquimans County, NC; San Patricio County, TX.

growth stage has been indicated by way of the phenological description previously outlined for cotton (Munger et al, 1998).

The results from the protein analysis (Table 5) showed the highest expression of mCry51Aa2 was in the leaf. This was demonstrated in leaf material collected at the leaf development (2-6 leaf) and inflorescence emergence (cut out) stages. The lowest expression was seen in pollen, seed and root tissue.

Table 5: Expression of mCry51Aa2 in plant tissue samples

Tissue	Growth Stage	mCry51Aa2 (µg/g DW)[2]		
		Mean	SD	Range
Leaf	2-6 leaf	1200	380	550-1700
Leaf	Cut out	1000	160	700-1300
Root	Peak bloom	190	41	150-290
Pollen[1]	Peak bloom	2.6	0.41	2.0-2.9
Seed	Maturity	130	17	91-170

1. Data for pollen was generated from a pooled sample of the four replicates collected per site. 2. DW - dry weight.

4.3 Characterisation of mCry51Aa2 protein expressed in cotton and equivalence to bacterially-produced forms

Protein was extracted from 100 g ground cotton seed (R4) and mCry51Aa2 was purified from the protein extract by affinity chromatography. Aliquots of purified protein were run on SDS-PAGE then visualised with Brilliant Blue G-Colloidal stain, and were shown to be pure and have an apparent molecular mass above 31 kDa. Using an *in silico*-based molecular weight analysis tool, the apparent molecular mass of the purified protein was calculated to be 34.2 kDa, which is in the expected size range of the theoretical protein. Western blot analysis showed that the protein being expressed was immunoreactive to a mCry51Aa2-specific antibody. N-terminal sequencing confirmed that the first 16 amino acids were as expected (Figure 5), although the N-terminal methionine residue was missing, likely due to cleavage. Peptide mapping also showed that the protein being expressed in cotton was mCry51Aa2, with 55% sequence coverage achieved (Figure 5). The protein was shown by an enhanced chemiluminescence glycoprotein detection procedure to be non-glycosylated. Finally, the functional activity of mCry51Aa2 was verified in a Western tarnished plant (WTP) bug (*L. herperus*) diet-incorporated insect bioassay. This bioassay confirms that the protein acts as an insecticide and allows determination of the concentration that would kill 50% (LC50) of the bugs (Bachman et al, 2017).

The mCry51Aa2 protein was also expressed in *B. thuringiensis*, transformed with a plasmid to express the modified protein. The bacterially-derived protein was shown to be pure by electrophoresis, had an apparent molecular mass of 34.7 kDa and was immunoreactive to mCry51Aa2 antibodies in a western blot. N-terminal sequencing and peptide mapping showed that the sequence matched the expected mCry51Aa2 sequence (as shown in Figure 5) with no N-terminal methionine. The protein was also shown to be non-glycosylated and was functional in the insect bioassay, with an equivalent LC50 to that obtained with the plant-produced form. This data further demonstrated that the bacterial-produced mCry51Aa2 is a

suitable standard for the ELISA used to detect mCry51Aa2 protein in plant tissues (as discussed in Section 4.1.1) and a suitable surrogate for use in the safety assessment experiments described in Section 4.4.

```
  1 -AILDLKSLV LNAINYWGPK NNNGIQGGDF GYPISEKQID
 41 TSIITSTHPR LIPHDLTIPQ NLETIFTTTQ VLTNNTDLQQ
 81 SQTVSFAKKT TTTTATSTTN GWTEGGKISD TLEEKVSVSI
121 PFIGEGGGKN STTIEANFAH NSSTTTSQEA STDIEWNISQ
161 PVLVPPRKQV VATLVIMGGN FTIPMDLMTT IDSTEHYSGY
201 PILTWISSPD NSYSGRFMSW YFANWPNLPS GFGPLNSDNT
241 VTYTGSVVSQ VSAGVYATVR FDQYDIHNLW TIEKTWYARH
281 ATLHNGKKIS INNVTEMAPT SPIKTN
```

Figure 5: Tryptic peptide map of the MON88702-produced mCry51Aa2 protein. The deduced amino acid sequence is 305 amino acids, with the expected N-terminal methionine residue missing. Boxed regions correspond to peptide sequence coverage achieved using MALDI-TOF MS. The italicised amino acids from position 2-16 were confirmed by N-terminal sequencing.

4.4 Safety of the introduced mCry51Aa2

4.4.1 Bioinformatic analyses of mCry51Aa2

The Applicant provided the results of *in silico* analyses comparing the mCry51Aa2 amino acid sequence to known allergenic proteins in the COMPARE dataset, using the same search criteria as outlined in Section 3.4.5.1. The search did not identify any known allergens with homology to mCry51Aa2.

The Applicant also provided the results of *in silico* analyses comparing the amino acid sequence of mCry51Aa2 to proteins identified as “toxins” from the NCBI protein databases. The search identified one potential toxin with a sequence overlap of greater than 50% but similarity of less than 30%. The toxin (GenBank Accession: APC96725), from a pathogenic proteobacteria, is also a member of the aerolysin family of pore-forming toxin proteins that includes the Cry51Aa2 protein (Moar et al, 2017). Structural similarity was restricted to a region in both proteins associated with pore-formation. Given Cry51Aa2 belongs to the same family of proteins, such homology is unremarkable and does not implicate the Cry51Aa2 protein as being toxic to humans.

4.4.2 Structural stability of mCry51Aa2 after exposure to heat

Bacterial-produced mCry51Aa2 protein that had previously been shown to be equivalent to plant-produced protein (Section 4.3) was boiled for 15 min at temperatures ranging from 25-95°C (Farmer et al, 2017). A control sample was kept at 4°C. An aliquot of the control and boiled protein samples were run on SDS-PAGE and stained with Brilliant Blue G Colloidal stain to detect the extent of protein degradation. No visible degradation or decrease in band intensity was observed in the control, 25 and 37°C treated samples, however there were observable lower molecular weight bands in the 55, 75 and 95°C treated samples and some decrease in the 34 kDa band intensity was observed in the 75 and 95°C treated samples (data summarised in Table 6). These data indicate that the mCry51Aa2 protein is heat labile at temperatures from 55°C and above.

To correlate the impact of heat treatment on functionality, the remaining boiled protein was

tested in the insect bioassay (Gowda et al, 2016; Bachman et al, 2017). The feeding studies showed there was no loss in function in samples exposed to 25 and 37°C when compared to the control but at 55°C and above, an LC50 could not be calculated because the protein was no longer functional (Farmer et al, 2017).

Table 6: Structural stability of mCry51Aa2 after exposure to heat

Treatment	LC50[1] (µg/ml)	95% CI[2] (µg/ml)	Protein Degradation
Control	2.557	1.289 – 3.799	nc[3]
25°C	1.164	0.556 – 1.905	nc
37°C	2.557	1.646 – 3.675	nc
55°C	> 60	NA	+[4]
75°C	> 60	NA	++
95°C	> 60	NA	++++

1. LC50 – concentration that causes 50% mortality. 2. CI – confidence interval. 3. nc – no change observed. 4. +/++++ - indicates the extent of protein degradation observed from minor to major.

4.4.3 Susceptibility of mCry51Aa2 to digestion with pepsin and pancreatin

Bacterial-produced mCry51Aa2 protein (test protein) was incubated with pepsin (10U enzyme/µg protein) at 37°C for 0-60 min, in a simulated gastric fluid (SGF) system at an acidic pH range (Thomas et al, 2004). Controls included a no enzyme control (test protein only) and a test protein control (no enzyme) incubated for 0 and 60 min. The extent of digestion was visualised by protein gel staining and western blotting. A serial dilution of the reaction mix (test protein plus enzyme) without incubation was used to determine the limit of detection (LOD) for the protein after gel staining and western blotting. In the protein gel staining analysis, 1 µg test protein was loaded per lane and the LOD was calculated to be 6.3 ng. In the western blotting experiments, 20 ng test protein was loaded and the LOD was 0.31 ng.

The results from the pepsin digestions showed that by 0.5 min, there was no visible mCry51Aa2 remaining in the reaction mix. There was no loss of band intensity in the no enzyme control incubated for 60 min therefore the loss of protein in the reaction mix indicated that mCry51Aa2 was being fully digested by pepsin (Farmer et al, 2017).

Bacterial-produced mCry51Aa2 protein (test protein) was also incubated with porcine pancreatin (~61U enzyme/µg protein) at 37°C for 0-24 h, in a simulated intestinal fluid (SIF) system at a neutral pH range. Pancreatin is a mixture of enzymes, including trypsin, chymotrypsin and endoproteases. Controls for this experiment included a no enzyme control (test protein only) and a test protein control (no enzyme) incubated for 0 and 24 h. The extent of digestion was visualised by western blotting. A serial dilution of the reaction mix (test protein plus enzyme) without incubation was used to determine the limit of detection (LOD) for the protein after western blotting. In the western blotting experiments, 20 ng test protein was loaded and the LOD was 0.63 ng.

The results from the SIF digestion showed that mCry51Aa2 was partially resistant to pancreatin digestion. Intact protein was present in the reaction mix up to 24 h, although band intensity decreased over time and the presence of degradation products increased over time. There was no loss of band intensity in the no enzyme control incubated for 24 h therefore the decreased band intensity seen in the reaction mix samples showed that protein loss was due to pancreatin activity rather than protein instability. A further study to assess the function of the pancreatin showed that the enzyme was active and the activity was in the expected range.

In the human digestive system, gastric digestion (pepsin) occurs before intestinal digestion (pancreatin) therefore a sequential digestion was performed on mCry51Aa2. The pepsin digestion was run for 0 and 2 min followed by digestion of the 2 min sample by pancreatin from 0-2 h. The results showed that by 2 min, mCry51Aa2 was completely digested by the pepsin therefore there was no protein remaining in the mix that could be digested by pancreatin.

As a further evaluation of digestibility of mCry51Aa2, an *in silico* analysis of potential cleavage sites was investigated by FSANZ using the amino acid sequence of mCry51Aa2 and the PeptideCutter[4] tool in the ExPASy Proteomics Site. mCry51Aa2 has multiple cleavage sites for pepsin (41 sites at pH 1.3 and 67 sites at pH >2), trypsin (18 sites), chymotrypsin (25 high-specificity sites, 52 low-specificity sites) and endopeptidases (25 sites). Combining this data with the pepsin digestibility assay shows that mCry51Aa2 is as susceptible to protein digestion as the vast majority of dietary proteins.

4.4.4 Acute toxicity studies

Although not required, since no toxicity concerns were raised in the data considered in Sections 4.4.1 – 4.4.3, the Applicant also provided the results from an acute oral toxicity study as additional supporting information. As would be expected given the information presented above no treatment-related deaths or clinical symptoms were observed during the study.

The results from this study indicate that the no observed adverse effect level (NOAEL) would be > 5000 mg/kg. This is similar to the NOAELs reported for other Cry proteins (summarised in Koch et al, 2015), which range from 576-5000 mg/kg.

4.4.5 Conclusion

A range of characterisation studies were performed on plant-derived mCry51Aa2 confirming the identity and functionality of the protein and equivalence to the corresponding protein produced in a bacterial expression system. Expression of mCry51Aa2 in MON88702 was highest in leaf and lowest in pollen and cottonseed. The protein was shown to be heat labile and susceptible to pepsin digestion. A bioinformatic search showed mCry51Aa2 had minimal homology to known toxins and allergens. Taken together this indicates the protein is unlikely to be toxic to humans. An acute toxicity study in mice further confirmed the absence of toxicity.

5 Compositional analysis

The main purpose of compositional analyses is to determine if, as a result of the genetic modification, unexpected changes had occurred to the food. These changes could take the form of alterations in the composition of the plant and its tissues and thus its nutritional adequacy. Compositional analyses can also be important for evaluating the intended effect where there has been a deliberate change to the composition of the food.

The classic approach to the compositional analyses of GM food is a targeted one. Rather than analysing every possible constituent, which would be impractical, the aim is to analyse only those constituents most relevant to the safety of the food or that may have an impact on the whole diet. Important analytes therefore include the key nutrients, toxicants and anti-nutrients for the food in question. The key nutrients and anti-nutrients are those components in a particular food that may have a substantial impact in the overall diet. They may be major constituents (fats, proteins, carbohydrates or enzyme inhibitors such as anti-nutrients) or

[4] web.expasy.org/peptide_cutter/

minor constituents (minerals, vitamins). Key toxicants are those toxicologically significant compounds known to be inherently present in an organism, such as compounds whose toxic potency and level may be significant to health.

5.1 Key Components

Cottonseed oil is the primary cotton product used for human consumption. The key components to be analysed for the comparison of transgenic and conventional cotton are outlined in the OECD Consensus Document on Compositional Considerations for New Varieties of Cotton, and include proximates and fibre (cottonseed only), fatty acids, tocopherol (vitamin E), gossypol and the cyclopropenoid fatty acids - malvalic, sterculic and dihydrosterculic acids (OECD, 2009).

5.2 Study design

Five field trials were conducted for MON88702 in the USA during the 2015 growing season[5]. The agronomic practices and pest control measures used were location-specific and were typical for all aspects of cotton cultivation including soil preparation, fertiliser application, irrigation and pesticide-based control methods. The materials tested in the field trials included MON88702 and the parental control (DP393).

The field trials were established in a randomised complete block design, with four replicates of each plot. Cottonseed samples from all plots were harvested at maturity, ginned and acid-delinted. Seed samples were then ground before shipment to EPL Bio Analytical Services (Niantic, IL) for compositional analyses. The methodologies for the compositional analyses were based on internationally recognised procedures including official methods specified by the Association of Official Analytical Chemists (AOAC), the USDA and published articles or technical notes from industrial-based sources.

A total of 56 different analytes were measured. Statistical analyses were performed using SAS 9.4 (SAS Institute, Cary, NC). For each analyte, 'descriptive statistics' (mean and standard error of the mean (SEM)) were generated. A mixed model analysis of variance was then applied to the combined data covering the five replicated field trial sites. In assessing the significance of any difference between the mean analyte value for MON88702 and the parental control, a P-value of 0.05 was used.

In order to complete the statistical analysis for any component in this study, it was deemed that more than 50% of the values must be greater than the assay limit of quantitation (LOQ). If analytes had more than 50% of observations below the LOQ for that assay, they were excluded from the overall summary analysis. Values for all components were expressed on a dry weight basis with the exception of fatty acids, expressed as percent of total fatty acids.

To determine if the compositional changes observed in MON88702 reflected the natural variation seen in cottonseed, results were compared to those reported in the published literature (Hamilton et al, 2004; Betrand et al, 2005; Harrison et al, 2013), from non-GM cotton reference line data provided to FSANZ from previous applications (A1147[6], A1094[7], A1080[8]) and from the ILSI Crop Compositional Database (v.6)[9]. This combined data is

[5] Graham County, AZ; Rapides County, LA; Washington County, MS; Perquimans County, NC; San Patricio County, TX.
[6] http://www.foodstandards.gov.au/code/applications/Pages/A1147.aspx
[7] http://www.foodstandards.gov.au/code/applications/Pages/A1094-GM-Cotton.aspx
[8] http://www.foodstandards.gov.au/code/applications/Pages/A1080-Food-derived-from-Herbicide-tolerant-Cotton-Line-MON88701.aspx
[9] https://www.cropcomposition.org

shown as the Reference Range (minimum – maximum) in the data tables.

Unpublished studies

Compositional analyses of cottonseed from MON 88702 grown in the United States during the 2015 season (2017) Report MSL0028439. Monsanto Company.

5.3 Analyses of key components in fuzzy seed

5.3.1 Proximates and fibre

Analysis of the proximate and fibre levels showed there was a statistically significant difference in protein levels between MON88702 and the parental control (Table 7). The protein level was well in the range found in other cotton cultivars, shown in the reference range, indicating this change was not considered biologically significant.

Table 7: Comparison of Proximates and Fibre (% DW)

Parameter	Non-GM counterpart	MON88702	p value	Reference range
	Mean (SEM)	Mean (SEM)		Min - Max
Ash	4.11 (0.18)	4.17 (0.18)	ns[1]	3.006 – 5.476
Carbohydrate	52.33 (1.51)	52.22 (1.51)	ns	39.04 – 61.0
Protein	24.67 (0.63)	24.09 (0.63)	0.038	12.0 – 33.0
Total fat	18.89 (1.64)	19.54 (1.64)	ns	13.7 – 27.9
Acid Detergent Fibre	34.75 (0.55)	35.40 (0.55)	ns	19.7 – 54.1
Neutral Detergent Fibre	42.56 (0.90)	43.03 (0.90)	ns	25.6 – 58.6
Total Dietary Fibre	40.05 (1.69)	40.95 (1.69)	ns	25.2 – 77.0

1. ns – not significant

5.3.2 Minerals and alpha tocopherol (vitamin E)

There was no difference in the level of Vitamin E or phosphorus in MON88702 compared to the parental control but there was a statistically significant difference seen with calcium (Table 8). Calcium is acquired by plants from the soil thus the difference seen in MON88702 is likely related to the differences in levels of calcium found in the soil at the different field trial sites and within each site. The level of calcium also falls within the reference range and is therefore not considered to be biologically significant.

Table 8: Comparison of Minerals (% DW) and Vitamin E (mg/kg DW)

Parameter	Non-GM counterpart	MON88702	p value	Reference range
	Mean (SEM)	Mean (SEM)		Min - Max
Calcium	0.11 (0.0047)	0.13 (0.0047)	0.049	0.046 – 0.33
Phosphorus	0.63 (0.046)	0.66 (0.046)	ns[1]	0.31 – 1.03
Vitamin E	120.26 (8.09)	122.94 (8.09)	ns	26.6 - 224

1. ns – not significant

5.3.3 Amino acids

There were no statistically significant differences found in the level of the 18 amino acids analysed in MON88702 compared to the parental control (Table 9).

Table 9: Comparison of Amino Acids (% DW)

Parameter	Non-GM counterpart	MON88702	Reference range
	Mean (SEM)	Mean (SEM)	Min - Max
Alanine	0.88 (0.024)	0.88 (0.024)	0.49 – 1.75
Arginine	2.69 (0.110)	2.58 (0.110)	1.30 – 5.01
Aspartic Acid	2.15 (0.076)	2.08 (0.076)	1.08 – 4.08
Cysteine	0.50 (0.019)	0.50 (0.019)	0.18 – 0.78
Glutamic Acid	4.82 (0.200)	4.67 (0.200)	2.43 – 7.55
Glycine	0.95 (0.028)	0.95 (0.028)	0.52 – 1.89
Histidine	0.68 (0.020)	0.67 (0.020)	0.35 – 1.28
Isoleucine	0.76 (0.023)	0.74 (0.023)	0.37 – 1.47
Leucine	1.43 (0.043)	1.41 (0.043)	0.72 – 2.67
Lysine	1.14 (0.031)	1.12 (0.031)	0.55 – 2.18
Methionine	0.36 (0.013)	0.36 (0.013)	0.15 – 0.75
Phenylalanine	1.34 (0.048)	1.31 (0.048)	0.65 – 2.38
Proline	0.88 (0.028)	0.87 (0.028)	0.46 – 1.75
Serine	0.91 (0.037)	0.92 (0.037)	0.50 – 1.94
Threonine	0.78 (0.020)	0.78 (0.020)	0.32 – 1.41
Tryptophan	0.27 (0.008)	0.27 (0.008)	0.11 – 0.54
Tyrosine	0.48 (0.018)	0.47 (0.018)	0.32 – 1.14
Valine	1.05 (0.031)	1.03 (0.031)	0.54 – 2.06

5.3.4 Fatty Acids

The following fatty acids were excluded from the statistical summary as they had more than 50% of observations below the LOQ: 8:0 caprylic acid, 10:0 capric acid, 14:1 myristoleic acid, 15:0 pentadecanoic acid, 15:1 pentadecenoic acid, 18:3 gamma linolenic acid, 20:3 eicosatrienoic acid and 20:4 arachidonic acid.

Statistically significant differences between MON88702 and DP393 were observed for a range of fatty acids (Table 10). In the more prominent fatty acids (> 1% of total fatty acids), the changes were very minor and do not significantly change the composition of the oil. As the mean values fall within the reference ranges, these differences are not considered biologically significant.

Table 10: Comparison of Fatty Acids (% Total Fatty Acids)

Parameter	Non-GM counterpart	MON88702	p value	Reference range
	Mean (SEM)	Mean (SEM)		Min - Max
18:2 Linoleic Acid	56.13 (1.90)	56.60 (1.90)	ns[1]	42.50 – 67.80
16:0 Palmitic Acid	20.78 (0.60)	19.21 (0.60)	< 0.001	15.11 – 28.10
18:1 Oleic Acid	17.34 (1.16)	18.45 (1.16)	0.020	9.20 – 25.40
18:0 Stearic Acid	2.50 (0.15)	2.66 (0.15)	0.003	0.20 – 3.54
14:0 Myristic Acid	0.71 (0.047)[2]	0.57 (0.047)	< 0.001	0.22 – 2.40
16:1 Palmitoleic Acid	0.53 (0.018)	0.55 (0.018)	ns	0.33 – 1.74
20:0 Arachidic Acid	0.27 (0.028)	0.26 (0.028)	< 0.001	0.12 – 0.48
18:3 Linolenic Acid	0.22 (0.0095)	0.22 (0.0095)	ns	0.05 – 0.64
22:0 Behenic Acid	0.13 (0.017)	0.12 (0.017)	0.002	0.07 – 0.30
17:0 Heptadecanoic Acid	0.079 (0.0028)	0.077 (0.0028)	ns	0.05 – 1.12
20:1 Eicosenoic Acid	0.066 (0.0031)	0.065 (0.0031)	ns	0.01 – 0.10
17:1 Heptadecenoic Acid	0.042 (0.00087)	0.040 (0.00088)	ns	0.005 – 0.094
20:2 Eicosadienoic acid	0.019 (0.00085)	0.018 (0.00085)	ns	na[2]
12:0 Lauric acid	0.017 (0.0024)	0.015 (0.0024)	0.009	na

1. ns – not significant. 2. na – data not available.

5.3.5 Anti-nutrients

As shown in Table 11, there were no statistically significant differences observed between the anti-nutrient levels in MON88702 and those in the parental control.

Table 11: Comparison of Anti-nutrients

Parameter	Non-GM counterpart	MON88702	Reference Range
	Mean (SE)	Mean (SE)	Min - Max
Gossypol (% DW)			
Free gossypol	0.62 (0.093)	0.61 (0.093)	0.384 – 1.418
Total gossypol	1.01 (0.12)	1.01 (0.12)	0.350 – 1.613
Cyclopropenoic fatty acids (% total fatty acids)			
Malvalic acid	0.55 (0.046)	0.52 (0.046)	0.112 – 0.854
Sterculic acid	0.24 (0.015)	0.24 (0.015)	0.061 – 0.556
Dihydrosterculic acid	0.37 (0.038)	0.38 (0.038)	0.031 – 0.325

5.4 Conclusion

Of the 56 analytes measured in cotton fuzzy seed, mean values were provided for 47 analytes. A summary of the nine analytes that showed a statistically significant difference between MON88702 and the parental line DP393 is provided in Table 12.

For the majority of analytes presented in Table 12, the differences were small in magnitude (within 10%) with the exception of myristic acid, where there was an approximate change of minus (-) 20%. Regardless of these changes, the mean values were well within the reference ranges reported in the published literature (Hamilton et al, 2004; Betrand et al, 2005; Harrison et al, 2013), from non-GM cotton reference line data provided to FSANZ from previous applications (A1147, A1094, A1080) and from the ILSI Crop Compositional

Database. As the composition of cotton can vary significantly due to the cultivation site and agricultural practices, the differences reported here most likely reflect the normal biological variability that exists in cotton.

Overall, the compositional data are consistent with the conclusion that there are no biologically significant differences in the levels of key constituents in MON88702 when compared with conventional cotton cultivars already available in agricultural markets.

Table 12: Summary of statistically significant compositional differences between MON88702 and the parental control DP393.

Parameter	Non-GM counterpart	MON88702	Are values within the reference ranges?
	Mean ± SEM	Mean ± SEM	Yes / No
Protein	24.67 (0.63)[1]	24.09 (0.63)	Yes
Calcium	0.11 (0.0047)[1]	0.13 (0.0047)	Yes
Palmitic Acid	20.78 (0.60)	19.23 (0.60)	Yes
Oleic Acid	17.34 (1.15)	18.46 (1.15)	Yes
Stearic Acid	2.50 (0.15)	2.67 (0.15)	Yes
Myristic Acid	0.71 (0.047)	0.57 (0.047)	Yes
Arachidic Acid	0.27 (0.028)	0.26 (0.028)	Yes
Behenic Acid	0.13 (0.017)	0.12 (0.017)	Yes
Lauric acid	0.017 (0.0024)	0.015 (0.0024)	na[3]

1. Cells highlighted in red show data where MON88702 is significantly lower than the parental DP393 line.
2. Cells highlighted in green show data where MON88702 is significantly higher than the parental DP393 line. 3. na – data not available.

6 Nutritional impact

In assessing the safety of a GM food, a key factor is the need to establish that the food is nutritionally adequate and will support typical growth and wellbeing. In most cases, this can be achieved through a detailed understanding of the genetic modification and its consequences, together with an extensive compositional analysis of the food, such as that presented in Section 5 of this report.

Where a GM food has been shown to be compositionally equivalent to conventional varieties, the evidence to date indicates that feeding studies using target livestock or other animal species will add little to the safety assessment (Bartholomaeus et al., 2013; OECD, 2003). If the compositional analysis indicates biologically significant changes, either intended or unintended, to the levels of certain nutrients in the GM food, additional nutritional studies should be undertaken to assess the potential impact of the changes on the whole diet.

MON88702 is the result of a simple genetic modification to confer resistance to insects, with no intention to significantly alter nutritional parameters in the food. The compositional analyses have demonstrated that the genetic modification has not altered the nutritional adequacy of MON88702 as a source of food when compared with that of conventional cotton varieties. The introduction of foods derived from MON88702 into the food supply is therefore expected to have negligible nutritional impact.

7 References

ABCA (2012) GM cotton in Australia: a resource guide. Resource guide 3, Agricultural Biotechnology Council of Australia. http://www.abca.com.au/wp-content/uploads/2012/ 09/ABCA_Resource_Guide _3_v2.pdf (accessed October 2017)

ABSA (2017) *Risk group database* hosted by the American Biological Safety Association International. https://my.absa.org/tiki-index.php?page=Riskgroups (accessed October 2017)

Adang MJ, Crickmore N, Jurat-Fuentes JL (2014) Diversity of *Bacillus thuringiensis* crystal toxins and mechanism of action. In: Dhadialla TS, Gill SS (eds) Insect midgut and insecticidal proteins, First edition. Academic Press, Amsterdam, pp 39-88

Anderson HM, Bowen DJ, Chay CA, Flasinski S, Kesanapalli UR, Milligan JS, Slightom RN, Yin Y (2015) Pesticidal toxin active against coleopteran and/or hemipteran insects. World Patent 2015021354 A3.

APVMA (2017) *Bacillus thuringiensis* products identified by searching the APVMA website for "evaluation" AND "thuringiensis" and recording those products that have been gazetted.

Bachman PM, Ahmad A, Ahrens JE, Akbar W, Baum JA, Brown S, Clark TL, Fridley JM, Gowda A, Greenplate JT, Jensen PD, Mueller GM, Odegaard ML, Tan J, Uffman JP, Levine SL (2017) Characterization of the activity spectrum of MON 88702 and the plant-incorporated protectant Cry51Aa2.834_16. PLoS. One. 12:e0169409

Barker RF, Idler KB, V TD, Kemp JD (1983) Nucleotide sequence of the T-DNA region from the Agrobacterium tumefaciens octopine Ti plasmid pTi15955. Plant Mol.Biol. 2:335-350

Bartholomaeus A, Parrott W, Bondy G, Walker K (2013) The use of whole food animal studies in the safety assessment of genetically modified crops: Limitations and recommendations. Crit Rev Toxicol 43(S2):1-24

Baum JA, Sukuru UR, Penn SR, Meyer SE, Subbarao S, Shi X, Flasinski S, Heck GR, Brown RS, Clark TL (2012) Cotton plants expressing a hemipteran-active *Bacillus thuringiensis* crystal protein impact the development and survival of *Lygus hesperus* (Hemiptera: Miridae) nymphs. J Econ Entomol 105:616-624

Baum JA, Flasinski S, Heck GR, Penn SR, Sukuru UR, Shi X (2017) Hemipteran-and coleopteran active toxin proteins from *Bacillus thuringiensis*. United States Patent 9546378

Bertrand JA, Sudduth TQ, Condon A, Jenkins TC, Calhoun MC (2005) Nutrient content of whole cottonseed. J Dairy Sci 88:1470-1477

Bowling AJ, Pence HE, Li H, Tan SY, Evans SL, Narva KE (2017) Histopathological effects of Bt and TcdA insecticidal proteins on the midgut epithelium of Western corn rootworm larvae (Diabrotica virgifera virgifera). Toxins (Basel) 9:e156

Bravo A, Gómez I, Conde J, Muñoz-Garay C, Sánchez J, Miranda R, Zhuang M, Gill SS, Soberón M (2004) Oligomerization triggers binding of a Bacillus thuringiensis Cry1Ab pore-forming toxin to aminopeptidase N receptor leading to insertion into membrane microdomains. Biochim Biophys Acta 1667:38-46

Bridge R, Bridge JM, McGowen R (2005) Cotton cultivar DP 393. United States Patent 6930228 B2

CERA (2011) A review of the environmental safety of the Cry1Ab protein. Center for Environmental Risk Assessment, ILSI Research Foundation, Washington http://ilsirf.org/publication/a-review-of-the-environmental-safety-of-the-cry1ab-protein/?kp1=2#publications1 (accessed November 2017)

Chen Y, Rivlin A, Lange A, Ye X, Vaghchhipawala Z, Eisinger E, Dersch E, Paris M, Martinell B, Wan Y (2014) High throughput Agrobacterium tumefaciens-mediated germline transformation of mechanically isolated meristem explants of cotton (*Gossypium hirsutum* L.). Plant Cell Rep 33:153-164

Cho S-H, Kang S-H, Lee Y-E, Kim S-J, Yoo Y-B, Bak Y-S, Kim J-B (2015) Distribution of toxin genes and enterotoxins in *Bacillus thuringiensis* isolated from microbial insecticide products. J Microbiol Biotechnol 25:2043-2048

Conlee KM, Stephens ML, Rowan AN, King LA (2005) Carbon dioxide for euthanasia: Concerns

regarding pain and distress, with special reference to mice and rats. Lab Anim 39:137-161

Delaney B, Astwood JD, Cunny H, Eichen Conn R, Herouet-Guicheney C, MacIntosh S, Meyer LS, Privalle LS, Gao Y, Mattsson J, Levine M, ILSI (2008) Evaluation of protein safety in the context of agricultural biotechnology. Food Chem Toxicol 46:S71-S97

Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM (1982) Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet 1:561-573

FAO (2017) FAOSTAT. Production: Crops. Food and Agriculture Organization of the United Nations, Rome. www.fao.org/faostat/en/#data/QC/visualize (accessed October 2017)

Farmer DR, Edrington TC, Kessenich CR, Wang C, Petrick JS (2017) Improving insect control protein activity for GM crops: A case study demonstrating that increased target insect potency can be achieved without impacting mammalian safety. Regul Toxicol Pharmacol 89:155-164

Gowda A, Rydel TJ, Wollacott AM, Brown RS, Akbar W, Clark TL, Flasinski S, Nageotte JR, Read AC, Shi X, Werner BJ, Pleau MJ, Baum JA (2016) A transgenic approach for controlling Lygus in cotton. Nat Commun 7(18):e_12213

Hamilton KA, Pyla PD, Breeze M, Olson T, Li M, Robinson E, Gallagher SP, Sorbet R, Chen Y (2004) Bollgard II cotton: compositional analysis and feeding studies of cottonseed from insect-protected cotton (Gossypium hirsutum L.) producing the Cry1Ac and Cry2Ab2 proteins. J Agric Food Chem 52:6969-6976

Harrison JM, Howard D, Malven M, Halls SC, Culler AH, Harrigan GG, Wolfinger RD (2013) Principal variance component analysis of crop composition data: A case study on herbicide-tolerant cotton. J Agric Food Chem 61:6412-6422.

Jerga A, Chen D, Zhang C, Fu J, Kouadio J-LK, Wang Y, Duff SMG, Howard JE, Rydel TJ, Evdokimov AG, Ramaseshadri P, Evans A, Bolognesi R, Park Y, Haas JA (2016) Mechanistic insights into the first Lygus-active β-pore forming protein. Arch Biochem Biophys 600:1-11

Jurat-Fuentes JL, Crickmore N (2017) Specificity determinants for Cry insecticidal proteins: Insights from their mode of action. J Invertebr Pathol 142:5-10

Kabaluk JT, Svircev AM, Goettal MS, Woo SG (eds) (2010) Use and regulation of microbial pesticides in representative jurisdictions worldwide. International Organization for Biological Control of Noxious Animals and Plants. http://www.iobc-global.org/publications iobc_use_ and_regulation_of_microbial_pesticides.html (accessed November 2017)

Koch MS, Ward JM, Levine SL, Baum JA, Vicini JL, Hammond BG (2015) The food and environmental safety of Bt crops. Front Plant Sci 6:283.

Moar WJ, Evans AJ, Kessenich CR, Baum JA, Bowen DJ, Edrington TC, Haas JA, Kouadio J-LK, Roberts JK, Silvanovich A, Yin Y, Weiner BE, Glenn KC, Odegaard ML (2017) The sequence, structural, and functional diversity within a protein family and implications for specificity and safety: The case for ETX_MTX2 insecticidal proteins. J Invertebr Pathol 142:50-59

Mogen BD, MacDonald MH, Graybosch R, Hunt AG (1990) Upstream sequences other than AAUAAA are required for efficient messenger RNA 3'-end formation in plants. Plant Cell 2:1261-1272

Munger P, Bleiholder H, Hack H, Hess M, Stauss R, van den Boom T, Weber E (1998) Phenological Growth Stages of the Cotton plant (*Gossypium hirsutum* L.) Codification and Description according to the BBCH Scale – with figures. J Agro Crop Sci 180, 143 - 149.

NCPA (2002) Cottonseed and its products. National Cottonseed Products Association. http://www.cottonseed.org/publications/cottonseedanditsproducts.asp (accessed October 2017)

NCPA (unknown) Functional properties of cottonseed oil. National Cottonseed Products Association http://www.cottonseedoiltour.com/pdf/PropertiesofCSO_whitepaper.pdf (accessed October 2017)

Nexter E, Thomashow LS, Metz M, Gordon M (2002) 100 years of *Bacillus thuringiensis*: a critical scientific assessment. American Academy of Microbiology. http://www.asmscience.org/content/report/colloquia/number

OECD (2003) Considerations for the Safety Assessment of Animal Feedstuffs Derived from Genetically Modified Plants. Series on the Safety of Novel Foods and Feeds, No. 9. Organisation for Economic Cooperation and Development, Paris.

OECD (2007) Consensus document on safety information on transgenic plants expressing *Bacillus thuringiensis* - derived insect control protein. Organisation for Economic Co-operation and Development, Paris

OECD (2008) Consensus document on the biology of cotton (Gossypium spp.). Series on Harmonisation of Regulatory Oversight in Biotechnology, no. 45. Organisation for Economic Co-operation and Development, Paris.

OECD (2009) Consensus document on compositional considerations for new varieties of cotton (*Gossypium hirsutum* and *Gossypium barbadense*): key food and feed nutrients and anti-nutrients. Organisation for Economic Co-operation and Development, Paris.

OGTR (2008) The Biology of *Gossypium hirsutum* L. and *Gossypium barbadense* L. (cotton). Office of the Gene Technology Regulator, Canberra.

Raymond B, Federici BA (2017) In defence of *Bacillus thuringiensis*, the safest and most successful microbial insecticide available to humanity—a response to EFSA. FEMS Microbiol Ecol 93(7) e_fix084

Richins RD, Scholthof HB, Shepherd RJ (1987) Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Res 15:8451-8466

Rogers SG (2000) Promoter for transgenic plants. United States Patent 6018100.

Samples JR, Buettner H (1983) Ocular infection caused by a biological insecticide. J Infect Dis 148:614

Shimada N, Kim YS, Miyamoto K, Yoshioka M, Murata H (2003) Effects of Bacillus thuringiensis Cry1Ab toxin on mammalian cells. J Vet Med Sci 65:187-191

Thomas K, Aalbers M, Bannon GA, Bartels M, Dearman RJ, Esdaile DJ, Fu T-J, Glatt CM, Hadfield N, Hatzos C, Hefle SL, Heylings JR, Goodman RE, Henry B, Herouet C, Holsapple M, Ladics GS, Landry TD, MacIntosh SC, Rice EA, Privalle LS, Steiner HY, Teshima R, van Ree R, Woolhiser M, Zawodny J (2004) A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. Regul Toxicol Pharmacol 39:87-98

USDA (2001) Oilseeds: World markets and trade. United States Department of Agriculture: Foreign Agricultural Service. Cotton, Oilseeds, Tobacco and Seeds Division, Washington DC

USDA (2016) Australia: Oilseeds and Products Annual. United States Department of Agriculture: Foreign Agricultural Service. Global Agricultural Information Network, Washington DC

USDA (2017) Oilseeds: World markets and trade. United States Department of Agriculture: Foreign Agricultural Service. Office of Global Analysis, Washington DC

van Frankenhuyzen K (2013) Cross-order and cross-phylum activity of *Bacillus thuringiensis* pesticidal proteins. J Invertebr Pathol 114:76-85.

WHO (1999) Microbial pest control agent *Bacillus thuringiensis*. Environmental Health Criteria 217. World health Organization, Geneva.

Yabe N, Takahashi T, Komeda Y (1994) Analysis of tissue-specific expression of *Arabidopsis thaliana* HSP90-family gene HSP81. Plant Cell Physiol 35:1207-1219

Zambryski P, Depicker A, Kruger K, Goodman HM (1982) Tumor induction by Agrobacterium tumefaciens: analysis of the boundaries of T-DNA. J Mol Appl Genet 1:361-370