

Supporting document 1

Safety assessment – Application A1087

Food derived from Insect-protected Soybean Line DAS-81419-2

Summary and Conclusions

Background

A genetically modified (GM) soybean line with OECD Unique Identifier DAS-81419-2, hereafter also referred to as soybean 81419, has been developed that is protected against several lepidopteran pests of soybean, including soybean looper (*Chrysodeixis includens*, formerly *Pseudoplusia includens*), velvetbean caterpillar (*Anticarsia gemmatalis*), fall armyworm (*Spodoptera frugiperda*) and tobacco budworm (*Heliothis virescens*).

The line contains two insecticidal genes, *cry1Ac(synpro)* and *cry1Fv3*, derived from the common soil bacterium *Bacillus thuringiensis* (often referred to just as 'Bt'). These genes express two insecticidal proteins which, for the purposes of this assessment are referred to as Cry1Ac and Cry1F. These two proteins have the same amino acid sequence as that found in WideStrike cotton considered in FSANZ Application A518.

In addition to the two *cry* genes, soybean 81419 contains a selectable marker gene (*pat*) from the bacterium *Streptomyces viridochromogenes*, which produces an enzyme (phosphinothricin acetyltransferase, PAT) that detoxifies the herbicide glufosinate ammonium. PAT functions as a selectable marker in the initial laboratory stages of plant cell selection and thus soybean 81419 is also tolerant to the herbicide glufosinate ammonium. However, the Applicant states it is not intended that this trait be used in commercial production of soybean 81419, and no MRL for glufosinate ammonium will be sought. The *pat* gene has been widely used for genetic modification of a number of crop species, including soybean.

In conducting a safety assessment of food derived from soybean line DAS-81419-2, a number of criteria have been addressed including: a characterisation of the transferred gene and its origin, function and stability in the soybean genome; the changes at the level of DNA, protein and in the whole food; compositional analyses; evaluation of intended and unintended changes; and the potential for the newly expressed proteins to be either allergenic or toxic in humans.

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

- any risks to the environment that may occur as the result of growing GM plants used in food production
- any risks to animals that may consume feed derived from GM plants
- the safety *per se* of food derived from the non-GM (conventional) plant.

History of Use

Soybean (*Glycine max*) is grown as a commercial crop in over 35 countries worldwide. Soybean-derived products have a range of food and feed as well as industrial uses and have a long history of safe use for both humans and livestock. Oil, in one form or another, accounts for the major food use of soybean and is incorporated in salad and cooking oil, bakery shortening, and frying fat as well as processed products such as margarine.

Molecular Characterisation

Comprehensive molecular analyses of soybean line DAS-81419-2 indicate that it contains a single intact copy of each of the three cassettes *cry1Fv3*, *cry1Ac(synpro)* and *pat* together with a small partial fragment of *cry1Ac(synpro)* at the 5' end of the insert. No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred to the plant. There has not been any disruption of endogenous genes as a result of the transformation procedure. The introduced genetic elements are stably inherited from one generation to the next.

Characterisation of Novel Protein

Soybean line DAS-81419-2 expresses three novel proteins, Cry1Ac, Cry1F and PAT. Expression analyses of the three proteins showed that all were detected in the plant parts tested. In general terms, it can be concluded that all three proteins are present in highest concentration in the leaves and lowest concentration in the roots.

A number of studies were used to confirm the identity and physicochemical properties of the plant-derived Cry1Ac, Cry1F and PAT proteins. These studies demonstrated that the three proteins conform in molecular weight and amino acid sequence to that expected, and do not exhibit any post-translational modification including glycosylation.

For all three proteins, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens; digestibility studies suggest the proteins would be rapidly degraded in the stomach following ingestion; and thermolability studies indicate that all three proteins are inactivated by heating. Taken together, the evidence indicates that Cry1Ac, Cry1F and PAT are unlikely to be toxic or allergenic to humans.

Compositional Analyses

Detailed compositional analyses were done to establish the nutritional adequacy of seed from soybean line DAS-81419-2. Analyses were done of proximates (moisture, crude protein, fat, ash, fibre), amino acids, fatty acids, vitamins, minerals, phytic acid, trypsin inhibitor, lectin, isoflavones, stachyose and raffinose. The levels were compared to levels in the seeds of a non-GM control line ('Maverick') grown alongside the GM line.

These analyses did not indicate any differences of biological significance between the seed from soybean DAS-81419-2 and the non-GM control 'Maverick'.

In an overall analysis, statistically significant differences were noted in a few constituents. However the differences were typically small, and all mean values were within both the reference range obtained for non-GM reference varieties grown at the same time and (where it exists) the literature range. Any observed differences are therefore considered to represent the natural variability that exists within soybean.

Conclusion

On the basis of the data provided in the present Application, and other available information, food derived from soybean line DAS-81419-2 is considered to be as safe for human consumption as food derived from conventional soybean cultivars.

Table of Contents

SUMMARY AND CONCLUSIONS.....	1
LIST OF TABLES.....	2
LIST OF FIGURES.....	2
LIST OF ABBREVIATIONS.....	3
1. INTRODUCTION.....	4
2. HISTORY OF USE.....	4
2.1 HOST ORGANISM.....	4
2.2 DONOR ORGANISMS.....	6
3. MOLECULAR CHARACTERISATION.....	7
3.1 METHOD USED IN THE GENETIC MODIFICATION.....	8
3.2 DESCRIPTION OF THE INTRODUCED GENES.....	8
3.3 BREEDING TO OBTAIN SOYBEAN LINE DAS-81419-2.....	11
3.4 CHARACTERISATION OF THE GENES IN THE PLANT.....	12
3.5 STABILITY OF THE GENETIC CHANGE.....	14
3.6 ANTIBIOTIC RESISTANCE MARKER GENES.....	14
3.7 CONCLUSION.....	15
4. CHARACTERISATION OF NOVEL PROTEINS.....	15
4.1 POTENTIAL ALLERGENICITY/TOXICITY OF NOVEL ORFs CREATED BY THE TRANSFORMATION PROCEDURE.....	15
4.2 FUNCTION AND PHENOTYPIC EFFECTS OF THE CRY1Ac, CRY1F AND PAT PROTEINS.....	16
4.3 NOVEL PROTEIN EXPRESSION IN PLANT TISSUES.....	18
4.4 PROTEIN CHARACTERISATION STUDIES.....	19
4.5 POTENTIAL TOXICITY.....	22
4.6 POTENTIAL ALLERGENICITY.....	25
5. COMPOSITIONAL ANALYSIS.....	28
5.1 KEY COMPONENTS OF SOYBEAN.....	28
5.2 STUDY DESIGN AND CONDUCT.....	29
5.3 SEED COMPOSITION.....	30
5.4 CONCLUSION.....	35
6. NUTRITIONAL IMPACT.....	35
REFERENCES.....	36

List of Tables

Table 1: Description of the genetic elements contained in the T-DNA of pDAB9582	9
Table 2: DAS-81419-2 generations used for various analyses	12
Table 3: Average concentration (ug/g dw) over ten locations of Cry1Ac, Cry1F and PAT proteins in various plant parts from soybean line DAS-81419-2	19
Table 4: Mean percentage of proximates and fibre in seed from 'Maverick' and DAS-81419-2	30
Table 5: Mean percentage composition, relative to total fat, of major fatty acids in seed from 'Maverick' and DAS-81419-2.....	31
Table 6: Mean percentage composition, relative to total amino acids, of amino acids in seed from 'Maverick' and DAS-81419-2.....	32
Table 7: Mean weight ($\mu\text{g/g}$ dry weight expressed as aglycon equivalents) of isoflavones in 'Maverick' and DAS-81419-2 seed.....	33
Table 8: Mean levels of anti-nutrients in 'Maverick' and DAS-81419-2 seed.	33
Table 9: Mean values for mineral levels in seed from 'Maverick' and DAS-81419-2.....	34
Table 10: Mean weight (mg/g dry weight) of vitamins in seed from 'Maverick' and DAS-81419-2	34

List of Figures

Figure 1: Vector map of plasmid pDAB9582	8
Figure 2: Representation of the genetic elements in the T-DNA insert of plasmid pDAB9582.....	9
Figure 3: Breeding strategy for plants containing event DAS-81419-2	11
Figure 4: Map of the DAS-81419-2 insert.....	13

List of Abbreviations

ADF	acid detergent fibre
AOAC	Association of Analytical Communities
APVMA	Australian Pesticides and Veterinary Medicines Authority
BLAST	Basic Local Alignment Search Tool
BLOSUM	Blocks Substitution Matrix
bp	base pairs
<i>Bt</i>	<i>Bacillus thuringiensis</i>
Cry	Crystal protein
CsVMV	Cassava vein mosaic virus
Cyt	Cytolytic protein
DIG	digoxigenin
DNA	deoxyribonucleic acid
T-DNA	transferred DNA
dw	dry weight
ELISA	enzyme linked immunosorbent assay
EPA	(U.S.) Environment Protection Agency
FAO	Food and Agriculture Organization of the United Nations
FARRP	Food Allergy Research and Resource Program
FASTA	Fast Alignment Search Tool - All
FDR	False discovery rate
FSANZ	Food Standards Australia New Zealand
fw	fresh weight
GM	genetically modified
IgG	Immunoglobulin G
kDa	kilo Dalton
LOQ	limit of quantitation
LSM	Least squares mean
MALDI-TOF MS	matrix assisted laser desorption/ionization time-of-flight mass spectrometry
MRL	maximum residue limit
MS/MS	Tandem mass spectrometry
MT	Million tonnes
NDF	neutral detergent fibre
NPTN	National Pesticide Telecommunications Network
OECD	Organisation for Economic Co-operation and Development
OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
PAT	Phosphinothricin acetyltransferase
PCR	polymerase chain reaction
PFT	pore-forming toxins
L-PPT	L-phosphinothricin
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
Ti	tumour inducing
U.S.	United States of America
USDA	United States Department of Agriculture
WHO	World Health Organization

1. Introduction

A genetically modified (GM) soybean line with OECD Unique Identifier DAS-81419-2, hereafter also referred to as soybean 81419, has been developed that is protected against several lepidopteran pests of soybean, including soybean looper (*Chrysodeixis includens*, formerly *Pseudoplusia includens*), velvetbean caterpillar (*Anticarsia gemmatilis*), fall armyworm (*Spodoptera frugiperda*) and tobacco budworm (*Heliothis virescens*).

The line contains two insecticidal genes, *cry1Ac(synpro)* and *cry1Fv3*, derived from the common soil bacterium *Bacillus thuringiensis* (often referred to just as 'Bt'). The *cry1Ac(synpro)* gene is a synthetic chimera comprising sequences from: the *cry1Ac1* gene originally isolated from *B. thuringiensis* subsp. *kurstaki* strain HD73; the *cry1Ca3* gene originally isolated from *B. thuringiensis* subsp. *aizawai* strain PS811; and the *cry1Ab1* gene originally isolated from *B. thuringiensis* subsp. *berliner* 1715. Similarly the *cry1Fv3* gene is chimeric and comprises sequences from the *cry1Fa2* gene originally isolated from *Bacillus thuringiensis* subsp. *aizawai* strain PS811; the *cry1Ca3* gene originally isolated from *B. thuringiensis* subsp. *aizawai* strain PS811; and at the *cry1Ab1* gene originally isolated from *B. thuringiensis* subsp. *berliner* 1715.

These genes express two insecticidal proteins which, for the purposes of this assessment are referred to as Cry1Ac and Cry1F. These two proteins have the same amino acid sequence as those expressed in WideStrike cotton considered in FSANZ Application A518 (FSANZ, 2005).

The Applicant claims that the genetic modification will provide growers in a number of countries around the world with improved insect resistance management and an alternative to chemical insecticides. Using two *B. thuringiensis*-derived insecticidal proteins, rather than one, in the same plant improves the spectrum of control and the seasonal efficacy and significantly reduces the chances of selecting insects resistant to the toxins. *Bt* formulations are widely used as biopesticides on a variety of cereal and vegetable crops grown organically or under conventional agricultural conditions.

In addition to the two *cry* genes, soybean 81419 contains a selectable marker gene (*pat*) from the bacterium *Streptomyces viridochromogenes*, which produces an enzyme (phosphinothricin acetyltransferase, PAT) that detoxifies the herbicide glufosinate ammonium. PAT functions as a selectable marker in the initial laboratory stages of plant cell selection and thus soybean 81419 is also tolerant to the herbicide glufosinate ammonium. However, the Applicant states it is not intended that this trait be used in commercial production of soybean 81419 and no MRL for glufosinate ammonium is being sought. The *pat* gene has been widely used for genetic modification of a number of crop species, including soybean.

It is anticipated that soybean 81419 may be grown predominantly in North and South America. The Applicant has not indicated that there is any intention to grow the plant line in Australia or New Zealand.

2. History of use

2.1 Host organism

The host organism is a conventional soybean (*Glycine max* (L.) Merr.), belonging to the family Leguminosae. The commercial soybean cultivar 'Maverick' 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 comparative assessment with soybean 81419. It was developed by the Missouri and Illinois Agricultural Experiment Stations and released in 1996 (Sleper *et al.*, 1998).

Soybean is grown as a commercial food and feed crop in many countries worldwide, with some 76 countries listed as producers in 2012 (FAOSTAT 2013), and has a long history of safe use for both humans and livestock. The major producers of soybean seed, accounting for 90% of world production, are the U.S. (2.45 MT), Argentina (1.03 MT), Brazil (1.00 MT), China (0.78 MT) and India (0.65 MT) (FAOSTAT 2013). Australia, while a net importer of soybean seed, grows crops in latitudes extending from the tropics (16° S) to temperate regions (37° S), mainly in the eastern states and as a rotational crop (James and Rose, 2004). The seed is used mainly to produce meal for use in animal feed (Grey, 2006).

In many soybean producing countries, GM soybean (mainly with a herbicide tolerant trait) accounts for a significant proportion of the total soybean grown e.g. U.S. (91%); Argentina (99%); Brazil (63%); South Africa (87%); Uruguay (99%) (Brookes and Barfoot, 2009). Australia does not currently grow any commercial GM soybean lines¹.

Soybean food products are derived either from whole or cracked soybeans:

- Whole soybeans are used to produce soy sprouts, baked soybeans, roasted soybeans and traditional soy foods such as miso, tofu, soy milk and soy sauce.
- Cracked soybeans have the hull (seed coat) removed and are then rolled into flakes which undergo solvent extraction to remove the oil.
- Crude oil is further refined to produce cooking oil, shortening and lecithin as well as being incorporated into a variety of edible and technical/industrial products. The flakes are dried and undergo further processing to form products such as meal (for use in livestock, pet and poultry food), protein concentrate and isolate (for use in both edible and technical/industrial products), and textured flour (for edible uses). The hulls are used in mill feed.

Unprocessed (raw) soybeans are not suitable for food use, and have only limited feed uses, as they contain toxicants and anti-nutritional factors, such as lectins and trypsin inhibitors (OECD, 2012). Appropriate heat processing inactivates these compounds.

Soybean oil constitutes approximately 30% of global consumption of edible fats and oils (The American Soybean Association, 2011; OECD, 2012), and is currently the second largest source of vegetable oil worldwide (USDA, 2009). Oil, in one form or another, accounts for the major food use of soybean (Shurtleff and Aoyagi, 2007) and is incorporated in salad and cooking oil, bakery shortening, and frying fat as well as processed products such as margarine.

Another possible food product that can be derived from the soybean plant is bee pollen. (Krell, 1996).

¹ See information on approved commercial releases of GM crops in Australia on the website of the Office of the Gene Technology Regulator - <http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/ir-1>

2.2 Donor organisms

2.2.1 *Bacillus thuringiensis* (Bt)

The two Cry proteins expressed in soybean line 81419 are derived from *B. thuringiensis*, a facultative anaerobic, gram-positive spore-forming bacterium that, while typically referred to as a soil bacterium, probably has its main ecological niche in insects (Federici, 1999).

The species is more appropriately regarded as a complex of over sixty subspecies that are characterised by the production of a proteinaceous crystal structure known as the parasporal body during the sporulation phase (Federici, 1999). The subspecies can be distinguished from one another on the basis of immunological differences in flagellar (H antigen) serotype or molecular techniques (see eg Yu *et al.*, 2002). The parasporal body contains one or more crystalline protein inclusions that are toxic to insects and are categorised as either Crystal (Cry) or Cytolytic (Cyt) toxins, also called δ -endotoxins. The Cry toxins are specifically toxic to the insect orders Coleoptera, Diptera, Hymenoptera, and Lepidoptera, and also to nematodes. The Cyt toxins are mostly found in *B. thuringiensis* strains that are active against Diptera. Both types of toxins are innocuous to humans, vertebrates and plants, and are completely biodegradable (Bravo *et al.*, 2007).

Studies on mammals, particularly laboratory animals, demonstrate that *B. thuringiensis* is mostly non-pathogenic and non-toxic. *B. thuringiensis* has been demonstrated to be highly specific in its insecticidal activity and has demonstrated little, if any, direct toxicity to non-target insects (see NPTN, 2000; OECD, 2007 and references therein). Infection in humans is unusual although there have been at least two clinical reports, one in the wounds of a soldier (Hernandez *et al.*, 1998) and one in burn wounds (Damgaard *et al.*, 1997), and in both cases impaired immunosuppression was implicated in the cause of the infection. *B. thuringiensis* has also been rarely associated with gastroenteritis (see eg Jackson *et al.*, 1995) but generally, *B. thuringiensis* present in drinking water or food has not been reported to cause adverse effects on human health (WHO, 1999; NPTN, 2000; OECD, 2007).

The effect of *B. thuringiensis* products on human health and the environment was the subject of a critical review by the WHO International Programme on Chemical Safety (WHO, 1999). The review concluded that '*B. thuringiensis* products are unlikely to pose any hazard to humans or other vertebrates or the great majority of non-target invertebrates provided that they are free from non-*B. thuringiensis* microorganisms and biologically active products other than the insecticidal proteins'.

With the exception of case reports on ocular irritation (Samples and Buettner, 1983) and inflammation after a needle stick injury (Warren *et al.*, 1984), no adverse health effects have been documented after occupational exposure to *B. thuringiensis* products. The use of *B. thuringiensis* products in the field can result in considerable aerosol and dermal exposure in humans. Studies of human populations exposed to *Bacillus thuringiensis* subsp. *kurstaki*, aerial spraying in Oregon (U.S.) and British Columbia (Canada) did not indicate any association with short-term clinical illness, except possibly in people who were immunocompromised (Green *et al.*, 1990; Valadares de Amorim *et al.*, 2001).

2.2.2 *Streptomyces viridochromogenes*

The source of the *pat* gene is the bacterial species *Streptomyces viridochromogenes*, strain Tü494 (Wohlleben *et al.*, 1988). The *Streptomycetaceae* bacteria were first described in the early 1900's. These organisms are generally soil-borne, although they may also be isolated from water. They are not typically pathogenic to animals including humans, and few species have been shown to be phytopathogenic (Kützner, 1981; Bradbury, 1986).

Although these organisms are not used in the food industry, the *pat* gene from *S. viridochromogenes*, has been used to confer glufosinate ammonium-tolerance in a range of food producing crops. The *bar* gene from the closely related *S. hygrosopicus* produces a protein that is structurally and functionally equivalent to the protein encoded by the *pat* gene (Wehrmann *et al.*, 1996) and has similarly been used widely for genetic modification of crop species.

2.2.3 Other organisms

Genetic elements from several other organisms have been used in the genetic modification of soybean 81419 (refer to Table 1). These non-coding sequences are used to drive, enhance or terminate expression of the two novel genes. None of the sources of these genetic elements is associated with toxic or allergenic responses in humans. The genetic elements derived from plant pathogens (Cassava vein mosaic virus, *Agrobacterium tumefaciens*) are not pathogenic in themselves and do not cause pathogenic symptoms in soybean 81419.

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.

Studies submitted:

Zhuang M. 2012. Transformation information for plasmid pDAB9582. *Study ID 110688*, Dow AgroSciences LLC, Indianapolis, IN

Hoffman T, Shan G. 2012. Event Sorting and Selection Process for the Development of DAS-81419-2 *Study ID 120687*, Dow AgroSciences LLC, Indianapolis, IN

Guttikonda S. 2012a. Bioinformatics Evaluation of the Putative Reading Frames across the Whole T-DNA Insert and Junctions in DAS-81419-2 Soybean for Potential Protein Allergenicity and Toxicity *Study ID: 120934*, Dow AgroSciences, LLC, Indianapolis, IN

Guttikonda S. 2012b. Molecular characterization of DAS-81419-2 soybean. *Study ID 110813*, Dow AgroSciences LLC, Indianapolis, IN

Guttikonda S, Richey K. 2012. Cloning and characterization of the DNA sequence for the insert and its flanking border regions of DAS-81419-2 soybean. *Study ID 102126*, Dow AgroSciences LLC, Indianapolis IN

Mo J. 2012a. Molecular characterization of DAS-81419-2 soybean within a single segregating generation. *Study ID 110814*, Dow AgroSciences LLC, Indianapolis, IN

3.1 Method used in the genetic modification

Soybean cultivar 'Maverick' was transformed via *Agrobacterium*-mediated transformation (Deblaere *et al.*, 1987) following the method of Zeng *et al.*(2004). The genes of interest were inserted into plasmid pDAB9582 (refer to Figure 1) between DNA sequences known as the Left and Right Borders (Border A and Border B in Figure 1). These border sequences were isolated from the tumour-inducing (Ti) plasmid of *Agrobacterium tumefaciens* and normally delimit the DNA sequence (T-DNA) transferred into the plant (Zambryski, 1988).

Basically, the cotyledonary nodes of *in vitro* germinated seedlings were co-cultivated with the *Agrobacterium tumefaciens* strain EHA101 (Hood *et al.*, 1986) containing the binary vector pDAB9582. Following shoot development, putative transformed shoots were selected on a medium containing glufosinate ammonium as the selection agent. The selected shoots were then rooted and transferred to soil, and the terminal leaflets of the resulting plantlets were leaf painted with glufosinate ammonium as a further screen. Selected plantlets (T_0) were sampled for molecular analysis that included verification of the absence of vector backbone and presence of the *pat*, *cry1Ac(synpro)* and *cry1Fv3* genes.

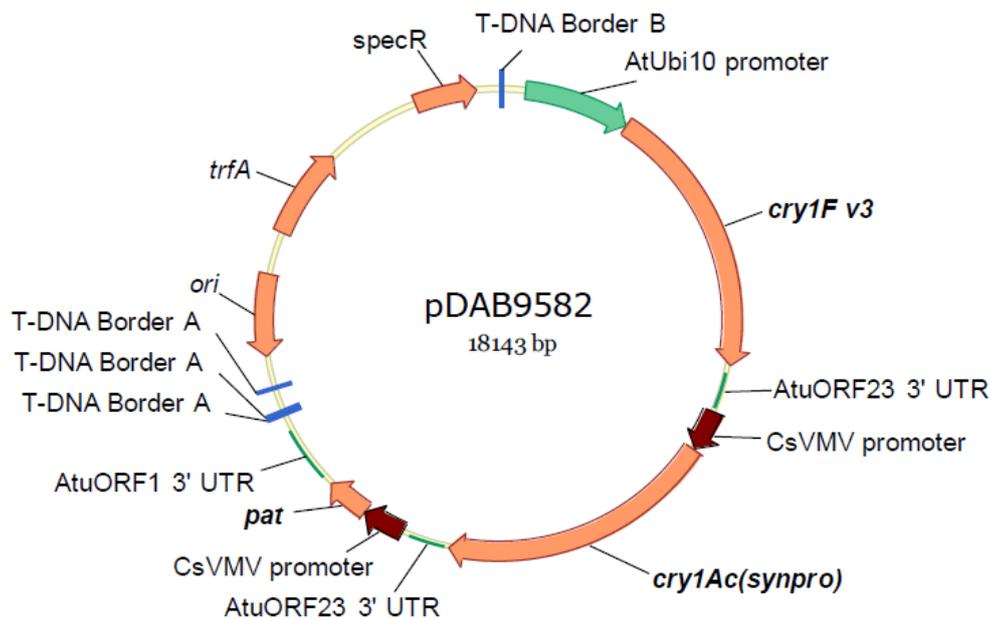


Figure 1: Vector map of plasmid pDAB9582

3.2 Description of the introduced genes

A diagram of the T-DNA insert in plasmid pDAB8264 is given in Figure 2. Information on the genetic elements in the T-DNA insert is summarised in Table1.

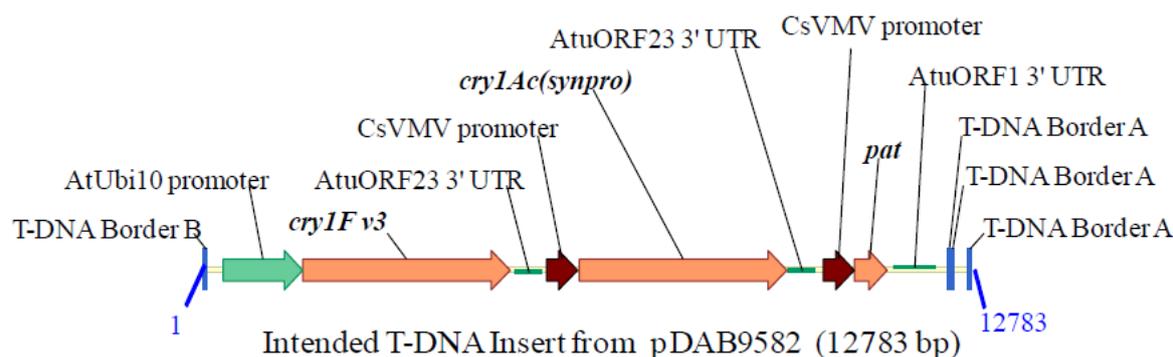


Figure 2: Representation of the genetic elements in the T-DNA insert of plasmid pDAB9582

Information on the genetic elements in the T-DNA insert is summarised in Table 1.

Table 1: Description of the genetic elements contained in the T-DNA of pDAB9582

Genetic element	bp location on pDAB89582	Size (bp)	Source	Orient.	Description & Function	References
Border B	1 - 24	24	<i>Agrobacterium tumefaciens</i>		• Required for the transfer of the T-DNA into the plant cell	Barker et al (1983)
cry1Fv3 cassette						
Intervening sequence	25 - 295	271			• Cloning sequence	
AtUbi10 promoter	296 - 1617	1322	<i>Arabidopsis thaliana</i>	Clockwise	• Polyubiquitin 10 promoter, 5'UTR and intron • Drives constitutive expression of the cry1Fv3 gene	Norris et al. (Norris et al., 1993)
Intervening sequence	1618 - 1625	8			• Cloning sequence	
Cry1Fv3	1626 - 5072	3447	<i>Bacillus thuringiensis</i>	Clockwise	• Coding sequence of a synthetic, chimeric construct comprising sequences from 3 genes but predominantly cry1F • Optimised for expression in plants	Cardineau et al. (2001); Gao et al. (2006);
Intervening sequence	5073 - 5174	102			• Cloning sequence	
AtuORF23 terminator	5175 - 5631	457	<i>Agrobacterium tumefaciens</i>	Clockwise	• Transcriptional terminator and polyadenylation site of open reading frame 23	Barker et al. (1983)
Intervening sequence	5632 - 5694	63			• Cloning sequence	
cry1Ac(synpro) cassette						
CsVMV promoter	5695 - 6211	517	Cassava vein mosaic virus	Clockwise	• Drives constitutive expression of the cry1Ac(synpro) gene	Verdaguer et al. (1996)
Intervening sequence	6212 - 6220	9			• Cloning sequence	
Cry1Ac(synpro)	6221 - 9691	3471	<i>Bacillus thuringiensis</i>	Clockwise	• Coding sequence of a synthetic, chimeric construct comprising sequences from 3 genes but predominantly cry1Ac1 • Optimised for expression in plants	Adang et al. (1985); Cardineau et al. (2001); Gilroy & Wilcox (1992)
Intervening sequence	9692 - 9724	33			• Cloning sequence	
AtuORF23 terminator	9725 - 10181	457	<i>Agrobacterium tumefaciens</i>	Clockwise	• Transcriptional terminator and polyadenylation site of open reading frame 23	Barker et al. (1983)
Intervening sequence	10182 - 10295	8			• Cloning sequence	

Genetic element	bp location on pDAB89582	Size (bp)	Source	Orient.	Description & Function	References
-----------------	--------------------------	-----------	--------	---------	------------------------	------------

pat cassette

CsVMV promoter	10296 - 10812	517	Cassava vein mosaic virus	Clockwise	• Drives constitutive expression of the <i>cry1Ac(synpro)</i> gene	Verdaguer <i>et al.</i> (1996)
Intervening sequence	10813 - 10819	7			• Cloning sequence	
<i>pat</i>	10820 - 11371	552	<i>Streptomyces viridochromogenes</i> strain Tü494	Clockwise	• Providing glufosinate ammonium tolerance. • Optimized for plant codon usage.	Wohleben <i>et al.</i> (1988)
Intervening sequence	11372 - 11473	102			• Cloning sequence	
<i>AtuORF1</i> terminator	11474 - 12177	704	<i>Agrobacterium tumefaciens</i>	Clockwise	• Transcriptional terminator and polyadenylation site of open reading frame 1	Barker <i>et al.</i> (1983)
Intervening sequence	12178 - 12405	228			• Cloning sequence	
					•	
Border A	12406 - 12429	25	<i>Agrobacterium tumefaciens</i>		• Required for the transfer of the T-DNA into the plant cell	Barker <i>et al.</i> (1983)

3.2.1 *cry1Fv3* expression cassette

The *cry1Fv3* gene is a synthetic, chimeric construct comprising 3447 base pairs (bp). The three components are:

- the *cry1Fa2* gene originally isolated from *Bacillus thuringiensis* subsp. *aizawai* strain PS811 (nucleotides 1 – 1810);
- the *cry1Ca3* gene originally isolated from *B. thuringiensis* subsp. *aizawai* strain PS811 (nucleotides 1811 – 1917); and
- the *cry1Ab1* gene originally isolated from *B. thuringiensis* subsp. *berliner* 1715 (nucleotides 1918 – 3447).

The sequences have been optimised for expression in plants.

The *cry1Fv3* coding region is controlled by the Arabidopsis (*Arabidopsis thaliana.*) ubiquitin-10 promoter. It is terminated by the polyadenylation and termination sequences derived from the open reading frame-23 (ORF23) terminator from *Agrobacterium tumefaciens* pTi15955.

3.2.2 *Cry1Ac(synpro)* expression cassette

The *Cry1Ac(synpro)* gene is also a synthetic, chimeric construct modified for optimum plant codon usage. It comprises 3,471 bp made up of

- the *cry1Ac1* gene originally isolated from *B. thuringiensis* subsp. *kurstaki* strain HD73 (nucleotides 1 – 1834);
- the *cry1Ca3* gene originally isolated from *B. thuringiensis* subsp. *aizawai* strain PS811 (nucleotides 1835 – 1941); and
- the *cry1Ab1* gene originally isolated from *B. thuringiensis* subsp. *berliner* 1715 (nucleotides 1942 – 3471).

Transcription of the *cry1Ac(synpro)* gene is controlled by the constitutive promoter from Cassava vein mosaic virus (CsVMV) and terminated by the ORF23 polyadenylation sequence.

Together, the *cry1Fv3* and *cry1Ac(synpro)* genes confer protection against several lepidopteran pests of soybean, including soybean looper (*Chrysodeixis includens*, formerly *Pseudoplusia includens*), velvetbean caterpillar (*Anticarsia gemmatilis*), fall armyworm (*Spodoptera frugiperda*) and tobacco budworm (*Heliothis virescens*).

3.2.3 *pat* gene expression cassettes

The *pat* gene encodes the PAT enzyme, which confers resistance to the herbicide glufosinate ammonium. This gene was introduced as a selectable marker for the identification of transformed plants. The *pat* gene was originally isolated from *Streptomyces viridochromogenes* Tü494 (Wohlleben *et al.*, 1988), but in this construct has been modified in order to alter the guanosine and cytosine codon bias to a level more typical for plant codons. The deduced amino acid sequence is identical to the native bacterial PAT enzyme.

Transcription of the *pat* gene is controlled with a constitutive promoter from CsVMV and terminated by a polyadenylation sequence from ORF1 of *Agrobacterium tumefaciens* pTi15955.

3.3 Breeding to obtain soybean line DAS-81419-2

A breeding programme was undertaken for the purposes of:

- obtaining generations suitable for analysing the molecular and genetic characteristics of soybean 81419
- ensuring that the 81419 event is incorporated into elite proprietary breeding line(s) for commercialisation..

The breeding pedigree for the various generations is given in Figure 3.

Following selection of initial transformants (T_0 plants) a series of self-fertilisation and seed bulking crosses proceeded up to generation T_6 . At the T_3 generation, plants were crossed with a non-GM elite commercial line to produce an F_1 generation which was either self-fertilised to produce an F_2 generation, or backcrossed to the appropriate parental elite cultivar to produce a BC_1F_1 generation which, in turn, was selfed to produce the BC_1F_2 generation.

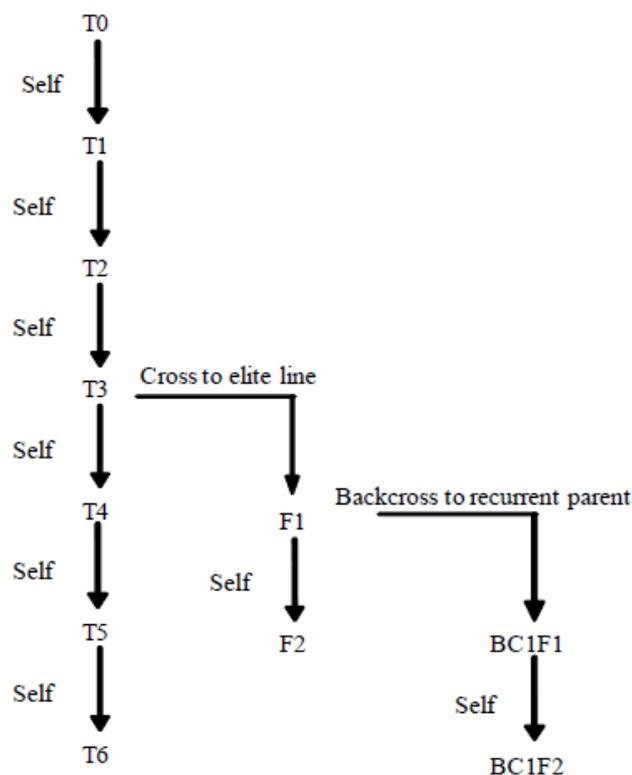


Figure 3: Breeding strategy for plants containing event DAS-81419-2

Table 2 indicates the generations that were used in the various studies characterising soybean 81419-2.

Table 2: DAS-81419-2 generations used for various analyses

Analysis	DAS-81419-2 Generation used	Control used
Molecular characterisation	T ₁ , T ₂ , T ₃ , T ₄ , F ₂	'Maverick'
Mendelian inheritance	F ₂ BC ₁ F ₂	
Genetic stability	T ₁ , T ₂ , T ₃ , T ₄ and F ₂	'Maverick'
Protein characterisation and comparison of plant and microbial proteins	T ₄ , T ₅	'Maverick'
Protein expression in plant parts	T ₄	'Maverick'
Compositional analyses	T ₄	'Maverick'

3.4 Characterisation of the genes in the plant

A range of analyses was undertaken to characterise the genetic modification in soybean line 81419. These included: DNA sequence, determination of insert copy number and integrity; and Open Reading Frame (ORF) analysis of inserted DNA as well as flanking and junction regions.

3.4.1 Insert characterisation

Genomic DNA was obtained from verified leaf tissue from the T₄ generation of soybean 81419. These samples were used to characterise the DNA sequence in the transgene insertion and its flanking border regions.

Ten overlapping polymerase chain reaction (PCR) fragments spanning the inserted sequences and border regions in event 81419 were amplified, purified and then cloned into a bacterial vector. For each fragment, the DNA from vector colonies was sequenced individually and the sequences were aligned to obtain a consensus sequence. Commercially available software (Sequencher®) was then used to assemble the consensus sequences to obtain a final sequence for the 81419 insert. This sequence was then compared to the T-DNA sequence in pDAB9582.

A total of 15,172 base pairs (bp) of DAS-81419-2 sequences were obtained comprising 1,297 bp of 5' flanking border sequence, 1,379 bp of 3' flanking border sequence and 12,496 bp of insert.

The sequences of the insert in event 81419 were shown to comprise (see Figure 4):

- A single intact copy of each of the *cry1Fv3*, *cry1Ac(synpro)* and *pat* cassettes.
- A 135 bp re-arranged sequence at the 5' end of the insert, of which 98 bp showed 99% identity with 1990 – 2087 bp of *cry1Ac(synpro)* in complementary orientation.
- A 9 bp re-arranged sequence at the 3' end of the insert

The Basic Local Alignment Search Tool for Nucleotides (BLASTN)(Agostino, 2013) was used to screen the DNA sequences, obtained for each of the 5' and 3' regions of the 81419-2 insert, for similarity to a sequence assembly of soybean genome from 'Williams 82', the elite U.S. cultivar that was chosen for sequencing of the whole soybean genome (Schmutz

et al., 2010). This indicated a 99% similarity with a segment of Chromosome 2 (GenBank Accession CM000835.1) of 'Williams 82' and therefore confirmed that the flanking sequences in event 81419 are of *Glycine max* origin.

Additionally, the border sequences of event 81419 were used to design primers for cloning the parental locus in 'Maverick'. The sequence thus obtained was also compared to the 'Williams 82' sequence assembly and similarly showed 99% identity to the segment in Chromosome 2 therefore confirming the parental locus of the insert in soybean 81419.

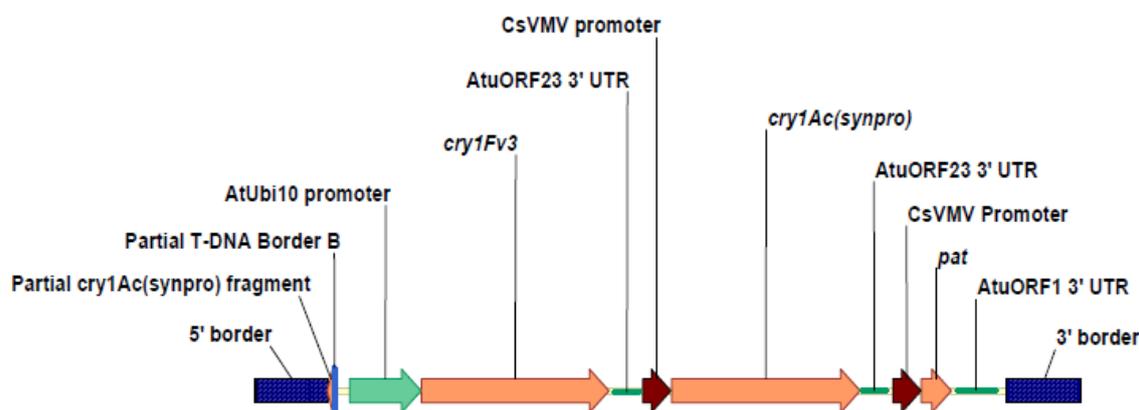


Figure 4: Map of the DAS-81419-2 insert

3.4.2 Transgene copy number, insertion integrity and plasmid backbone analysis

Total genomic DNA from pooled leaf tissue (at least three plants) of 5 generations (T_1 , T_2 , T_3 , T_4 and F_2) from lateral flow strip-verified (for presence of PAT) soybean line 81419 and from 'Maverick' (negative control) was used for Southern blot analyses. A positive control (DNA from 'Maverick' spiked with restriction enzyme--digested DNA from the pDAB9582 plasmid) was also included in the Southern blot analyses.

DNA from 81419, 'Maverick' and the positive control was digested with one, or a combination, of three restriction enzymes that cut within the T-DNA of pDAB9582. The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with 19 different digoxigenin (DIG)-labelled probes that represent various sections of the T-DNA (13 functional regions) and vector backbone.

The Southern blot analyses verified that

- Soybean line DAS-81419-2 contains a single intact copy of each of the three cassettes (see Table 1) *cry1Fv3*, *cry1Ac(synpro)* and *pat* together with a small partial fragment of *cry1Ac(synpro)* at the 5' end of the insert.
- There are no pDAB9582 backbone sequences present in the 81419 genome.

3.4.3 Novel open reading frame (ORF) analysis

The sequence of 15,172 bp obtained in the insert characterisation (Section 3.4.1) were analysed using an in-house Perl script to search for six-frame translations from stop codon to stop codon across the junction regions at the 5' and 3' borders and across the insert itself (without the flanking regions).

Within the insert, a total of 737 ORFs (greater than 8 amino acids) were identified. Nine ORFS (greater than 8 amino acids) were identified spanning the junctions across the insert

and its border regions. A discussion of the bioinformatic analysis of the novel ORFs is given in Section 4.1.

3.4.4 Analysis of possible disruption to endogenous genes at the pre-insertion locus

Sequence alignment of the parental locus in 'Maverick' with the border sequences in event 81419 indicated that a 57 bp fragment from the parental locus was deleted in 81419 as a consequence of the transformation. According to sequence information available at the time of analysis, no coding sequences or regulatory elements were identified in this Chromosome 2 (see Section 3.4.1) part of the parental locus from which the deletion was made.

3.5 Stability of the genetic change

The concept of stability encompasses both the genetic and phenotypic stability of the introduced trait over a number of generations. Genetic stability refers to maintenance of the modification over successive generations, as produced in the initial transformation event. It is best assessed by molecular techniques, such as Southern analysis or PCR, using probes and primers that cover the entire insert and flanking regions. Phenotypic stability refers to the expressed trait remaining unchanged over successive generations. It is often quantified by a trait inheritance analysis to determine Mendelian heritability via assay techniques (chemical, molecular, visual).

Phenotypic stability was assessed using greenhouse-grown plants of a segregating F₂ generation of soybean 81419 generated by crossing T₃ plants with an elite non-GM line. The F₁ plants were self-pollinated to obtain the F₂ generation (refer to Figure 3).

Leaves of 123 F₂ soybean 81419 plants were analysed by lateral flow strip testing for expression of the PAT protein. A Chi squared (χ^2) test for specified proportions was used to compare the observed segregation data to the hypothesised ratio of 3:1 (PAT positive : PAT negative). A total of 91 plants were positive for PAT while 32 were negative. The χ^2 value of 0.0678 ($P > 0.05$) indicated that the segregation ratio was consistent with the Mendelian inheritance pattern of a single dominant trait.

One hundred and twenty-two plants (one plant died before genomic extraction) were also tested by event-specific PCR for the presence/absence of the DAS-81419-2 insert. The results were entirely consistent with the PAT protein results i.e. all plants testing positive for PAT also tested positive for the insert, and all plants testing negative for PAT also tested negative for the insert.

Genetic stability was assessed by the Southern blot analyses described in Section 3.4.2 which utilized DNA isolated from five generations of 81419. The hybridization patterns were identical across the five generations, thus indicating the integrity and stable inheritance of the insert.

3.6 Antibiotic resistance marker genes

No antibiotic marker genes are present in soybean 81419. Plasmid backbone analysis (refer to Section 3.4.2) shows that no plasmid backbone has been integrated into the soybean genome during transformation, i.e. the *specR* gene, which was used as a bacterial selectable marker gene, is not present in soybean 81419.

3.7 Conclusion

Comprehensive molecular analyses of soybean line DAS-81419-2 indicate that it contains a single intact copy of each of the three cassettes *cry1Fv3*, *cry1Ac(synpro)* and *pat* together with a small partial fragment of *cry1Ac(synpro)* at the 5' end of the insert. No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred to the plant. There has not been any disruption of endogenous genes as a result of the transformation procedure. The introduced genetic elements are stably inherited from one generation to the next.

4. Characterisation of novel proteins

In considering the safety of novel proteins it is important to recognise that a large and diverse range of proteins are ingested as part of the normal human diet without any adverse effects, although a small number have the potential to impair health, e.g. because they are allergens or anti-nutrients (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-nutritional and allergenic effects. To effectively identify any potential hazards requires knowledge of the characteristics, concentration and localisation of all novel proteins expressed in the organism as well as a detailed understanding of their biochemical function and phenotypic effects. It is also important to determine if the novel protein is expressed as expected, including whether any post-translational modifications have occurred.

Two types of novel proteins were considered:

- those that may be potentially generated as a result of the creation of novel ORFs during the introduction of the T-DNA of plasmid pDAB9582 (see Section 3.4.3)
- those that were expected to be produced as a result of the expression of the introduced genes. Soybean 81419 expresses three novel proteins, Cry1Ac, Cry1F and PAT.

4.1 Potential allergenicity/toxicity of novel ORFs created by the transformation procedure

Study submitted:

Guttikonda S. 2012a. Bioinformatics Evaluation of the Putative Reading Frames across the Whole T-DNA Insert and Junctions in DAS-81419-2 Soybean for Potential Protein Allergenicity and Toxicity
Study ID: 120934, Dow AgroSciences, LLC, Indianapolis, IN

Nine novel ORFs were identified in the flanking regions and 737 in the T-DNA insert itself (refer to Section 3.4.3). The amino acid sequences corresponding to these ORFs were analysed for potential allergenicity and toxicity using an *in silico* approach. These analyses are entirely theoretical since there is no reason to expect that any of the identified ORFs would, in fact, be expressed.

4.1.1 Allergenicity assessment

The amino acid sequence of each identified ORF was compared with a peer-reviewed database containing 1,603 known and putative allergens, as well as coeliac-induction sequences residing in the FARRP (Food Allergy Research and Resource Program) dataset (Version 12) within AllergenOnline (University of Nebraska; <http://www.allergenonline.org/>). The allergen search utilised the Fast Alignment Search Tool - All (FASTA) search algorithm,

version 34, with Blocks Substitution Matrix50 (BLOSUM50) scoring matrix². ORFs shorter than 29 amino acids were not evaluated since a minimum 35% identity requires at least a match of 29 amino acids over 80 amino acids. The 35% identity is a recommended criterion for indicating potential allergenicity (FAO/WHO, 2001; Codex, 2003).

A separate eight-amino-acid search comparing every possible peptide of eight contiguous amino acids in the query sequence with the sequences in the FARRP AllergenOnline database was also carried out.

Of the nine ORFs identified in the flanking regions, four were less than 29 amino acids and of the 737 ORFs in the insert, 507 were less than 29 amino acids. For the remaining ORFs > 29 amino acids, in both the flanking regions and insert, no similarities with known allergens that exceeded the minimum 35% shared identity over a minimum of 80 amino acids were found. No matches of eight or more contiguous amino acids were found between any sequence and any entry in the FARRP AllergenOnline database.

Toxicity assessment

The sequences corresponding to the nine identified ORFs in the flanking regions were compared with protein sequences present in a number of large public reference databases including Swissprot, PIR (Protein Information Resource), PRF (Protein Research Foundation) and PDB (Protein Data Bank). The similarity searches used the BLASTP (Basic Local Alignment Search Tool Protein) algorithm (refer to Section 4.5.2 for an explanation). One significant similarity with an E-value of <1 (see Section 4.5.2 for explanation) was returned). This ORF encompassed the already identified fragment of *cry1Ac(synpro)* at the 5' end of the insert (see Section 3.4.1).

The BLASTP search of the 737 ORFs in the T-DNA insert returned 17 ORFs that showed alignments with an E-value <1.0. As expected, three of the alignments were with Cry1Ac, Cry1F and PAT. None of the remaining 14 ORFs returned alignments with any known protein toxins.

4.1.2 Conclusion

It is concluded that, in the unlikely event any of the identified novel ORFs were expressed, there is no significant similarity between the encoded sequences and any known protein toxins or allergens.

4.2 Function and phenotypic effects of the Cry1Ac, Cry1F and PAT proteins

4.2.1 Cry1Ac and Cry1F proteins

The general mechanism of insecticidal activity of Cry proteins is well understood (Gill *et al.*, 1992; Schnepf *et al.*, 1998; see eg Bravo *et al.*, 2007; OECD, 2007), with the mode of action being characterised principally in lepidopteran insects. The Cry proteins belong to a class of bacterial toxins known as pore-forming toxins (PFT) that are secreted as water-soluble

² The BLOSUM series of matrices tabulate the frequency with which different substitutions occur in conserved blocks of protein sequences and are effective in identifying distant relationships (Henikoff and Henikoff, 1992). The most commonly used BLOSUM matrices are BLOSUM50 (the default for use with the FASTA algorithm, with the matrix being built using sequences with no more than 50% similarity) and BLOSUM62 (the matrix used by BLAST and derived from blocks that are ≤ 62% identical).

proteins which, after undergoing conformational change, are able to insert into, or translocate across, the cell membranes of their host. There are two main groups of PFT: (i) the α -helical toxins in which the α -helix regions form the trans-membrane pore; and (ii) the β -barrel toxins, that insert into the membrane by forming a β -barrel composed of β -sheet hairpins from each monomer (Parker and Feil, 2005). The Cry proteins belong to the α -helical group of PFT, along with other toxins such as exotoxin A (from *Pseudomonas aeruginosa*) and diphtheria toxin.

The primary action of Cry toxins is to lyse midgut epithelial cells in the target insect by forming pores in the apical microvilli membrane of the cells, which subsequently leads to ion leakage and cell lysis. The crystal inclusions ingested by susceptible larvae dissolve in the alkaline environment of the gut, and the solubilised inactive protoxins are cleaved by midgut proteases yielding 60-70 kDa protease resistant core toxins (Bravo *et al.*, 2007). Toxin activation involves the proteolytic removal of an N-terminal peptide. The activated toxin then binds to specific receptors on the brush border membrane of the midgut epithelium columnar cells (Hofmann *et al.*, 1988; Aronson and Shai, 2001) before inserting into the membrane. Toxin insertion leads to formation of lytic pores in microvilli apical membranes (de Maagd *et al.*, 2001; Aronson and Shai, 2001) and eventually to cell lysis and disruption of the gut epithelium. The septicaemia that inevitably follows may be mediated by an influx of enteric bacteria into the haemocoel (Broderick *et al.*, 2006).

Cry1 insecticidal crystal proteins from *Bacillus thuringiensis* are a class of structurally related delta endotoxins. These proteins are generally toxic to a subset of lepidopteran larvae, and of all the Cry proteins, this class of Cry proteins has been investigated most thoroughly with regard to their mode of action. Cry proteins have been expressed in a range of agricultural crops (Sanahuja *et al.*, 2011).

The Cry1Ac protein produced in soybean 81419 is chimeric and comprises the core toxin of Cry1Ac1 and the non-toxic portions of Cry1Ca3 and Cry1Ab1 proteins. Together, the portions of the Cry1Ca3 and Cry1Ab1 proteins comprise the C-terminal domain and are removed by alkaline proteases during the formation of the Cry1Ac core toxin. The full length Cry1Ac is approximately 131 kDa and comprises 1156 amino acids, however it is digested by plant enzymes into the insecticidally active 65 kDa core toxin.

The Cry1F protoxin (1148 amino acids in length, molecular weight approximately 130 kDa) is also chimeric and comprises the core toxin of Cry1Fa2 and the non-toxic portions of Cry1Ca3 and Cry1Ab1 proteins. Together the portions of Cry1Ca3 and Cry1Ab1 that comprise the chimeric C-terminal domain are approximately those removed by alkaline proteases during the formation of the active Cry1Fa2 core toxin. The expressed protoxin (Cry1F synpro) is truncated to an active core toxin of approximately 65kDa.

4.2.2 PAT protein

Members of the genus *Streptomyces* produce antibiotics, one of which is bialaphos. These bacteria have evolved a mechanism to avoid the toxicity of their own products. Thus the *pat* gene from *Streptomyces viridochromogenes* and the *bar* gene from *S. hygrosopicus* both confer tolerance to bialaphos (Wehrmann *et al.*, 1996). Bialaphos, now also used as a non-selective herbicide, is a tripeptide composed of two L-alanine residues and an analogue of glutamate known as L-phosphinothricin (PPT) (see Thompson *et al.*, 1987) more recently known also as glufosinate ammonium. Free glufosinate ammonium released from bialaphos by peptidases (or applied directly as a synthetic herbicide) inhibits glutamine synthetase which in turn leads to rapid accumulation of ammonia and subsequent cell death.

The homologous polypeptide produced by the *bar* and *pat* genes (see Section 3.2.3) is known as phosphinothricin acetyltransferase (PAT); it is an acetyl transferase with enzyme

specificity for both PPT and demethylphosphinothricin (DMPT) in the acetylation reaction (Thompson *et al.*, 1987). In the presence of acetyl-CoA, PAT catalyses the acetylation of the free amino group of PPT to N-acetyl-PPT, a herbicidally-inactive compound. The kinetics and substrate specificity of the PAT enzyme are well characterised; it has a high specificity for PPT and has been shown to have a very low affinity to related compounds and amino acids; even excess glutamate is unable to block the PPT-acetyltransferase reaction (Thompson *et al.*, 1987).

The acetyltransferase activity is heat- and pH-dependent (Wehrmann *et al.*, 1996). PAT is active between temperatures of 25-55°C, with maximum activity occurring between 40° and 45°C. Complete thermoinactivation occurs after 10 minutes at 60°C and above. The optimum pH for PAT activity is 8.5, but it is active over a broad pH range of 6 to 11. The protein is expressed in a wide range of GM crop plants and is regarded as safe (see e.g. Hérouet *et al.*, 2005).

4.3 Novel protein expression in plant tissues

Study submitted:

Maldonado PM. 2012. Protein Expression of a Transformed Soybean Cultivar Containing Cry1Ac, Cry1F, and Phosphinothricin Acetyltransferase (PAT) - Event DAS-81419-2. *Study ID 110000.02*, Dow AgroSciences LLC, Indianapolis, IN

The Cry1Ac, Cry1F and PAT proteins are expected to be expressed in all plant tissues since the genes encoding them are driven by constitutive promoters (refer to Table 1). Ten locations in the U.S.³ representing regions where soybean is commercially grown, were planted with soybean 81419 (generation T₄) and 'Maverick' in 2011. This study used the same plots described in Section 6 – Compositional Analysis. Samples for analysis of expression of Cry1Ac, Cry1F and PAT were taken from a number of plant parts at specific growth stages (refer to Table 3).

The Cry1Ac, Cry1F and PAT protein levels were determined by enzyme linked immunosorbent assay (ELISA) using commercial ELISA kits (from Romer labs, Inc. for Cry1Ac and Cry1F; and from EnviroLogix Inc. for PAT).

No Cry1Ac or PAT proteins were detected in samples taken from 'Maverick' plants. In one out of 600 'Maverick' samples, a detectable level of Cry1F was found. Most likely this represented a sampling error or contamination.

For soybean 81419 plants, Cry1Ac, Cry1F and PAT proteins were detected in all plant parts analysed (Table 3). In general terms, it can be concluded that all three proteins are present in highest concentration in the leaves and lowest concentration in the roots. Plant expression of the Cry1F protein is higher than that of the Cry1Ac protein. This is probably due to the use of different promoters for the two genes. The Applicant states that the use of a different promoter for each gene provides less opportunity for negative interaction between the two inserts. The Applicant reports that expression of the two *cry* genes was sufficient to achieve good protection against the target pest species.

³ The test sites were located in Iowa (2 sites), Illinois (2 sites), Missouri (2 sites), Nebraska (2 sites), Indiana and Pennsylvania refer to Section 6.2 for more detail)

Table 3: Average concentration (ug/g dw) over ten locations of Cry1Ac, Cry1F and PAT proteins in various plant parts from soybean line DAS-81419-2

Sample source	Growth Stage*	Average protein content in µg/g dry weight ±SD		
		Cry1Ac	Cry1F	PAT
trifoliolate leaves	V5	25.44 ± 6.61	56.75 ± 15.03	5.23 ± 0.88
trifoliolate leaves	V10 - 12	23.16 ± 6.17	39.07 ± 16.6	5.6 ± 1.14
roots	R3	0.39 ± 0.24	5.23 ± 3.74	0.63 ± 0.12
forage	R3	5.54 ± 2.54	20.28 ± 11.29	4.06 ± 1.30
seed	R8	1.04 ± 0.1	13.8 ± 1.24	0.86 ± 0.13

*For information on soybean growth stages see e.g. Iowa State University (2009).

4.4 Protein characterisation studies

Studies submitted:

- Embrey S. 2012a. Certificate of Analysis for Full Length Cry1Ac (TSN102591) lyophilized protein standard. *BIOT 01-5808*, Dow AgroSciences LLC, Indianapolis, IN
- Embrey S. 2012b. Certificate of Analysis for Full Length Cry1F (TSN103748) lyophilized protein standard. *BIOT 02-7994*, Dow AgroSciences LLC, Indianapolis, IN
- Embrey SK, Schafer BW. 2009. Certificate of analysis of the test/reference/control substance: Phosphinothricin Acetyltransferase (PAT - TSN031116-0001). *BIOT09-203839*, Dow AgroSciences LLC, Indianapolis, IN
- Gao Y, Fencil KJ, Xu X, Schwedler DA, Gilbert JR, Herman RA. 2006. Purification and Characterization of a Chimeric Cry1F δ-Endotoxin Expressed in Transgenic Cotton Plants. *Journal of Agricultural and Food Chemistry* 54: 829-35
- Gao Y, Gilbert JR, Ni W, Xu X. 2002a. Characterization of Cry1Ac(synpro) Delta-Endotoxin Derived from Recombinant *Pseudomonas Fluorescens*. *Study ID GH-C 5508*, Dow AgroSciences LLC, Indianapolis, IN
- Schafer BW, Juba AN. 2012. Characterization of the Phosphinothricin Acetyltransferase (PAT) Protein Derived from Transgenic Soybean Event DAS-81419-2 *Study ID 120046*, Dow AgroSciences LLC, Indianapolis, IN
- Schafer BW, Oman TJ, Clement JM, Juba AN, Embrey SK. 2012a. Characterization of the Full Length Cry1Ac Protein Derived from Transgenic Soybean Event DAS-81419-2. *Study ID 110840*, Dow AgroSciences LLC, Indianapolis, IN
- Schafer BW, Oman TJ, Clement JM, Juba AN, Embrey SK. 2012b. Characterization of the Full Length Cry1F Protein Derived from Transgenic Soybean Event DAS-81419-2. *Study ID 110841*, Dow AgroSciences LLC, Indianapolis, IN

None of the novel proteins in soybean 81419 is produced in sufficient quantity to isolate for the studies required for a safety assessment. A standard procedure to overcome this problem is to produce the protein in a bacterial system and, if this protein shows equivalence to the *in planta*-produced protein, to then use the bacterially-produced protein as a surrogate for the plant-produced protein. The Cry1Ac, Cry1F and PAT proteins were therefore expressed in recombinant *Pseudomonas fluorescens* (Cry1Ac and Cry1F) or *Escherichia coli* (PAT). Characterisation tests were done to confirm the identity and equivalence of these bacterially-produced proteins to those produced in soybean 81419.

For Cry1Ac and Cry1F, crude protein extracts and immunopurified proteins were obtained from lateral flow assay-verified DAS-814819-2 seed of generation T₄ and immunopurified protein was obtained from a *P. fluorescens* bacterial expression systems. For PAT, crude protein extracts and immunopurified protein were obtained from lateral flow assay-verified

DAS-814819-2 seed of generation T₅ and immunopurified protein was obtained from an *E. coli* bacterial expression system. Crude extracts from 'Maverick' were also tested. The following analyses of the protein samples were then undertaken to confirm their identity as well as to compare the plant- and bacterially-produced proteins:

- Immunoreactivity – western blot and lateral flow strip
- glycosylation status (immunopurified Cry1Ac and Cry1F only)
- peptide mass mapping (Cry1Ac and Cry1F only)
- N-terminal sequencing (bacterially produced Cry1Ac and Cry1F only)
- Insecticidal activity (bacterially-produced Cry1F only)

Immunoreactivity

Lateral flow test strips specific to each protein were prepared commercially by EnviroLogix.

For the western blots, blotted polyvinylidene fluoride membranes were tested separately as follows:

- For Cry1Ac: polyclonal rabbit anti-Cry1Ac followed by conjugated goat anti-rabbit IgG horseradish peroxidase secondary antibody.
- For Cry1F: polyclonal rabbit anti-Cry1F followed by conjugated goat anti-rabbit IgG horseradish peroxidase secondary antibody.
- For PAT: polyclonal rabbit anti-PAT followed by conjugated goat anti-rabbit IgG horseradish peroxidase secondary antibody.

Cry1Ac and Cry1F

In the lateral flow strip test (capable of detecting 1 transgenic seed in 400), crude extract from 81419 tested positive for Cry1Ac and Cry1F while extract from 'Maverick' did not.

The sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE) gel of both bacterially-derived Cry1Ac and Cry1F showed three major bands and several minor bands with the major highest molecular weight band occurring at approximately 130 kDa and corresponding to the full length Cry1Ac or Cry1F (see discussion in Section 4.2.1). Western blot analysis showed that a number of the major and minor bands were immunoreactive to the specific Cry1Ac and Cry1F antibodies and that the bands below 130 kDa were therefore partially truncated forms of the proteins. The occurrence of the truncated forms indicated that there had been cleavage of some of the Cry1Ac/Cry1F by the proteases found in the *P. fluorescens* suspension. This is not unexpected since Cry proteins are known to have protease recognition sites on their exposed surface (Gao *et al.*, 2006). In contrast, the crude extract from DAS-81419-2, while showing many bands on a Coomassie stained SDS-PAGE gel, showed only one immunoreactive band (at approximately 130 kDa) in the western blot for both Cr1Ac and Cry1F. No bands were detected in the western blot of 'Maverick' crude extract.

PAT

In the lateral flow strip test, crude extract from 81419 tested positive for PAT while extract from 'Maverick' did not.

The SDS-PAGE gel of bacterially-derived PAT showed a major band at approximately 20 kDa and very faint staining of a band at approximately 40 kDa; which is consistent with the presence of a PAT dimer. This same band showed immunoreactivity in the western blot analysis. As expected, the crude plant extract showed numerous bands on a Coomassie-stained SDS-PAGE gel but the band at 20 kDa showed immunoreactivity in the western blot. No bands were detected in the western blot of 'Maverick' crude extract.

The Western blot analysis for each protein confirmed the plant- and microbial-derived proteins were of equivalent molecular weight and immunoreactivity. Immunoreactivity was further confirmed for each protein by the results of the lateral flow strip assays.

Glycosylation status

Many eukaryotic proteins are glycoproteins that have been post-translationally modified by the addition of carbohydrate moieties (glycans) covalently linked to the polypeptide backbone.

N-glycosylated proteins are glycosylated on an asparagine residue and commonly contain an asparagine-X-serine/threonine sequence (N-X-(P)-[S/T], where X-(P) indicates any amino acid except proline (Orlando and Yang, 1998). Although rare, the sequence asparagine-X-Cysteine (N-X-C) can also be an N-glycosylation site (Miletich and Broze Jr., 1990). No N-glycosylation sites were predicted from the amino acid sequence.

Analysis of immunoaffinity-purified plant- and microbial-derived proteins was done using a commercial kit (GelCode Glycoprotein Staining Kit from ThermoScientific) following SDS-PAGE. The kit detects carbohydrates that may be covalently linked to the protein. A glycosylated protein (horseradish peroxidase) was applied to each gel as a positive control while the non-glycosylated protein, soybean trypsin inhibitor, was used as a negative control. A visible band was obtained for horseradish peroxidase while the soybean trypsin inhibitor and the Cry1Ac and Cry1F immunopurified proteins from both plant and microbial sources gave no visible bands.

These results support the conclusion that neither microbially nor DAS-81419-2-derived Cry1Ac and Cry1F proteins are glycosylated.

Peptide mass mapping

A protein identification made by peptide mass fingerprinting is considered to be reliable if the measured coverage of the sequence is 15% or higher with a minimum of five peptide matches (Jensen *et al.*, 1997).

Purified Cry1Ac and Cry1F proteins from microbial and plant sources were denatured (i.e. heated or run on SDS-PAGE and the corresponding 130 kDa protein bands were excised) and digested with trypsin (as well as chymotrypsin in the case of the plant-derived proteins) followed by matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) to determine the peptide mass fingerprint coverage. The sequences thus obtained were verified by tandem mass spectrometry (MS/MS). The masses of the detected peptides were searched *in silico* against those deduced from potential trypsin/chymotrypsin cleavage sites within Cry1Ac/Cry1F amino acid sequences, using Protein Analysis Worksheet freeware from Proteometrics LL.

For microbially-derived Cry1Ac, 27 peptides were identified matching the theoretical deduced peptide masses of Cry1Ac(synpro). Two internal peptides were also sequenced and matched residues 182 – 192 and 350 – 360 of Cry1Ac (synpro). For the plant-derived Cry1Ac, sequence coverage was approximately 85% of the theoretical total peptide sequence and confirmed that the plant-derived protein amino acid sequence matched that of the microbial-derived Cry1Ac.

For microbially-derived Cry1F, 37 peptides were identified matching the theoretical deduced peptide masses of Cry1Ac(synpro) and covered 40% of the Cry1F protein sequence. The

plant derived protein covered approximately 82% of the sequence and confirmed that the plant-derived sequence matched the microbial-derived sequence.

N-terminal sequencing

Purified microbial-derived Cry1Ac and Cry1F proteins were separated from minor impurities by running on SDS-PAGE. Both full-length (~130 kDa) and various truncated forms of each protein were electro-blotted onto a polyvinylidene difluoride (PVDF) membrane, The bands were excised and subjected to Edman degradation using an automated commercial sequencer.

For the full length Cry1Ac, the first 15 amino acids were N-terminal sequenced and confirmed the expected sequence. For the 65 kDa core toxin, the first 15 amino acids of the N-terminal sequence matched residues #29 – 43 of the full length protein. This suggests that during proteolysis, in addition to the removal of a large piece of C-terminal sequence, the first 28 residues from the N-terminus are also cleaved.

The N-terminal sequences obtained from the full-length Cry1F and various forms of partially truncated Cry1F matched the first five residues of the theoretical N-terminal sequence of Cry1F(synpro). Similarly, for Cry1F, the first 10 amino acids of the core toxin were found to correspond to residues #28 – 37 of the full length protein. Again, the conclusion is that proteolysis leads to truncation from the C-terminal domain initially and then from the first 27 residues of the N-terminus, leaving the 65kDa trypsin-resistant core toxin.

This feature of the two Cry1 proteins aligns with their fate during insect ingestion (see Section 4.2.1) where, it is postulated, the full length protein undergoes the removal of an N-terminal peptide of 25 – 30 amino acids and the cleavage of approximately half of the sequences from the C-terminus, leaving a core toxin of 60 – 70 kDa. Both the C-terminal and N-terminal residues are believed to be dispensable for toxicity *per se* (see references and discussion in Gao *et al.*, 2006).

Insecticidal activity

The Applicant submitted supplementary evidence of insecticidal activity in one of the Cry proteins, which involved a diet-overlay bioassay (Gao *et al.*, 2004; Gao *et al.*, 2009) using microbial-derived Cry1F. Three insect pests were selected based on their susceptibility to the Cry1F endotoxin – tobacco budworm (highly susceptible), beet armyworm (susceptible) and cotton bollworm (less susceptible). GI₈₀ (80% growth inhibition) values were calculated. In this particular study, the microbial-derived Cry1F was compared to Cry1F (identical coding region to the Cry1F in soybean 81419) obtained from a transgenic cotton and the proteins from both sources had similar and the expected potencies.

Conclusion

A range of characterisation studies confirmed the identity and equivalence of the Cry1Ac, Cry1F and PAT proteins produced in both a bacterial expression system and in soybean DAS-81419-2. Based on weight-of-evidence, it is concluded that microbially-derived Cry1Ac, Cry1F and PAT proteins are suitable surrogates for use in safety assessment studies.

4.5 Potential toxicity

While the vast majority of proteins ingested as part of the diet are not typically associated with toxic effects, a small number may be harmful to health. Therefore, if a GM food differs from its conventional counterpart by the presence of one or more novel proteins, these proteins should be assessed for their potential toxicity. The main purpose of an assessment

of potential toxicity is to establish, using a weight of evidence approach, that the novel proteins will behave like any other dietary protein.

The assessment focuses on:

- whether the novel proteins have a prior history of safe human consumption, or are sufficiently similar to proteins that have been safely consumed in food;
- amino acid sequence similarity with known protein toxins and anti-nutrients;
- structural properties of the novel proteins including whether they are resistant to heat or processing and/or digestion.

An oral toxicity study is only deemed necessary if the results of biochemical, bioinformatic, digestibility or stability studies indicate further investigation of potential toxicity is warranted.

4.5.1 History of human consumption

Bt proteins are used widely as an insecticide in both conventional and organic agriculture. In Australia, various *Bt* insecticidal products are registered with the Australian Pesticides and Veterinary Medicines Authority (APVMA) for use on cotton, vegetables, fruits, vines, oilseeds, cereal grains, herbs, tobacco, ornamentals, forestry and turf. This wide use of formulations containing the *Bt* insecticidal proteins indicates that people eating and handling fresh foods are commonly in contact with this protein. No safety issues have been raised with the use of food products derived from *Bt* crops (Mendelsohn *et al.*, 2003; OECD, 2007; Delaney *et al.*, 2008; Shelton *et al.*, 2009).

Insecticidal products using *Bt* were first commercialised in France in the late 1930s (Nester *et al.*, 2002) and were first registered for use in the United States by the Environment Protection Agency (EPA) in 1961 (EPA, 1998). The EPA thus has a historical toxicological database for *B. thuringiensis*, which indicates that no adverse health effects have been demonstrated in mammals in any infectivity/ pathogenicity/ toxicity study (McClintock *et al.*, 1995; EPA, 1998; Betz *et al.*, 2000). This confirms the long history of safe use of *Bt* formulations.

Streptomyces hygroscopicus and *S. viridochromogenes* (refer to Section 2.2.1) are common soil bacteria, therefore humans have a long history of exposure to the PAT protein through the consumption of roots and vegetables. Since 1995, humans have also been directly exposed to the PAT protein through the consumption of foods derived from GM glufosinate ammonium-tolerant canola, soybean, cotton and corn, without any evidence of toxicity (Hérouet *et al.*, 2005; Delaney *et al.*, 2008).

4.5.2 Amino acid sequence similarity to known protein toxins

Bioinformatic analyses are useful for assessing whether introduced proteins share any amino acid sequence similarity with known protein toxins.

Studies submitted:

Guttikonda S. 2012d. Sequence Similarity Assessment of Cry1Ac Protein to Known Toxins by Bioinformatics Analysis (Update, May, 2012). *Study ID 120761*, Dow AgroSciences LLC, Indianapolis, IN

Guttikonda S. 2012f. Sequence Similarity of PAT Protein to Known Toxins by Bioinformatics Analysis (Update, February, 2012). *Study ID 120480*, Dow AgroSciences LLC, Indianapolis IN

Song P. 2012. Sequence Similarity Assessment of Cry1F Protein to Known Toxins by Bioinformatics Analysis. *Study ID 120762*, Dow AgroSciences LLC, Indianapolis, IN

A similarity search was done for the Cry1Ac, Cry1F and PAT proteins, using the BLASTP⁴ (Basic Local Alignment Search Tool Protein – version 2.2.21) algorithm (Altschul *et al.*, 1997), and BLOSUM62 scoring matrix (see Footnote 2), against non-redundant known protein sequences present in a number of large public reference databases including GenBank, RefSeq, Uniprot_Swissprot, PIR (Protein Information Resource), PRF (Protein Research Foundation) and PDB (Protein Data Bank).

BLASTP is now frequently applied for searching for similarities in protein sequences by performing local alignments of domains or short sequence similarities; this detects more similarities than would be found using the entire query sequence length. The search generates a parameter known as the *E* value (see eg Baxevanis, 2005). Comparisons between highly homologous proteins yield *E*-values approaching zero, indicating the very low probability that such matches would occur by chance. A larger *E*-value indicates a lower degree of similarity. All database sequences with an *E*-value of 1 or lower were identified by default by the BLASTP program. A statistically significant sequence similarity generally requires a match with an *E*-value of less than 0.01 (Pearson, 2000). Commonly, for protein-based searches, hits with *E*-values of 10^{-3} or less and sequence identity of 25% or more are considered significant although any conclusions reached need to be tempered by an investigation of the biology behind the putative homology (Baxevanis, 2005).

It would be expected that both the Cry1Ac and Cry1F protein sequences would show similarity with other delta toxins and endotoxins. The Cry1Ac similarity search identified 662 alignments with an *E* value of < 0.01, of which 610 were identified as being of the same class of insecticidal proteins from *B. thuringiensis* and two were from other *Bacillus* species. The remaining 52 alignments were related to other delta toxin and endotoxin proteins, parasporal crystal proteins and hypothetical proteins from non-*Bacillus* species. None of the other significant sequence alignments were related to any other known protein toxins. A further 22 alignments were found with *E*-values between 0.01 and 1.0 (i.e. weak alignments) but none was associated with known toxins.

The Cry1F similarity search identified 646 alignments with an *E* value of < 0.01, of which 618 were identified as being of the same class of insecticidal proteins from *B. thuringiensis*. The remaining 28 alignments were related to other insecticidal proteins, parasporal crystal proteins and hypothetical proteins from non-*Bacillus* species. None of the other significant sequence alignments were related to other known protein toxins. A further 26 alignments were found with *E*-values between 0.01 and 1.0 (i.e. weak alignments) but none was associated with known toxins.

The PAT similarity search identified 1,855 alignments with an *E* value of < 0.01. As expected, the PAT protein showed a high degree of homology with other acetyltransferases, none of which is considered to be a toxin. There were also matches with putative uncharacterised proteins. There were 283 alignments with *E*-values between 0.01 and 1.0 (i.e. weak alignments) but none was associated with known toxins.

Conclusion

The Cry1Ac, Cry1F and PAT proteins do not have significant similarity with known toxins that are harmful to human health.

4.5.3 *In vitro* digestibility

See Section 4.6.3.

⁴ BLASTP is used to compare a protein sequence with a database of protein sequences.

4.5.4 Stability to heat

The thermolability of a protein provides an indication of the stability of the protein under cooking/processing conditions.

Study submitted:

Embrey, S.K.; Shan, G. (2005). Heat lability of insecticidal proteins Cry1Ac and Cry1F. Dow AgroSciences LLC, Indianapolis, Indiana. Study ID GH-C 5777

The Applicant previously supplied two studies concerning the thermolability of Cry1Ac and Cry1F with Application A518 (FSANZ, 2005). Aqueous solutions of the proteins were heated at 60°C, 75°C, and 90°C for 30 minutes and then assayed through an artificial insect diet bioassay. Both proteins were totally inactivated after treatment at 90°C; Cry1F was inactivated at 75°C and Cry1Ac was almost entirely inactivated at this temperature.

In a more recent study, thermolability of Cry1F and Cry1Ac proteins (produced in recombinant *Pseudomonas fluorescens*) was investigated by incubating aqueous formulations of each protein at 91°C for 60 min. The samples were then cooled on ice and assayed by ELISA using a sequential double-antibody format (1^o antibody was polyclonal rabbit anti-Cry1Ac or anti-Cry1F; 2^o antibody was conjugated goat anti-rabbit IgG horseradish peroxidase). The development of colour in the assay is directly proportional to the presence of epitopes (immunoreactivity). The results indicated that almost all of the epitopes in both proteins were destroyed during the heat treatment. Therefore it can be concluded that both the Cry1Ac and Cry1F proteins are immunochemically unstable at 91°C.

The Applicant did not supply a study for PAT. However, it has been established that the PAT protein is completely inactivated after 10 min at 50° C (Hérouet *et al.*, 2005).

4.5.5 Acute oral toxicity studies

Studies submitted

Brooks KJ, Andrus AK. 1999. Cry1f Microbial Protein (FL): Acute Oral Toxicity Study in CD-1 Mice. Study ID 991178, Dow AgroSciences LLC, Indianapolis, IN
Brooks KJ, Yano BL. 2001. Cry1Ac-(SYNPRO) Microbial Protein: Acute Oral Toxicity Study In CD-1 Mice. Study ID 011126, Dow AgroSciences LLC, Indianapolis, IN

In view of the extensive and long term human exposure to Cry proteins (see sections 2.2.1 and 4.5.1), an acute oral toxicity study in mice is considered redundant for the purpose of this safety assessment. Nonetheless, the Applicant supplied the above-mentioned studies for Cry1Ac and Cry1F both of which concluded that there were no adverse effects.

Similarly, since the PAT protein has already been considered in 15 FSANZ approvals (A372, A375, A380, A385, A386, A446, A481, A518, A533, A543, A589, A1028, A1040, A1046, A1073), as well as being accepted in the literature as not having toxicity concerns (see e.g. Delaney *et al.*, 2008) no study is required for this safety assessment.

4.6 Potential allergenicity

The potential allergenicity of novel proteins was evaluated using an integrated, step-wise, case-by-case approach relying on various criteria used in combination. This is because no

single criterion is sufficiently predictive of either allergenicity or non-allergenicity (see e.g. Thomas *et al.*, 2009). The assessment focuses on:

- the source of the novel protein;
- any significant amino acid sequence similarity between the novel protein and known allergens;
- the structural properties of the novel protein, including susceptibility to digestion, heat stability and/or enzymatic treatment; and
- specific serum screening if the novel protein is derived from a source known to be allergenic or has amino acid sequence similarity to a known allergen, additional *in vitro* and *in vivo* immunological testing may be warranted.

Applying this approach systematically provides reasonable evidence about the potential of a novel protein to act as an allergen.

4.6.1 Source of the protein

The Cry1Ac, Cry1F and PAT proteins are derived from common soil bacteria to which humans have been naturally exposed and which may have been inadvertently ingested on fresh produce without eliciting adverse effects. Neither *Bacillus thuringiensis* nor *Streptomyces hygroscopicus* is considered to be a source of allergenic proteins (OECD, 2007; see eg EFSA, 2007). Bacterial proteins are rarely allergenic because of the low exposure levels and lack of allergic sensitisation (Taylor, 2002).

4.6.2 Amino acid sequence similarity to known allergens

Bioinformatic analysis is part of a 'weight of evidence' approach for assessing potential allergenicity of novel proteins introduced to GM plants (Thomas *et al.*, 2005; Goodman, 2006). It is a method for comparing the amino acid sequence of the introduced protein with sequences of known allergens in order to indicate potential cross-reactivity between allergenic proteins and the introduced protein. As with the bioinformatic analysis that looked at similarities of the novel proteins with known protein toxins (refer to Section 4.5.2), the generation of an *E* value provides an important indicator of significance of matches (Pearson, 2000; Baxevanis, 2005). The criteria used to indicate potential allergenicity were a minimum of eight-contiguous amino acid identity or 35% identity on a window of 80 amino acids within the sequence of an allergenic protein (FAO/WHO, 2001; Codex, 2003). Refer also to Section 4.1.1 for a discussion of the bioinformatic analysis.

Studies submitted:

Guttikonda S. 2012c. Sequence Similarity Assessment of Cry1Ac Protein to Known Allergens by Bioinformatics Analysis (Update, May, 2012). *Study ID 120763*, Dow AgroSciences LLC, Indianapolis, IN

Guttikonda S. 2012e. Sequence Similarity Assessment of PAT Protein to Known Allergens by Bioinformatics Analysis (Update, February, 2012). *Study ID 120143*, Dow AgroSciences LLC, Indianapolis IN

Mo J. 2012b. Sequence Similarity Assessment of Cry1F to Known Allergens by Bioinformatics Analysis. *Study ID 120764*, Dow AgroSciences LLC, Indianapolis, IN

The Cry1Ac, Cy1F and PAT sequences were compared with known and putative allergen sequences contained in a reference allergen database, (FARRP version 12 – released in February 2012 and containing 1,603 non-redundant entries) using the FASTA algorithm and BLOSUM50 scoring matrix (refer to Section 4.1).

No matches were found for any of the three proteins with known allergenic proteins or with known allergenic epitopes and neither did any of the three proteins share a sequence of eight or more consecutive identical amino acids with any potential allergens.

4.6.3 *In vitro* digestibility

Typically, food proteins that are allergenic tend to be stable to enzymes such as pepsin and the acidic conditions of the digestive system, exposing them to the intestinal mucosa and leading to an allergic response (Astwood and Fuchs, 1996; Metcalfe *et al.*, 1996; Kimber *et al.*, 1999). Therefore some correlation exists between resistance to digestion by pepsin and potential allergenicity although it does not necessarily follow that resistance to digestion is always an indicator of an allergenic protein (Thomas *et al.*, 2004; Herman *et al.*, 2007). As a consequence, one of the criteria for assessing potential allergenicity is to examine the stability of novel proteins in conditions mimicking human digestion. Proteins that are rapidly degraded in such conditions are considered less likely to be involved in eliciting an allergic response. However, evidence of slow or limited protein digestibility does not necessarily indicate that the protein is allergenic.

Studies submitted

Korjagin VA. 2001a. *In vitro* simulated gastric fluid digestibility study of microbially derived Cry1Ac (synpro). Study ID 010026, Dow AgroSciences LLC, Indianapolis, IN
Korjagin VA. 2001b. *In vitro* simulated gastric fluid digestibility study of microbially derived Cry1F (synpro). Study ID 010081, Dow AgroSciences LLC, Indianapolis, IN

It is noted that the two studies supplied by the Applicant are the same studies submitted in the Application dossier for A518 (FSANZ, 2005) and have therefore already been considered by FSANZ. In the safety assessment for that Application, FSANZ made the following comments:

“Samples of both Cry1Ac and Cry1F (produced in recombinant P. fluorescens) were incubated with simulated gastric fluid (SGF) at 37°C to determine if these two proteins would be digested. The digestions were performed at time intervals of 1, 3, 6, 10, 15, 20, 30, and 60 minutes. Following digestion, the protein samples were analysed by SDS-PAGE and Western blotting. Both Cry1Ac and Cry1F were fully digested in SGF in under 1 minute.”

The Applicant did not supply a study for PAT. However, as stated in Section 4.5 there has already been adequate consideration of the protein in previous applications to FSANZ. Additionally, there is no evidence in the literature (Wehrmann *et al.*, 1996; Delaney *et al.*, 2008) to suggest that the PAT protein may be associated with any allergenicity concerns.

4.6.4 *Stability to heat*

See Section 4.5.4.

4.6.5 *Conclusion*

Soybean line DAS-81419-2 expresses three novel proteins, Cry1Ac, Cry1F and PAT. Expression analyses of the three proteins showed that all were detected in the plant parts tested. In general terms, it can be concluded that all three proteins are present in highest concentration in the leaves and lowest concentration in the roots.

A number of studies were used to confirm the identity and physicochemical properties of the plant-derived Cry1Ac, Cry1F and PAT proteins. These studies demonstrated that the three

proteins conform in size and amino acid sequence to that expected, and do not exhibit any post-translational modification including glycosylation.

For all three proteins, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens; digestibility studies suggest the proteins would be rapidly degraded in the stomach following ingestion; and thermolability studies indicate that all three proteins are inactivated by heating. Taken together, the evidence indicates that Cry1Ac, Cry1F and PAT are unlikely to be toxic or allergenic to humans. Additionally the extensive and long term exposure to Cry proteins through conventional agricultural practices confirm that they do not cause any adverse effects in humans.

5. Compositional analysis

The main purpose of compositional analysis is to determine if any unexpected changes in composition have occurred to the food and to establish its nutritional adequacy. Compositional analysis can also be important for evaluating the intended effect where there has been a deliberate change to the composition of food.

The classic approach to the compositional analysis of GM food is a targeted one; rather than analysing every single 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 as anti-nutrients) or 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 (eg solanine in potatoes).

5.1 Key components of soybean

For soybean intended for human food use, the key components considered important for compositional analysis include the proximates (moisture, crude protein, fat, ash, fibre), amino acids, fatty acids, minerals, vitamins, isoflavones, phospholipids, sterols, saponins and the anti-nutrients phytic acid, trypsin inhibitors, stachyose, raffinose and lectins, (OECD, 2012). It is noted that the OECD recommendations for analysis of phospholipids, sterols and saponins are not emphasised in the previous version of the consensus document (OECD, 2001) and that the compositional studies reported by the Applicant were done in 2011 and therefore were based on this previous version. It is emphasised that in all of the OECD Consensus documents the recommended choice of analytes in relation to the food use of a crop is a suggestion and not a mandatory requirement to demonstrate safety. The inclusion of the extra analytes in the recent version of the OECD document does not imply that the safety of a compositional consideration would be compromised if these analytes were not considered, and certainly does not negate the conclusions of safety assessments associated with the numerous approvals for food from GM soybean lines in which the compositional analyses were based on the recommendation of the OECD 2001 document.

Analyses for key components were done on seed and forage. In general, soybean is cultivated for the production of seed, which is used as a source of both human food and animal feed, and is only infrequently used as a forage crop for livestock. As there are no human food products derived from forage, only the results of the compositional analyses for seed and its processed fractions are presented in this report. The compositional analyses for forage focussed only on proximates, fibre and two minerals (Ca and P). The analyses

showed that there was no significant difference between 81419 and 'Maverick' for mean levels of any of the analytes measured.

Study submitted:

Fast BJ, Johnson TY. 2012. Nutrient composition of a transformed soybean cultivar expressing Cry1Ac, Cry1F, and PAT: event DAS-81419-2. *Study ID 110000.01*, Dow AgroSciences LLC, Indianapolis, IN

5.2 Study design and conduct

The test (verified 81419, seed of T₄ lineage), and control (verified 'Maverick') lines were grown under similar conditions at 10 field sites across North America⁵ during 2011. The sites were representative of where soybean is commercially grown. 'Maverick' is the original transformed line and therefore represents the near-isogenic control line for the purposes of the comparative analyses (see Section 2.1). Six different commercial, non-GM soybean lines were also grown under the same conditions in order to generate a reference range for each analyte. The reference varieties were randomized across sites in a balanced incomplete-block design with three reference varieties at each site and each reference variety present at five sites.

All lines were treated with conventional pesticides (termed 'untreated' for the purposes of this safety assessment). There were no specific pesticide treatments.

Seed and forage from soybean 81419 and 'Maverick' were harvested from all replicated plots and analysed for composition. Forage was collected at the R3 plant growth stage, and seed was harvested at physiological maturity (R8 stage) (refer to Table 3 for reference to growth stages). Methods of composition analysis were based on internationally recognised procedures (e.g. those of the Association of Analytical Communities - AOAC), methods specified by the manufacturer of the equipment used for analysis, or other published methods.

Data were transformed into Statistical Analysis Software⁶ (SAS) data sets and analysed using SAS. A least squares mean (LSM) value was generated and used for each analyte comparison, and standard error and minimum and maximum values were also calculated for each analyte. The calculated means are summarised in Tables 4 – 10. Analysis of Variance was used for over-all analysis. The significance of an overall treatment effect was estimated using an F-test, while paired contrasts were made using t-tests. Probability values were adjusted using False Discovery Rate (FDR) procedures to improve discrimination of true differences (Benjamini and Hochberg, 1995; OECD, 2012). In assessing the significance of any difference between the mean analyte value for soybean 81419 and 'Maverick' an FDR-adjusted P-value of 0.05 was used. This means that approximately 5% of statistically significant differences are expected to occur due to chance alone.

The results for the treatments were compared to

- The reference range (min – max) compiled from the results of the six non-GM reference lines, in order to assess whether any differences were likely to be biologically meaningful.

⁵ The 10 sites were: Richland and Atlantic, IA; York, NE; Fisk and La Plata, MO; Wyoming and Carlyle, IL; Brunswick, NE; Germanville, PA; Frankfort, IN

⁶ SAS website - <http://www.sas.com/technologies/analytics/statistics/stat/index.html>

- A combined literature range for each analyte, compiled from published literature⁷. Any mean value for a soybean 81419 analyte that fell within the combined literature range was considered to be within the normal variability of commercial soybean cultivars even if the mean value was statistically different from the 'Maverick' control. It is noted, however, that information in the published literature is limited and is unlikely to provide a broad reflection of the natural diversity that occurs within soybean. Therefore, even if means fall outside the published range, this is not necessarily a concern.

5.3 Seed composition

5.3.1 Proximates and fibre

Results of the proximate and fibre analysis are shown in Table 4. For fat, ash and moisture, there was a statistically significant overall treatment effect but the FDR-adjusted P value for pairwise t-test comparisons of the same analytes was not significant. In the case of all analytes, the means were within both the reference range and the combined literature range.

Table 4: Mean percentage of proximates and fibre in seed from 'Maverick' and DAS-81419-2

Analyte	'Maverick'	81419	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range	Combined literature range
Protein (%dw)	37.9	38.1	NS		34.4 – 46.0	32 – 48.4
Fat (%dw)	18.2	17.7	0.008	NS	4.1 – 22.7	8.1 – 24.7
Ash (%dw)	5.06	5.18	0.026	NS	3.79 – 6.79	3.8 – 6.9
Moisture (%fw)	12.3	11.7	0.019	NS	7.91 – 22.7	4.7 – 34.4
Carbohydrate (%dw) ¹	38.8	39.0	NS		29.9 – 40.8	29.3 – 50.2
ADF ² (%dw)	15.3	15.2	NS		10.2 – 21.0	7.81 – 26.6
NDF ³ (%dw)	17.5	17.7	NS		10.6 – 22.6	8.53 – 23.9
Total dietary fibre (%dw)	23.8	24.0	NS		16.1 – 29.5	Not reported

¹ Carbohydrate calculated as 100% - (protein %dw + fat %dw + ash %dw)

² ADF = acid detergent fibre

³ NDF = neutral detergent fibre

5.3.2 Fatty acids

The levels of 22 fatty acids were measured. Of these, the following 14 were below the level of quantification (LOQ) in more than 50% of samples and were therefore not statistically analysed: caprylic (8:0), capric (10:0), lauric (12:0), myristic (14:0), myristoleic (14:1), pentadecanoic (15:0), pentadecenoic (15:1), palmitoleic (16:1), heptadecanoic (17:0), heptadecenoic (17:1), γ -linolenic (18:3), eicosadienoic (20:2), eicosatrienoic (20:3), and arachidonic (20:4). Results for the remaining eight fatty acids are given in Table 5 and can be summarised as follows:

⁷ References included: Kakade et al (1972); Iskander (1987); Hartwig & Kilen (1991); Padgett et al (1996); Taylor et al (1999); OECD (2001); McCann et al (2005); Harrigan et al (2007); Bilyeu et al (2008); Lundry et al (2008); Berman et al (2009); Berman et al (2010); Harrigan et al (2010); ILSI (2011)

- There was no significant difference between 'Maverick' and soybean 81419 in terms of the levels of stearic, oleic, linoleic, arachidic and behenic acids.
- For palmitic acid, the level in 81419 was significantly higher than the level in 'Maverick' both at the overall treatment level and at the pairwise comparison. The mean for 81419 was within the combined literature range and just slightly outside (higher) than the reference range; it does not raise a concern.
- The mean levels of linolenic and eicosenoic acids were significantly different from the levels in 'Maverick'; overall but the FDR-adjusted P value for pairwise t-test comparisons of the same analytes was not significant. Both 81419 means were within the reference range and combined literature range.

Table 5: Mean percentage composition, relative to total fat, of major fatty acids in seed from 'Maverick' and DAS-81419-2

Fatty acid	'Maverick' (% total)	81419 (% total) ¹	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range % total	Combined literature range % total
Palmitic (16:0)	11.12	11.56	<0.001	<0.001	9.12 – 11.53	1.4 – 15.7
Stearic (18:0)	4.40	4.46	NS		3.19 – 5.07	2.59 – 5.88
Oleic (18:1)	21.6	21.2	NS		18.8 – 24.6	2.6 – 45.6
Linoleic (18:2)	54.1	53.8	NS		53.6 – 57.5	7.58 – 58.8
Linolenic (18:3)	7.97	8.17	0.04	NS	6.58 – 9.88	1.27 – 12.52
Arachidic (20:0)	0.319	0.325	NS	<	0.254 – 0.383	0.038 – 0.57
Eicosenoic (20:1)	0.157	0.153	0.048	NS	< LOQ – 0.196	<LOQ – 0.35.
Behenic (22:0)	0.317	0.321	NS		0.277 – 0.390	0.043 – 0.65

¹ Mauve shading represents 81419 means where a pairwise contrast t-test showed a significantly lower value than for the 'Maverick' mean, while orange shading represents 81419 means that were significantly higher than 'Maverick'

5.3.3 Amino acids

Levels of 18 amino acids were measured. Since asparagine and glutamine are converted to aspartate and glutamate respectively during the analysis, levels for aspartate include both aspartate and asparagine, while glutamate levels include both glutamate and glutamine. Results of the analysis are given in Table 6.

The only significant difference (overall) between 81419 and 'Maverick' was in the mean level of phenylalanine but the FDR-adjusted P value for a pairwise t-test comparison of the same analyte was not significant. The mean phenylalanine level in 81419 was within both the reference range and combined literature range.

Table 6: Mean percentage composition, relative to total amino acids, of amino acids in seed from 'Maverick' and DAS-81419-2

Amino Acid	'Maverick' %total amino acid	81419 %total amino acid ¹	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range %total amino acid	Combined literature range %total amino acid
Alanine	4.57	4.59	NS		4.28 – 4.75	4.16 – 4.74
Arginine	7.49	7.44	NS		7.22 – 8.20	6.41 – 8.41
Aspartate	11.51	11.48	NS		9.99 – 11.74	11.37 – 12.68
Cysteine	01.61	1.629	NS		1.240 – 1.792	1.02 – 1.87
Glutamate	17.32	17.23	NS		17.04 – 18.56	17.71 – 20.48
Glycine	4.50	4.50	NS		4.14 – 4.54	4.19 – 4.62
Histidine	2.714	2.703	NS		2.434 – 2.776	2.49 – 2.89
Isoleucine	4.80	4.80	NS		4.61 – 4.99	4.13 – 5.11
Leucine	7.65	7.64	NS		7.45 – 7.98	7.46 – 8.29
Lysine	6.32	6.44	NS		5.61 – 7.29	6.23 – 7.38
Methionine	1.419	1.435	NS		1.223 – 1.618	1.18 – 1.71
Phenylalanine	5.15	5.11	0.009	NS	4.88 – 5.37	4.91 – 5.44
Proline	5.14	5.20	NS		4.80 – 6.02	4.75 – 5.62
Serine	5.13	5.16	NS		4.81 – 5.53	3.25 – 6.04
Threonine	4.19	4.20	NS		3.86 – 4.25	3.15 – 4.24
Tryptophan	1.519	1.524	NS		1.271 – 1.686	0.95 – 1.49
Tyrosine	3.97	3.96	NS		3.82 – 4.16	2.62 – 3.72
Valine	4.98	4.96	NS		4.59 – 5.18	4.28 – 5.57

5.3.4 Isoflavones

In total, there are 12 different soybean isoflavone isomers, namely three parent isoflavones (genistein, daidzein and glycitein), their respective β -glucosides (genistin, daidzin, and glycitin), and three β -glucosides each esterified with either malonic or acetic acid (Messina, 2005). The parent isoflavones are also referred to as free or aglycon isoflavones, while the glucosides and their esters are also referred to as conjugated isoflavones.

The Applicant used an AOAC International method (AOAC, 2005), to measure the levels of the three parent isoflavones and the conjugates in seed from soybean 44406 and the control. Results are expressed in aglycon equivalents by summing the concentrations of the aglycons and the aglycon equivalents of the corresponding glucosides. The mean level of total glycitein overall was significantly lower in 81419 compared to the control but the FDR-adjusted P value for a pairwise t-test comparison of the same analyte was not significant. Mean levels of all analytes were within both the reference range and combined literature range.

Table 7: Mean weight ($\mu\text{g/g}$ dry weight expressed as aglycon equivalents) of isoflavones in 'Maverick' and DAS-81419-2 seed

Isoflavone	'Maverick' ($\mu\text{g/g dw}$)	81419 ($\mu\text{g/g dw}$)	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range ($\mu\text{g/g dw}$)	Combined literature range ($\mu\text{g/g/dw}$)
Total Daidzein (aglycon equivalents)	950	932	NS		585 - 1460	25 - 2453
Total Genistein (aglycon equivalents)	1296	1276	NS		753 - 1950	28 - 2837
Total Glycitein (aglycon equivalents)	197	180	0.002	NS	40.3 - 259	15 - 349

5.3.5 Anti-nutrients

Levels of key anti-nutrients are given in Table 8. Overall, there were no significant differences between the 81419 means and the control means for any of the analytes. All means were within both the reference range and the combined literature range.

Table 8: Mean levels of anti-nutrients in 'Maverick' and DAS-81419-2 seed.

Anti-nutrient	'Maverick'	81419	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range	Combined literature range
Lectin (Haemagglutinin Units/mg protein dw)	30.8	32.2	NS		7.89 – 45.2	0.1 - 323
Phytic acid (%dw)	1.22	1.24	NS		0.678 – 1.71	0.41 – 2.74
Raffinose (%dw)	0.750	0.766	NS		0.570 – 1.16	0.21 – 1.85
Stachyose (%dw)	3.68	3.69	NS		3.01 – 5.28	1.21 – 6.65
Trypsin inhibitor (trypsin inhibitor units/mg)	29.1	30.2	NS		19.5 – 53.8	18.14 – 118.68

5.3.6 Minerals

Levels of 10 minerals were measured. Average sodium levels were below the LOQ. The means for the remaining nine minerals are given in Table 9 and show:

- There was no significant difference between the mean of 81419 and the control for any analyte.
- The mean potassium levels for both lines were outside (lower than) the literature range but within the reference range.

Table 9: Mean values for mineral levels in seed from 'Maverick' and DAS-81419-2.

Mineral	'Maverick' (mg/100g dw)	81419 (mg/100g dw) ¹	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range (mg/100g dw)	Combined literature range (mg/100g dw)
Calcium	270	267	NS		181 - 339	116 - 510
Copper	1.31	1.33	NS		0.693 – 1.86	0.632 – 1.092
Iron	9.56	810.26	NS		6.33 - 151	3.73 – 10.95
Magnesium	233	232	NS		205 - 278	219 - 312
Manganese	2.64	2.67	NS		2.22 – 7.18	2.52 – 3.876
Phosphorus	607	619	NS		471 - 759	506 - 935
Potassium	1799	1819	NS		1650 - 2050	1868 - 2510
Selenium	468	507	NS		<LOQ - 3060	Not reported ^{4,5,3}
Zinc	4.53	4.63	NS		3.12 – 6.33	4.98 – 7.578

5.3.7 Vitamins (including Vitamin E compounds)

Levels of 13 vitamins were measured. Those of Vitamin A (β -carotene) and β -tocopherol were generally below the LOQ and were not statistically analysed. The means for the remaining 11 vitamins are given in Table 10. Overall there were no significant differences between the 81419 means and the control means for Vitamins B₁, B₂, B₃, B₆, B₉ or C or for δ -tocopherol or total tocopherol. The overall analysis showed a significant difference between the 81419 and control means for vitamin B₅ and γ -tocopherol but the FDR-adjusted P value for a pairwise t-test comparison of the same analytes was not significant. All means were within literature ranges (when available) and/or within ranges for the reference varieties.

Table 10: Mean weight (mg/g dry weight) of vitamins in seed from 'Maverick' and DAS-81419-2

Vitamin	'Maverick' (mg/kg dw)	81419 (mg/kgdw)	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range (mg/kg dw)	Combined literature range (mg/kg dw)
Vitamin B ₁ (thiamin)	3.51	3.43	NS		1.82 – 4.92	1.01 – 2.54
Vitamin B ₂ (riboflavin)	3.40	3.51	NS		2.42 – 5.00	1.9 – 3.21
Vitamin B ₃ (niacin)	25.0	25.6	NS		20.5 – 29.0	Not reported
Vitamin B ₅ (pantothenic acid)	14.8	14.0	0.004	NS	8.97 – 18.0	Not reported
Vitamin B ₆ (pyridoxine)	5.23	5.18	NS		3.01 – 6.36	Not reported
Vitamin B ₉ (folic acid)	4.21	4.15	NS		2.94 – 5.59	2.38 – 4.70
Vitamin C (ascorbic acid)	141	133	NS		49.2 - 210	Not reported
α -tocopherol	14.3	13.6	NS		6.51 – 25.0	1.93 – 84.9
γ -tocopherol	69.4	74.7	0.004	NS	49.7 – 104.0	Not reported
δ -tocopherol	168	172	NS		77.5 - 240	Not reported

Vitamin	'Maverick' (mg/kg dw)	81419 (mg/kgdw)	Overall treatment effect (P value)	Pairwise comparison (FDR-adjusted P)	Reference range (mg/kg dw)	Combined literature range (mg/kg dw)
Total tocopherol	237.5	247.1	NS		150.5 – 299.9	Not reported

5.3.8 Summary of analysis of key components

Mean values for a total of 62 analytes are presented in Tables 4 – 10. For 52 of these analytes an overall analysis showed there were no significant differences between the levels found in seed of soybean DAS-81419 and 'Maverick'. However, for only one analyte – palmitic acid – of the 10 showing an overall significant difference did a pairwise comparison using an FDR-adjusted probability show a significant difference (with the mean level being significantly higher in 81419). The palmitic acid mean for 81419 was within the combined literature range and just slightly outside (higher) than the reference range; and does not raise a safety concern.

5.4 Conclusion

The compositional analyses do not indicate any differences of biological significance between seed from soybean DAS-81419-2 and the non-GM control 'Maverick'. In an overall analysis, statistically significant differences were noted in a few constituents. However the differences were typically small, and all mean values were within both the reference range obtained for non-GM reference varieties grown at the same time and (where it exists) the literature range. Any observed differences are therefore considered to represent the natural variability that exists within soybean.

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 well-being. In most cases, this can be achieved through an understanding of the genetic modification and its consequences, together with an extensive compositional analysis of the food.

If the compositional analysis indicates biologically significant changes to the levels of certain nutrients in the GM food, additional nutritional assessment should be undertaken to assess the consequences of the changes and determine whether nutrient intakes are likely to be altered by the introduction of such foods into the food supply.

Where a GM food has been shown to be compositionally equivalent to conventional varieties, as is the case for soybean line 44406, the evidence to date indicates that feeding studies using target livestock species will add little value to the safety assessment and generally are not warranted (OECD, 2003; EFSA, 2008).

Soybean DAS-81419-2 is the result of a simple genetic modification to confer insect resistance with no intention to significantly alter nutritional parameters in the food. In addition, the extensive compositional analyses of seed that have been undertaken to demonstrate the nutritional adequacy of line 81419, indicate it is equivalent in composition to conventional soybean cultivars. The introduction of soybean line DAS-81419-2 into the food supply is therefore expected to have little nutritional impact.

References⁸

- Adang, M.J., Staver, M.J., Rocheleau, T.A., Leighton, J., Barker, R.F. and Thompson, D.V. (1985) Characterized full length and truncated plasmid clones of the crystal protein of *Bacillus thuringiensis* subsp. *kurstaki* HD-73 and their toxicity to *Manduca sexta*. *Gene* 36:289-300.
- Agostino, M. (2013) Introduction to the BLAST suite and BLASTN. In: *Practical Bioinformatics*. Chapter 3. Garland Science, Taylor & Francis Group, New York, pp. 47-71.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25(17):3389-3402.
- AOAC. (2005) AOAC Official Method 2001.10: Determination of Isoflavones in Soy and Selected Food Containing Soy. In: *Official Methods of Analysis of AOAC International*. 18th ed, AOAC International, Gaithersburg.
- Aronson, A.I. and Shai, Y. (2001) Why *Bacillus thuringiensis* insecticidal toxins are so effective: unique features of their mode of action. *FEMS Microbiology Letters* 195:1-8.
- Astwood, J.D. and Fuchs, R.L. (1996) Allergenicity of foods derived from transgenic plants. *Monographs in Allergy* 32 (Highlights in Food Allergy):105-120.
- Barker, R.F., Idler, K.B., Thompson D.V. and Kemp, J.D. (1983) Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. *Plant Molecular Biology* 2:335-350.
- Baxevanis, A.D. (2005) Assessing Pairwise Sequence Similarity: BLAST and FASTA. In: Baxevanis, A.D. and Ouellette, B.F.F. eds. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. Chapter 11. John Wiley & Sons, Inc., pp. 295-324.
- Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological* 57(1):289-300.
- Berman, K.H., Harrigan, G.G., Riordan, S.G., Nemeth, M.A., Hanson, C., Smith, M., Sorbet, R., Zhu, E. and Ridley, W.P. (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(23):11360-11369.
- Berman, K.H., Harrigan, G.G., Riordan, S.G., Nemeth, M.A., Hanson, C., Smith, M., Sorbet, R., Zhu, E. and Ridley, W.P. (2010) Compositions of forage and seed from second-generation glyphosate-tolerant soybean MON 89788 and insect-protected soybean MON 87701 from Brazil are equivalent to those of conventional soybean (*Glycine max*). *Journal of Agricultural and Food Chemistry* :
- Betz, F.S., Hammond, B.G. and Fuchs, R.L. (2000) Safety and Advantages of *Bacillus thuringiensis*-Protected Plants to Control Insect Pests. *Regulatory Toxicology and Pharmacology* 32:156-173.

⁸ All website references were current as at 31 October 2013

Bilyeu, K.D., Zeng, P., Coello, P., Zhang, Z.J., Krishnan, H.B., Bailey, A., Beuselinck, P.R. and Polacco, J.C. (2008) Quantitative conversion of phytate to inorganic phosphorus in soybean seeds expressing a bacterial phytase. *Plant Physiology* 146(2):468-477.

Bradbury, J.F. (1986) *Guide to plant pathogenic bacteria*. Cambridge News Ltd, Aberystwyth, p190-197.

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

Broderick, N.A., Raffa, K.F. and Handelsman, J. (2006) Midgut bacteria required for *Bacillus thuringiensis* insecticidal activity. *Proceedings of the National Academy of Sciences* 103:15196-15199.

Brookes, G. and Barfoot, P. (2009) *GM crops: global socio-economic and environmental impacts 1996-2007*. PG Economics Ltd.
<http://www.pgeconomics.co.uk/pdf/2009globalimpactstudy.pdf>.

Cardineau, G.A., Stelman, S.J. and Narva, K.E. (2001) Plant-optimized genes encoding pesticidal toxins. (US 6218188 B1): <http://patent.ipexl.com/US/6218188.html>.

Codex (2003) *Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants*. Report No. CAC/GL 45-2003, Codex Alimentarius.
http://www.codexalimentarius.net/web/standard_list.do?lang=en.

Damgaard, P.H., Granum, P.E., Bresciani, J., Torregrossa, M.V., Eilenberg, J. and Valentino, L. (1997) Characterization of *Bacillus thuringiensis* isolated from infections in burn wounds. *FEMS Immunology & Medical Microbiology* 18:47-53.

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

Deblaere, R., Reynaerts, A., Höfte, H., Hernalsteens, J.-P., Leemans, J. and van Montagu, M. (1987) Vectors for cloning in plant cells. *Methods in Enzymology* 153:277-292.

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

EFSA. (2007) Opinion of the Scientific Panel on Genetically Modified Organisms on an Application (Reference EFSA-GMO-UK-2004-04) for the Placing on the Market of Glufosinate Tolerant Genetically Modified Rice LLRICE62 for Food and Feed Uses, Import and Processing under Regulation (EC) No 1829/2003 from Bayer CropScience GmbH. *EFSA Journal* 588:1-25.

EFSA. (2008) Safety and nutritional assessment of GM plants and derived food and feed: The role of animal feeding trials. *Food and Chemical Toxicology* 46:S1-S70, doi:10.1016/j.fct.2008.02.008.

EPA (1998) *RED Facts: Bacillus thuringiensis*. Report No. EPA 738-F-98-001., Environmental Protection Agency.

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, January 22 - 25, 2001, Rome.

FAOSTAT. (2013) Online database of the Food and Agriculture Organization of the United Nations. <http://faostat3.fao.org/faostat-gateway/go/to/home/E>.

Federici, B.A. (1999) *Bacillus thuringiensis* in Biological Control. In: Bellows, T.W. and Fisher, T.S. eds. *Handbook of Biological Control*. Chapter 21. Academic Press, San Diego, pp. 575-593.

FSANZ (2005) *Application A518 - Food derived from Insect Protected cotton line MXB-13*. Report prepared by Food Standards Australia New Zealand. <http://www.foodstandards.gov.au/code/applications/pages/applicationa518foodd2314.aspx>.

Gao, S., Dagnaes-Hansen, F., Nielsen, E.J.B., Wengel, J., Besenbacher, F., Howard, K.A. and Kjems, J. (2009) The effect of chemical modification and nanoparticle formulation on stability and biodistribution of siRNA in mice. *Molecular Therapy* 17(7):1225-1233.

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

Gao, Y., Schafer, B.W., Collins, R.A., Herman, R.A., Xu, X., Gilbert, J.R., Ni, W., Langer, V.L. and Tagliani, L.A. (2004) Characterization of Cry34Ab1 and Cry35Ab1 insecticidal crystal proteins expressed in transgenic corn plants and *Pseudomonas fluorescens*. *Journal of Agricultural and Food Chemistry* 52:8057-8065.

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

Gilroy, T.E. and Wilcox, E.R. (1992) Hybrid *Bacillus thuringiensis* gene, plasmid and transformed *Pseudomonas fluorescens*. (US 5128130): <http://www.patentlens.net/daisy/patentlens/patentlens.html>.

Goodman, R.E. (2006) Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. *Molecular Nutrition and Food Research* 50:655-660.

Green, M., Heumann, M., Sokolow, R., Foster, L.R., Bryant, R. and Skeels, M. (1990) Public health implications of the microbial pesticide *Bacillus thuringiensis*: an epidemiological study, Oregon 1985 - 86. *American Journal of Public Health* 80:848-852.

Grey, D. (2006) Growing soybeans in northern Victoria. State of Victoria, Department of Primary Industries, http://aof.clients.squiz.net/data/assets/pdf_file/0008/7667/Growing_Soybeans_in_Northern_Victoria.pdf.

Harrigan, G.G., Glenn, K.C. and Ridley, W.P. (2010) Assessing the natural variability in crop composition. *Regulatory Toxicology and Pharmacology* 58:S13-S20; doi:10.1016/j.yrtph.2010.08.023.

Harrigan, G.G., Ridley, W.P., Riordan, S., Nemeth, M.A., Sorbet, R., Trujillo, W.A., Breeze, M. and Schneider, R.W. (2007) Chemical composition of glyphosate-tolerant soybean 40-3-2

grown in Europe remains equivalent with that of conventional soybean (*Glycine max* L.). *Journal of Agricultural and Food Chemistry* 55(15):6160-6168.

Hartwig, E.E. and Kilen, T.C. (1991) Yield and composition of soybean seed from parents with different protein, similar yield. *Crop Science* 31(2):290-292.

Henikoff, S. and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences* 89:10915-10919.

Herman, R.A., Woolhiser, M.M., Ladics, G.S., Korjagin, V.A., Schafer, B.W., Storer, N.P., Green, S.B. and Kan, L. (2007) Stability of a set of allergens and non-allergens in simulated gastric fluid. *International Journal of Food Sciences and Nutrition* 58:125-141.

Hernandez, E., Ramiisse, F., Ducoureau, J.-P., Cruel, T. and Cavallo, J.-D. (1998) *Bacillus thuringiensis* subsp. *konkukian* (serotype H34) superinfection: case report and experimental evidence of pathogenicity in immunosuppressed mice. *Journal of Clinical Microbiology* 36:2138-2139.

Hérouet, C., Esdaile, D.J., Mallyon, B.A., Debruyne, E., Schulz, A., Currier, T., Hendrickx, K., van der Klis, R.-J. and Rouan, D. (2005) Safety evaluation of the phosphinothricin acetyltransferase proteins encoded by the *pat* and *bar* sequences that confer tolerance to glufosinate-ammonium herbicide in transgenic plants. *Regulatory Toxicology and Pharmacology* 41:134-149.

Hofmann, C., Vanderbruggen, H., Höfte, H., Van Rie, J., Jansens, S. and Van Mellaert, H. (1988) Specificity of *Bacillus thuringiensis* δ -endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target insect midguts. *Proceedings of the National Academy of Sciences* 85:7844-7848.

Hood, E.E., Helmer, G.L., Fraley, R.T. and Chilton, M.-D. (1986) The hypervirulence of *Agrobacterium tumefaciens* A281 is encoded in a region of pTiBo542 outside of T-DNA. *Journal of Bacteriology* 168(3):1291-1301.

ILSI (2011) *International Life Sciences Institute Crop Composition Database Version 4.0*. <http://www.cropcomposition.org/query/index.html>.

Iowa State University. (2009) Soybean growth stages. Soybean Extension and Research Program, Department of Agronomy, Iowa State University, Ames, IA. http://extension.agron.iastate.edu/soybean/production_growthstages.html.

Iskander, F.Y. (1987) Maturation stage and mineral content in soybeans. *Food Chemistry* 24:29-35.

Jackson, S.G., Goodbrand, R.B., Ahmed, R. and Kasatiya, S. (1995) *Bacillus cereus* and *Bacillus thuringiensis* isolated in a gastroenteritis outbreak investigation. *Letters in Applied Microbiology* 21:103-105.

James, A.T. and Rose, I.A. (2004) Integrating crop physiology into the Australian soybean improvement program. In: *4th International Crop Science Congress, Brisbane, Australia, September 2004*. http://www.cropspace.org.au/icsc2004/poster/3/4/6/327_jamesat.htm#TopOfPage.

Jensen, O.N., Podtelejnikov, A.V. and Mann, M. (1997) Identification of the components of simple protein mixtures by high-accuracy peptide mass mapping and database searching. *Annals of Chemistry* 69(23):4741-4750.

- Kakade, M.L., Simons, N.R., Liener, I.E. and Lambert, J.W. (1972) Biochemical and nutritional assessment of different varieties of soybeans. *Journal of Agricultural and Food Chemistry* 20(1):87-90.
- Kimber, I., Kerkvliet, N.I., Taylor, S.L., Astwood, J.D., Sarlo, K. and Dearman, R.J. (1999) Toxicology of protein allergenicity: prediction and characterization. *Toxicological Sciences* 48(2):157-162.
- Krell, R. (1996) *Value-Added Products from Beekeeping*. Chapter 3: Pollen, FAO Agricultural Services Bulletin No. 124, Food and Agriculture Organization of the United Nations, available online at <http://www.fao.org/docrep/w0076e/w0076e00.htm#con>.
- Kützner, H.J. (1981) The Family Streptomycetaceae. In: Starr, M.P., Stolp, H., Trüper, H.G., Ballows, A., and Schlegel, H.G. eds. *The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*. Springer Verlag, Berlin, pp. 2028-2090.
- Lundry, D.R., Ridley, W.P., Meyer, J.J., Riordan, S.G., Nemeth, M.A., Trujillo, W.A., Breeze, M.L. and Sorbet, R. (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(12):4611-4622.
- McCann, M.C., Liu, K., Trujillo, W.A. and Dobert, R.C. (2005) Glyphosate-tolerant soybeans remain compositionally equivalent to conventional soybeans (*Glycine max* L.) during three years of field testing. *Journal of Agricultural and Food Chemistry* 53:5331-5335.
- McClintock, J.T., Schaffer, C.R. and Sjoblad, R.D. (1995) A comparative review of the mammalian toxicity of *Bacillus thuringiensis*-based pesticides. *Pesticide Science* 45:95-105.
- Mendelsohn, M., Kough, J., Vaituzis, Z. and Matthews, K. (2003) Are *Bt* crops safe? *Nature Biotechnology* 21:1003-1009.
- Messina, M. (2005) Isoflavones. In: Coates, P., Blackman, M.R., Cragg, G., Levine, M., Moss, J., and White, J. eds. *Encyclopedia of Dietary Supplements*. Informa Healthcare.
- Metcalfe, D.D., Astwood, J.D., Townsend, R., Sampson, H.A., Taylor, S.L. and Fuchs, R.L. (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Critical Reviews in Food Science and Nutrition* 36 Suppl:S165-S186.
- Miletich, J.P. and Broze Jr., G.J. (1990) β Protein C is not glycosylated at asparagine 329. *The Journal of Biological Chemistry* 265:11397-11404.
- Nester, E.W., Thomashow, L.S., Metz, M. and Gordon, M. (2002) *100 years of Bacillus thuringiensis: A critical scientific assessment. A report from the American Academy of Microbiology based on a colloquium held on 16-18 November, 2001, in Ithaca, New York*.
- Norris, S.R., Meyer, S.E. and Callis, J. (1993) The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. *Plant Molecular Biology* 21(5):895-906.
- NPTN. (2000) *Bacillus thuringiensis* (Technical Fact Sheet). National Pesticide Telecommunications Network, Oregon State University & U.S. Environmental Protection Agency, <http://npic.orst.edu/factsheets/BTtech.pdf>.

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

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.

[http://www.olis.oecd.org/olis/2003doc.nsf/LinkTo/NT0000426A/\\$FILE/JT00147696.PDF](http://www.olis.oecd.org/olis/2003doc.nsf/LinkTo/NT0000426A/$FILE/JT00147696.PDF).

OECD (2007) *Consensus Document on Safety Information on Transgenic Plants Expressing Bacillus thuringiensis - Derived Insect Control Proteins*. Report No.

ENV/JM/MONO(2007)14, Organisation for Economic Co-operation and Development.

[http://www.olis.oecd.org/olis/2007doc.nsf/LinkTo/NT00002DF6/\\$FILE/JT03230592.PDF](http://www.olis.oecd.org/olis/2007doc.nsf/LinkTo/NT00002DF6/$FILE/JT03230592.PDF).

OECD (2012) *Revised consensus document on compositional considerations for new varieties of soybean [Glycine max (L.) Merr.]: Key food and feed nutrients, antinutrients, toxicants and allergens*. Report No. ENV/JM/MONO(2012)24, Organisation for Economic Co-operation and Development.

<http://www.oecd.org/env/chemicalsafetyandbiosafety/biosafety-biotrack/consensusdocumentsfortheworkonthesafetyofnovelfoodsandfeedsplants.htm>.

Orlando, R. and Yang, Y. (1998) Analysis of Glycoproteins. In: Larsen, B.S. and McEwen, C.N. eds. *Mass Spectrometry of Biological Materials*. 2nd ed, Chapter 9. Marcel Dekker, pp. 216-246.

Padgett, S.R., Taylor, N.B., Nida, D.L., Bailey, M.R., MacDonald, J., Holden, L.R. and Fuchs, R.L. (1996) The composition of glyphosate-tolerant soybean seeds is equivalent to that of conventional soybeans. *Journal of Nutrition* 126(3):702-716.

Parker, M.W. and Feil, S.C. (2005) Pore-forming protein toxins: from structure to function. *Progress in Biophysics and Molecular Biology* 88:91-142.

Pearson, W.R. (2000) Flexible sequence similarity searching with the FASTA3 program package. In: Misener, S. and Krawetz, S.A. eds. *Methods in Molecular Biology, Volume 132: Bioinformatics Methods and Protocols*. Chapter 10. Human Press Inc., Totowa, NJ, pp. 185-219.

Samples, J.R. and Buettner, H. (1983) Ocular infection caused by a biological insecticide. *The Journal of Infectious Diseases* 148:614.

Sanahuja, G., Banakar, R., Twyman, R.M., Capell, T. and Christou, P. (2011) *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnology Journal* 9:283-300.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C. and Jackson, S.A. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463(7278):178-183.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D.R. and Dean, D.H. (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews* 62(3):775-806.

Shelton, A.M., Zhao, J.-Z. and Roush, R.T. (2009) Economic, ecological, food safety and social consequences of the deployment of Bt transgenic plants. *Annual Review of Entomology* 47:845-881.

Shurtleff, W. and Aoyagi, A. (2007) History of soybean crushing: Soy oil and soybean meal. In: *History of soybeans and soyfoods: 1100 B.C. to the 1980s*. SoyInfo Center, available online at http://www.soyinfocenter.com/HSS/soybean_crushing1.php, Lafayette, California.

Sleper, D.A., Nickell, C.D., Noel, G.R., Cary, T.R., Thomas, D.J., Clark, K.M. and Rao Arelli, A.P. (1998) Registration of 'Maverick' soybean. *Crop Science* 38:549.

Taylor, N.B., Fuchs, R.L., MacDonald, J., Shariff, A.R. and Padgett, S.R. (1999) Compositional analysis of glyphosate-tolerant soybeans treated with glyphosate. *Journal of Agricultural and Food Chemistry* 47(10):4469-4473.

Taylor, S.L. (2002) Protein allergenicity assessment of foods produced through agricultural biotechnology. *Annual Review of Pharmacology and Toxicology* 42:99-112.

The American Soybean Association. (2011) Soy Stats 2011. <http://www.soystats.com/2011/Default-frames.htm>.

Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.-J., Glatt, C.M., Hadfield, N., Hatzos, C., Hefle, S.L., Heylings, J.R., Goodman, R.E., Henry, B., Herouet, C., Holsapple, M., Ladics, G.S., Landry, T.D., MacIntosh, S.C., Rice, E.A., Privalle, L.S., Steiner, H.Y., Teshima, R., Van Ree, R., Woolhiser, M. and Zawodny, J. (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., Bannon, G., Hefle, S., Herouet, C., Holsapple, M., Ladics, G., MacIntosh, S. and Privalle, L. (2005) In silico methods for evaluating human allergenicity to novel proteins: International Bioinformatics Workshop Meeting Report February 23 - 24, 2005. *Toxicological Sciences* 88:307-310.

Thomas, K., MacIntosh, S., Bannon, G., Herouet-Guicheney, C., Holsapple, M., Ladics, G., McClain, S., Vieths, S., Woolhiser, M. and Privalle, L. (2009) Scientific advancement of novel protein allergenicity evaluation: An overview of work from the HESI Protein Allergenicity Technical Committee (2000 - 2008). *Food and Chemical Toxicology* 47:1041-1050.

Thompson, C.J., Movva, N.R., Tizard, R., Cramer, R., Davies, J.E., Lauwereys, M. and Botterman, J. (1987) Characterization of the herbicide-resistance gene *bar* from *Streptomyces hygroscopicus*. *The EMBO Journal* 6:2519-2523.

USDA (2009) *Oilseeds: World markets and trade*. Circular Series FOP1-09. Foreign Agricultural Service, United States Department of Agriculture. <http://www.fas.usda.gov/oilseeds/circular/2009/January/Oilseedsfull0109.pdf>.

Valadares de Amorim, G., Whittome, B., Shore, B. and Levin, D.B. (2001) Identification of *Bacillus thuringiensis* subsp. *kurstaki* strain HD1-like bacteria from environmental and human samples after aerial spraying of Victoria, British Columbia, Canada, with Foray 48B. *Applied and Environmental Microbiology* 67:1035-1043.

Verdaguer, B., de Kochko, A., Beachy, R.N. and Fauquet, C. (1996) Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. *Plant Molecular Biology* 31:1129-1139.

Warren, R.E., Rubenstein, D., Ellar, D.J., Kramer, J.M. and Gilbert, R.J. (1984) *Bacillus thuringiensis* var *israelensis*: protoxin activation and safety. *The Lancet* 323:678-679.

Wehrmann, A., Van Vliet, A., Opsomer, C., Botterman, J. and Schulz, A. (1996) The similarities of *bar* and *pat* gene products make them equally applicable for plant engineers. *Nature Biotechnology* 14:1274-1278.

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

Wohlleben, W., Arnold, W., Broer, I., Hillemann, D., Strauch, E. and Punier, A. (1988) Nucleotide sequence of the phosphinothricin N-acetyltransferase gene from *Streptomyces viridochromogenes* Tü494 and its expression in *Nicotiana tabacum*. *Gene* 70:25-37.

Yu, J., Tan, L., Liu, Y. and Pang, Y. (2002) Phylogenetic analysis of *Bacillus thuringiensis* based on PCR amplified fragment polymorphisms of flagellin genes. *Current Microbiology* 45:139-143.

Zambryski, P. (1988) Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. *Annual Review of Genetics* 22:1-30.

Zeng, P., Vagnais, D.A., Zhang, Z. and Polacco, J.C. (2004) Refined glufosinate selection in *Agrobacterium*-mediated transformation of soybean [*Glycine max* (L.) Merrill]. *Plant Cell Reports* 22:478-482.