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

Supporting document 1

Safety assessment – Application A1073 (at Approval)

Food derived from Herbicide-tolerant Soybean Line DAS-44406-6

SUMMARY AND CONCLUSIONS

Background

A genetically modified (GM) soybean line DAS-44406-6, hereafter referred to as 44406, has been developed that is tolerant to three herbicides 2,4-dichlorophenoxyacetic acid (2,4-D), glufosinate ammonium and glyphosate. Tolerance to 2,4-D is achieved through expression of the enzyme aryloxyalkanoatedioxygenase-12 (AAD-12) encoded by the *aad-12* gene derived from the soil bacterium *Delftia acidovorans.* Tolerance to glufosinate ammonium is achieved through expression of the enzyme phosphinothricin acetyltransferase (PAT) encoded by the *pat* gene derived from another soil bacterium *Streptomyces viridochromogenes*. Tolerance to glyphosate is encoded through expression of a 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) encoded by the *2mepsps* gene from corn (*Zea mays*).

In conducting a safety assessment of food derived from soybean line 44406, a number of criteria have been addressed including: a characterisation of the transferred gene, 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 protein to be either allergenic or toxic in humans.

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

- environmental risks related to the environmental release of GM plants used in food production
- the safety of animal feed or animals fed with feed derived from GM plants
- the safety 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-44406-6 indicate that a single copy of T-DNA containing three expression cassettes for the genes *2m epsps, aad-12* and *pat* has been inserted at a single locus in Chromosome 6 of the soybean genome.

No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred to the plant. As a result of the integration of the three expression cassettes, there is a 3 bp insertion at the 5' junction region and a deletion of 4,385 bp of host DNA. The introduced genetic elements are stably inherited from one generation to the next.

Characterisation of Novel Protein

Soybean line DAS-44406-6 expresses three novel proteins, 2m EPSPS, AAD-12 and PAT. Expression analyses of the three proteins showed that all three were detected in the plant parts tested. The 2m EPSPS protein was lowest in the seed (approximately 22 µg/g dry weight) and highest in V10 – 12 leaves (approximately 2323 µg/g dry weight). AAD-12 was lowest in the roots (approximately 25 µg/g dry weight) and highest in V10 – 12 leaves (approximately 116 µg/g dry weight). PAT protein concentrations were much lower than those for AAD-12 but similarly, the V10 – 12 leaves contained the highest levels (approximately 10 µg/g dry weight) and the roots contained the lowest levels (approximately 2 µg/g dry weight).

Several studies were submitted with this application, or were known from historical data, to confirm the identity and physicochemical properties of the plant-derived 2m EPSPS, AAD-12 and PAT proteins, and demonstrated they conform in size and amino acid sequence to that expected, and do not exhibit any post-translational modification including glycosylation.

In relation to potential toxicity and allergenicity, the PAT protein has previously been demonstrated to be non-toxic to mammals and also does not exhibit any potential to be allergenic to humans.

For the 2m EPSPS and AAD-12 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 showed that both proteins are inactivated by heating. Taken together, the evidence indicates that neither 2m EPSPS nor AAD-12 are likely to be toxic or allergenic to humans.

Herbicide Metabolites

The metabolic profiles resulting from the novel protein x herbicide interactions in DAS-44406-6 have been assessed in previous applications or, in the case of PAT/glufosinate ammonium, have a significant history of use. There are no concerns that the spraying of soybean 44406 with glyphosate, 2,4-D or glufosinate ammonium would result in the production of metabolites that are not also produced in crops sprayed with the same herbicides and already used in the food supply.

Compositional Analyses

Detailed compositional analyses were done to establish the nutritional adequacy of seed from soybean line DAS-44406-6 sprayed with 2,4-D, glufosinate ammonium and glyphosate herbicides. Analyses were done of proximate (moisture, crude protein, fat, ash, fibre), amino acid, fatty acid, vitamin, mineral, phytic acid, trypsin inhibitor, lectin, isoflavone, stachyose and raffinose content. 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-44406-6 and the non-GM control 'Maverick'.

Significant differences were noted in a number of constituents. However the differences were typically small and all mean values were within both the tolerance 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. The spraying of soybean line DAS-44406-6 with 2,4-D, glufosinate ammonium and glyphosate, either alone or in combination, did not have a significant effect on seed composition.

Conclusion

No potential public health and safety concerns have been identified in the assessment of soybean line DAS-44406-6. On the basis of the data provided in the present Application, and other available information, food derived from soybean line DAS-44406-6 is considered to be as safe for human consumption as food derived from conventional soybean cultivars.

TABLE OF CONTENTS

SUMMARY AND CONCLUSIONS ..... **I**
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 5
3 MOLECULAR CHARACTERISATION **7**
3.1 METHOD USED IN THE GENETIC MODIFICATION 7
3.2 DESCRIPTION OF THE INTRODUCED GENES 8
3.3 BREEDING TO OBTAIN SOYBEAN LINE DAS-44406-6 11
3.4 CHARACTERISATION OF THE GENES IN THE PLANT 12
3.5 STABILITY OF THE GENETIC CHANGES 14
3.6 ANTIBIOTIC RESISTANCE MARKER GENES 15
3.7 CONCLUSION 15
4 CHARACTERISATION OF NOVEL PROTEINS **15**
4.1 POTENTIAL ALLERGENICITY/TOXICITY OF ORFS CREATED BY THE TRANSFORMATION PROCEDURE 16
4.2 FUNCTION AND PHENOTYPIC EFFECTS OF THE 2M EPSPS, AAD-12 AND PAT PROTEINS 17
4.3 NOVEL PROTEIN EXPRESSION IN PLANT TISSUES 19
4.4 PROTEIN CHARACTERISATION STUDIES 20
4.5 POTENTIAL TOXICITY 23
4.6 POTENTIAL ALLERGENICITY 26
4.7 CONCLUSION 28
5. OTHER NOVEL SUBSTANCES **29**
6. COMPOSITIONAL ANALYSIS **29**
6.1 KEY COMPONENTS OF SOYBEAN 29
6.2 STUDY DESIGN AND CONDUCT 30
6.3 SEED COMPOSITION 31
6.4 CONCLUSION 37
7. NUTRITIONAL IMPACT **38**
REFERENCES **38**

LIST OF TABLES

Table 1: Description of the genetic elements contained in the T-DNA of pDAB8264 9

Table 2: Location and characterisation of novel ORFs in the flanking regions 14

Table 3: Average concentration of 2m EPSPS, AAD-12 and PAT proteins in various plant parts from soybean 44406 20

Table 4: Mean percentage of proximates and fibre in seed from ‘Maverick’ and DAS-44406-6. 31

Table 5: Mean percentage composition, relative to total fat, of major fatty acids in seed from 'Maverick' and DAS-44406-6 32

Table 6: Mean percentage dry weight (dw), relative to total dry weight, of amino acids in seed from ‘Maverick’ and DAS-44406-6 33

Table 7: Mean weight (µg/g dry weight expressed as aglycon equivalents) of isoflavones in DAS- 44406-6 and ‘Maverick’ seed 34

Table 8: Mean levels of anti-nutrients in DAS-44406-6 and ‘Maverick’ seed. 34

Table 9: Mean values for mineral levels in seed from ‘Maverick’ and DAS-44406-6..... 35

Table 10: Mean weight (µg/g dry weight) of vitamins in seed from ‘Maverick’ and DAS-44406-6 36

Table 11: Summary of analyte means found in seed of DAS-44406-6 treatments that are significantly (adj. P<0.05) different from those found in seed of the control line 'Maverick' 37

LIST OF FIGURES

Figure 1: Vector map of plasmid PDAB8264 8

Figure 2: Representation of the genetic elements in the T-DNA insert of plasmid pDAB8264 8

Figure 3: Breeding strategy for plants containing event DAS-44406-6 12

Figure 4: General representation of the conversion of pyridyloxyacetate and phenoxyacetate herbicides to an inactive phenol in the presence of AAD-12. 18

LIST OF ABBREVIATIONS

AAD-12	Aryloxyalkanoate dioxygenase-12
ADF	acid detergent fibre
a.e.	Acid equivalent
a.i.	Active ingredient
AOAC	Association of Analytical Communities
bar	bialaphos resistance
BLAST	Basic Local Alignment Search Tool
bp	base pairs
2,4-D	2,4-dichlorophenoxyacetic acid
DCP	2,4-dichlorophenol
DNA	deoxyribonucleic acid
T-DNA	transferred DNA
dw	dry weight
ELISA	enzyme linked immunosorbent assay
EPSPS	5-enolpyruvylshikimate-3-phosphate-synthase
ESI-LC/MS	electrospray ionization liquid chromatography mass spectrometry
FAO	Food and Agriculture Organization of the United Nations
FARRP	Food Allergy Research and Resource Program
FASTA	Fast Alignment Search Tool - All
FSANZ	Food Standards Australia New Zealand
fw	fresh weight
GM	genetically modified
IgE	Immunoglobulin E
ILSI	International Life Sciences Institute
kDa	kilo Dalton
LC/MS	liquid chromatography mass spectrometry
LOQ	limit of quantitation
LSM	Least squares mean
MALDI-TOF	matrix-assisted laser desorption/ionization – time of flight
MAR	Matrix attachment region
Met	methionine
MS/MS	Tandem mass spectrometry
MW	Molecular weight
NDF	neutral detergent fibre
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
PEP	phosphoenolpyruvate
SAS	Statistical analysis Software
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
UTR	untranslated region

1. Introduction

A genetically modified (GM) soybean line DAS-44406-6, hereafter referred to as soybean 44406, has been developed that is tolerant to herbicides of the aryloxyalkanoate family including the phenoxy auxins such as 2,4-dichlorophenoxyacetic acid (2,4-D), and to the herbicides glufosinate ammonium and glyphosate.

Tolerance to 2,4-D is achieved through expression of the enzyme aryloxyalkanoatedioxygenase-12 (AAD-12) encoded by the *aad-12* gene derived from *Delftia acidovorans,* a gram-negative soil bacterium. The AAD-12 protein has previously been assessed in soybean by FSANZ (2011a).Tolerance to glufosinate ammonium is achieved through expression of the enzyme phosphinothricin acetyltransferase (PAT) encoded by the *pat* gene derived from another soil bacterium *Streptomyces viridochromogenes*. Tolerance to glyphosate is encoded through expression of a 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) encoded by the *2mepsps* gene from corn (*Zea mays*). Both the *pat* and *epsps* genes have been widely used for genetic modification of a number of crop species, including soybean.

It is anticipated that soybean 44406 will be grown in at least the United States of America (U.S.), Canada, Argentina and Brazil, subject to approval. There is currently no 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 44406. 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 over 35 countries worldwide (OECD, 2001) and has a long history of safe use for both humans and livestock. The major producers of soybeans, accounting for 90% of world production, are the U.S., Argentina, Brazil and China. Australia, while a net importer of soybean, grows crops in latitudes extending from the tropics (16^{o} S) to temperate regions (37^{o} 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 lines[1].

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

[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

- 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, 2001). Appropriate heat processing inactivates these compounds.

*S*oybean oil constitutes approximately 30% of global consumption of edible fats and oils (The American Soybean Association, 2011), 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. This substance is produced by bees during foraging and is taken back to the hive to be fed to larvae and young adult bees (Krell, 1996). It comprises pollen grains that are pelleted by the bee in the corbiculae ('pollen baskets') located on the posterior pair of legs. Beekeepers can collect the pellets by placing a screen at the entrance to a hive; as the bees pass through the screen, the pellets are dislodged and fall into a collection tray. The pellets are frozen or dried for storage and are then packaged for sale as bee pollen, which is generally consumed as the raw product without any further processing. It is highly unlikely that this product would be imported to Australia or New Zealand as domestic supply would satisfy market requirements.

2.2 Donor organisms

2.2.1 Delftia acidovorans

The *aad-12* gene was sourced from the bacterial species *Delftia acidovorans* strain MC1, a strain isolated from herbicide-contaminated building rubble (Müller *et al*., 1999). This bacterium is a member of the Pseudomonads, a ubiquitous group of environmental gram negative bacteria. It was originally classified in the genus P*seudomonas*, then renamed in the genus *Comamonas* (Tamaoka *et al*., 1987) and finally renamed again to *Delftia* (Wen *et al*., 1999).

Delftia spp. are aggressive colonisers of the rhizosphere of various crop plants and have a broad spectrum of antagonistic activity against plant pathogens (Han *et al*., 2005; see e.g. El-Banna, 2007). They have also been found to possess a variety of biodegradation mechanisms that could be exploited in the development of systems for the removal of chemicals that may be released into the environment (Patel *et al*., 1998; Müller *et al*., 1999; Urata *et al*., 2004). On rare occasions, *Delftia* spp. have been implicated in human infections (see e.g. Chun *et al*., 2009 and references therein).

D. acidovorans is one of several microorganisms that has been proposed as a bioconverter for use in the food industry to transform ferulic acid into vanillin and related flavour metabolites (Labuda *et al.*, 1992). However, commercial application has not been realised (see e.g. Yoon *et al*., 2005). The polyhydroxyalkanoates produced by *D. acidovorans* and other bacteria have

been proposed for use as biomaterial for use in tissue engineering and other medical applications (Sudesh, 2004).

2.2.2 Streptomyces viridochromogenes

The source of the *pat* gene is the bacterial species *Streptomyces viridochromogenes*, strain Tü494 (Wohlleben *et al*., 1988). The *Streptomycetae* 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. hygroscopicus* 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 Zea mays

Corn, *Zea mays*, is the source of the *epsps* gene that was modified to produce the *2mepsps* gene in soybean 44406 and is also the source of some of the regulatory gene elements. Corn is the world's third leading cereal crop, behind wheat and rice, and is grown in over 25 countries (OECD, 2002) and across a wide range of geographical conditions (OGTR, 2008). Also known as maize, corn has been grown in Mexico and Central America for some 8000 years and in Europe for 500 years and can thus be said to have a long history of safe use as a human food. The majority of corn that is grown however is destined for use as animal feed. In 2009, worldwide production of corn was over 818 million tonnes, with the United States and China being the major producers (~333 and 164 million tonnes, respectively) (FAOSTAT 2011).

The *epsps* gene was isolated from a cell suspension of 'Black Mexican' sweet corn. 'Black Mexican' is an heirloom cultivar of New England (USA) sweet corn originally introduced to the food supply in 1864 (http://www.southernexposure.com/productlist/prods/41103.html). Sweet corn is categorized as a vegetable and is mainly used for human consumption directly without processing.

2.2.4 Other organisms

Genetic elements from several other organisms have been used in the genetic modification of soybean 44406 (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 are not pathogenic in themselves and do not cause pathogenic symptoms in soybean 44406.

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:

Guttikonda, S.K. (2011). Cloning and characterization of the DNA sequence for the insert and its flanking border regions of DAS-44406-6 soybean. Study ID# 102117, Dow AgroSciences (unpublished).

Guttikonda, S. (2012). Bioinformatics evaluation of the putative reading frames across the whole T-DNA insert and junctions in DAS-44406-6 soybean for potential protein allergenicity and toxicity. Study ID# 120481, Dow AgroSciences (unpublished).

Han, L. Hoffman, T. (2011). Event sorting and selection process for the development of DAS-44406-6. Study ID# 110325, Dow AgroSciences (unpublished).

Mo, J. (2011). Molecular characterization of DAS-44406-6 soybean within a single segregating generation. Study ID# 102097, Dow AgroSciences (unpublished).

Poorbaugh, J. (2011). Molecular characterization of DAS-44406-6 soybean. Study ID# 101947, Dow AgroSciences (unpublished).

Zhuang, M.; Pareddy, D. (2011). Transformation information for plasmid pDAB8264. Study ID# 101880, Dow AgroSciences (unpublished).

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 pDAB8264 (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 pDAB8264. 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*, *aad-12* and *2mepsps* genes.

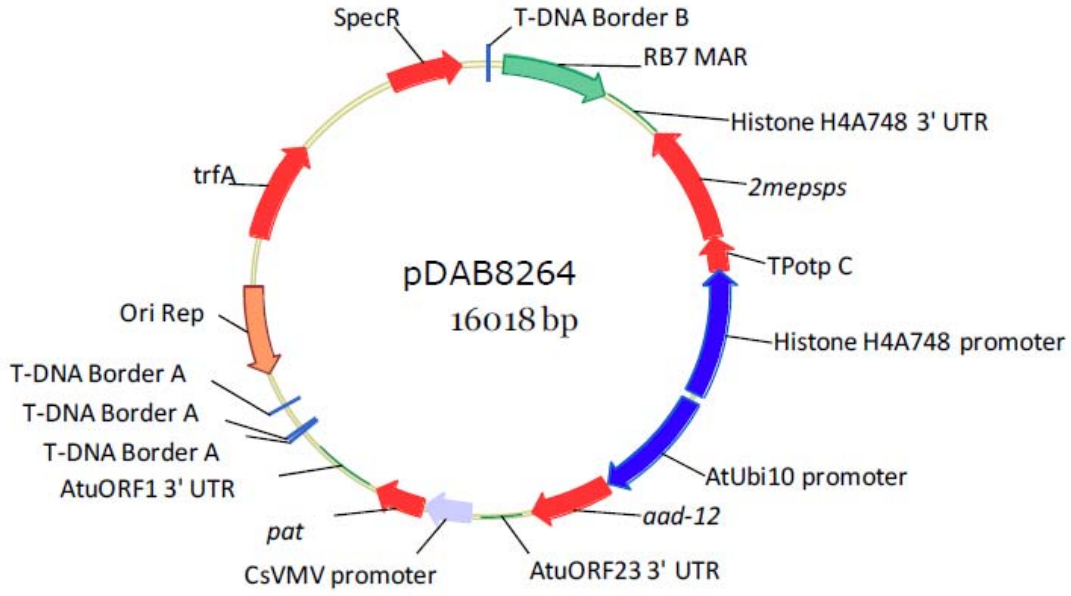


Figure 1: Vector map of plasmid PDAB8264

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.

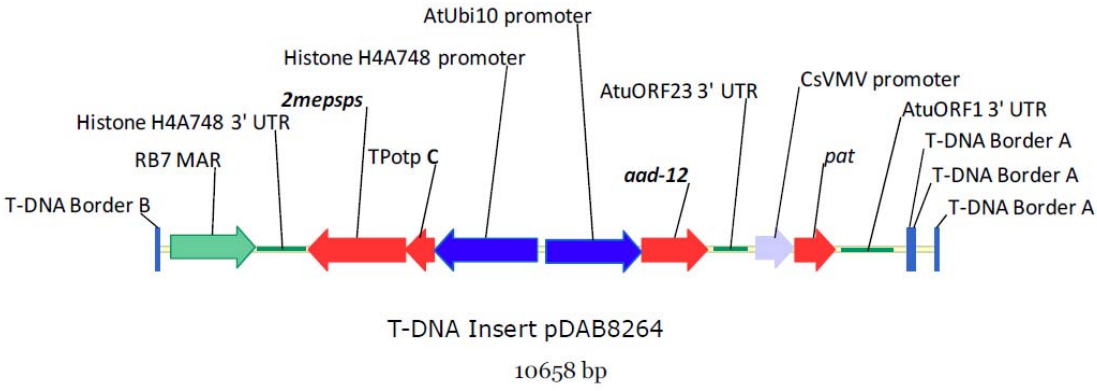


Figure 2: Representation of the genetic elements in the T-DNA insert of plasmid pDAB8264

Table 1: Description of the genetic elements contained in the T-DNA of pDAB8264

Genetic element	bp location on pDAB8264	Size (bp)	Source	Orient.	Description & Function	References
Border B	1 - 24	24	*Agrobacterium tumefaciens*		• Border repeat • Required for the transfer of the T-DNA into the plant cell	Zambryski (1988)
Intervening sequence	25 - 160	136	*Agrobacterium tumefaciens*		• Sequence from Ti plasmid pTi15955	Barker et al.(1983)
RB7-MAR	161 - 1326	1166	*Nicotiana tabacum*	Clockwise	• Matrix attachment region • Facilitates expression of the *aad-12* gene	Hall et al (1991)
Intervening sequence	1327 - 1365	39	Plasmid pENTR/D-TOPO		• The plasmid had been used as a cloning vector	Invitrogen Cat. # A10465
Histone H4A748	1366 - 2026	661	*Arabidopsis thaliana*		• Transcriptional terminator and polyadenylation site of histone H4A748	Chaboute et al (1987)
Intervening sequence	2027 - 2049	23			• Cloning sequence	
2mepsps	2050 - 3387	1338	*Zea mays*	Anti-clockwise	• Coding sequence of the *epsps* gene with two mutations providing glyphosate tolerance	Lebrun et al (1996); Lebrun et al (1997)
TPotp C	3388 - 3759	372	*Zea mays & Helianthus annuus*	Anti-clockwise	• Optimised chloroplast transit peptide derivative of the RuBisCo small unit genes from both species • Targets the 2mEPSPS protein to the plastids	Lebrun et al (1996)
Intervening sequence	3760 - 3763	4			• Cloning sequence	
Histone H4A748	3764 - 5193	1430	*Arabidopsis thaliana*	Anti-clockwise	• Promoter and 5' UTR • Drives constituitive expression of the *2mepsps* gene	Chaboute et al (1987)
Intervening sequence	5194 - 5285	92			• Cloning sequence	
AtUbi10	5286 - 6607	1322	*Arabidopsis thaliana*	Clockwise	• Polyubiquiton promoter, 5'UTR and intron • Drives constitutive expression of the *aad-12* gene	Norris et al (1993)
Intervening sequence	6608 - 6615	8			• Cloning sequence	
aad-12	6616 - 7497	882	*Delftia acidovorans*	Clockwise	• Coding sequence of the aryloxyalkanoate dioxygenase gene (adapted to plant codon usage)	Wright et al (2007); Wright et al (2010a)
Intervening sequence	7498 - 7599	102			• Cloning sequence	
AtuORF23	7600 - 8056	457	*Agrobacterium tumefaciens*		• Transcriptional terminator and polyadenylation site of ORF23 from Ti plasmid pTi15955	Barker et al.(1983)
Intervening sequence	8057 - 8170	114			• The plasmid had been used as a cloning vector	Invitrogen Cat. # A10465
CsVMV	8171 - 8687	517	Cassava vein mosaic virus	Clockwise	• Promoter and 5'UTR • Drives constitutive expression of the *pat* gene	Verdaguer et al (1996)
pat	8695 - 9246	552	*S.treptomyces viridochromogenes*	Clockwise	• Coding sequence of the phosphinothricin acetyl transferase gene (adapted to plant codon usage)	Wohlleben et al (1988)
Intervening sequence	9247 - 9348	102	Plasmid pCR2.1		• Cloning sequence	Invitrogen Cat. # K2050-01 cloning kit
ATuORF1	9349 - 10052	704	*Agrobacterium tumefaciens*		• Transcriptional terminator and polyadenylation site of ORF1 from Ti plasmid pTi15955	Barker et al.(1983)
Intervening sequence	10053 - 10280	228	*Agrobacterium tumefaciens*		• Sequence from Ti plasmid C58	Zambryski et al (1982); Wood et al (2001)

Genetic element	bp location on pDAB8264	Size (bp)	Source	Orient.	Description & Function	References
Border A	10281-10304	24	*Agrobacterium tumefaciens*		• Border repeat • Required for the transfer of the T-DNA into the plant cell	Zambryski (1988)
Intervening sequence	10305-10323	19	*Agrobacterium tumefaciens*		• Sequence from Ti plasmid C58	Zambryski et al (1982); Wood et al (2001)
Border A	10324-10347	24	*Agrobacterium tumefaciens*		• Border repeat • Required for the transfer of the T-DNA into the plant cell	Zambryski (1988)
Intervening sequence	10348-10634	287	*Agrobacterium tumefaciens*		• Sequence from Ti plasmid pTi15955	Barker et al.(1983)
Border A	10635-10658	24	*Agrobacterium tumefaciens*		• Border repeat • Required for the transfer of the T-DNA into the plant cell	Zambryski (1988)

3.2.1 *2mepsps expression cassette*

Homologues of the *epsps* gene are present in all plants, bacteria and fungi. The protein encoded by the gene is part of the shikamate pathway that is involved in aromatic amino acid synthesis.

The sequence of the *2mepsps* gene is derived from the wild type *epsps* gene from corn (*Zea mays*) with two single nucleotide mutations introduced by site directed mutagenesis. A methionine codon has been added to the N-terminal end of the 2mEPSPS protein sequence in order to restore the cleavage site of the optimized plastid transit peptide. The double mutant produces a 47.5 kDa protein with normal enzyme function and reduced affinity for the herbicide glyphosate.

The H4A748 promoter and terminator used to control expression of the *2mepsps* gene and are derived from the histone H4 gene of *Arabidopsis thaliana*. The use of the promoter directs high level constitutive expression, particularly in rapidly growing plant tissues.

TPotp C, encodes the optimized transit peptide derived from genes of corn and sunflower and targets the mature protein to the plastids where it is normally located in the cell.

3.2.2 *aad-12 expression cassette*

The *aad-12* gene of *D. acidovorans*, also referred to as *sdpA* (Schleinizt *et al*., 2004; Wright *et al.*, 2007; Wright *et al*., 2010b) has low homology (approximately 37% sequence identity) with the *tfdA* gene first isolated from *Ralstonia eutropha* (Streber *et al*., 1987) but found in phylogentically diverse bacteria (Baelum *et al*., 2008). The *tfdA* gene codes for an α-ketoglutarate-dependent dioxygenase which converts chlorinated phenoxyalkanoate herbicides such as 2,4-D into a harmless phenol and glyoxylate (refer to Section 4.2.2). Expression of the *aad-12* gene confers tolerance to both phenoxyalkanoate herbicides as well as to pyridyloxyacetic acids such as trichlopyr and fluroxypyr (Wright *et al.*, 2007). As well as occurring in *D. acidovorans*, *sdpa* genes have also been reported to occur in *Sphingomonas herbicidovorans* and *Rhodoferax* sp. but show considerable sequence diversity (Paulin *et al*., 2010).

The DNA sequence of the *aad-12* gene has been optimised for expression in plants and is approximately 80% identical to the DNA sequence of the native *aad-12* gene. The *aad-12* coding region in plasmid pDAB4468 is 882 bp in length and is driven by the constitutive polyubiquiton promoter from *Arabidopsis thaliana*. A matrix attachment region (MAR) from the root-specific *Rb7* gene (Hall, Jr. *et al*., 1991; Verma *et al*., 2005) of *Nicotiana tabacum* (tobacco) was included at the 5' end of cassette to potentially increase the consistency of

aad-12 expression (Abranches *et al*., 2005). When positioned on the flanking ends of gene cassettes, some MARs have been shown to increase expression of transgenes and to reduce the incidence of gene silencing. At the 3' untranslated region of the coding region is a transcript termination and polyadenylation region from *Agrobacterium tumefaciens.*

3.2.3 *pat* gene *expression cassette*

The *pat* gene from *Streptomyces viridochromogenes* and the *bar* gene from *S. hygroscopicus* confer tolerance to herbicides containing glufosinate ammonium (phosphinothricin). Both genes code for polypeptides of 183 amino acids and share 87% homology at the nucleotide sequence level (Wehrmann *et al*., 1996). Both genes have been widely used for genetic modification of food species.

The *pat* gene coding region (Wohlleben *et al*., 1988; Strauch *et al*., 1988) used in plasmid pDAB4468 is 552 bp in length and has been optimised for expression in plants. It is driven constitutively by a promoter region of the Cassava vein mosaic virus and terminated by a sequence of the 3'untranslated region of an open reading frame originating from plasmid pTi5955 of *Agrobacterium tumefaciens*.

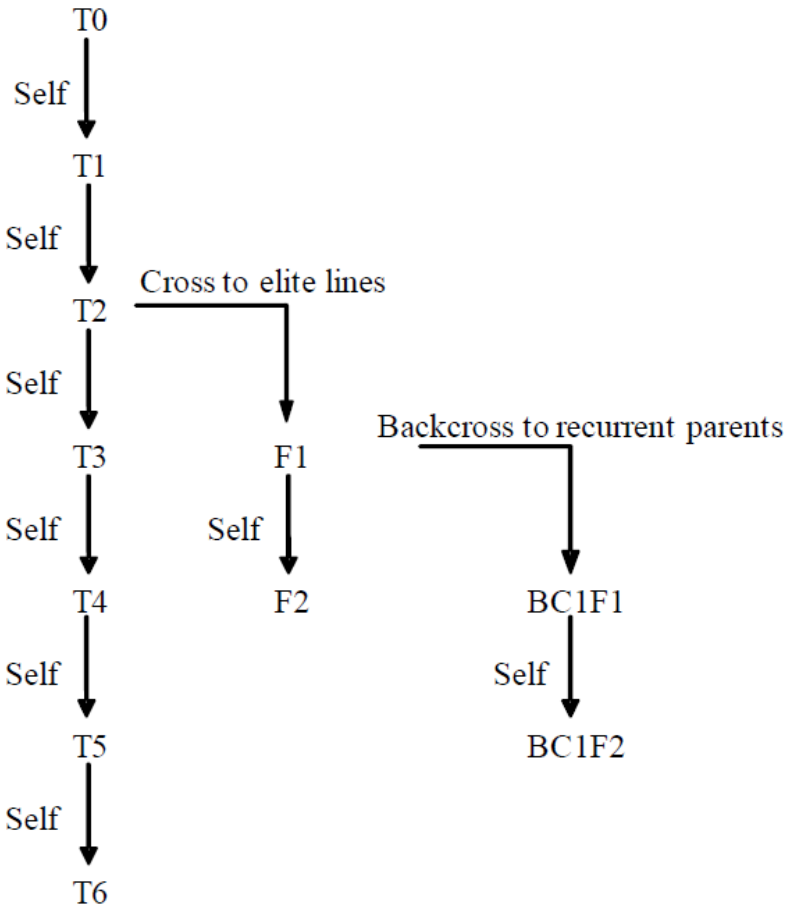
3.3 Breeding to obtain soybean line DAS-44406-6

A breeding programme was undertaken for the purposes of:

- obtaining generations suitable for analysing the molecular and genetic characteristics of soybean 44406
- ensuring that the DAS-44406-6 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 T_0 plants (see Section 3.1) a series of self-fertilisation and seed bulking crosses proceeded up to generation T_6. At the T_2 generation, plants were crossed with a number of elite lines to produce an F1 generation which was either self-fertilised to produce an F2 generation, or backcrossed to the appropriate parental elite cultivar. Figure 3 also indicates the generations that were used in the various studies characterising soybean 44406-6.



Analysis	DAS-444Ø6-6 Soybean Generations Used	Control
Molecular Analysis	T2, T3, T4, T6, F2	Maverick
Segregation Analysis	F2, BC1F2	Maverick
Protein Characterization	T4	Maverick
Protein Expression	T4	Maverick
Composition	T4	Maverick

Figure 3: Breeding strategy for plants containing event DAS-44406-6

3.4 Characterisation of the genes in the plant

A range of analyses was undertaken to characterise the genetic modification in soybean line 44406. These included: determination of insert copy number and integrity; and DNA sequence and ORF analysis of inserted DNA as well as flanking and junction regions.

3.4.1 Transgene copy number, insertion integrity and plasmid backbone analysis

Total genomic DNA from greenhouse-grown, leaf tissue of individual soybean 44406 seedlings (3 plants/generation) from each of five generations (T_2, T_3, T_4, T_6 and F_2) and a negative control (non-GM cultivar 'Maverick') was used for Southern blot analyses. A positive control (DNA from 'Maverick' spiked with T-DNA from plasmid pDAB8264) was also included in the Southern blot analyses. Lateral Flow Strip testing was done of the soybean 44406 seedlings to confirm the presence of the PAT protein. The DNA from soybean 44406 seedlings, the negative control and the positive control was digested with one, or a combination, of restriction enzymes. The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with 16 different digoxigenin (DIG)-labelled probes that represented various sections of the T-DNA and vector backbone.

The Southern blot analyses indicated there is a single insert in event DAS-44406 and that the arrangement of the T-DNA genetic material is the same as that in the pDAB8264 plasmid (refer to Figure 2). No vector backbone sequences are present in soybean 44406.

Additional to these data, the hybridization pattern was the same across the five generations, thus indicating there had been no rearrangements to the inserted DNA and therefore, that the insert was stably inherited across the generations.

3.4.2 Insert characterisation

Genomic DNA was obtained from leaf tissue of T_6 generation soybean 44406 plants and a negative control (cultivar 'Maverick'). These samples were used to characterise the DNA sequence in the transgene insertion and its flanking border regions. The sequence of the estimated parental locus in cultivar 'Williams 82' was used to design primers for cloning the parental locus in 'Maverick'. 'Williams 82' is the elite US cultivar that was chosen for sequencing of the whole soybean genome (Schmutz *et al*., 2010).

Standard polymerase chain reaction (PCR) was used to clone the insert and border sequences of event DAS-44406-6 and then obtain the sequences by primer walking. In total, eight fragments were cloned and sequenced. DNA sequencing analysis was done using commercially available software (Sequencher®).

The total cloned sequences of DAS-44406-6 comprised 13,659 bp, consisting of 1,495 bp of 5' flanking sequences, 1885 bp of 3' flanking sequences and 10,280 bp of the T-DNA insert from plasmid pDAB8264. The sequence analysis revealed a 3bp insertion at the 5' integration junction and also confirmed that all inserted sequences originate from the T-DNA of the transforming pDAB8264 plasmid.

Following the sequence analysis, a bioinformatic search was undertaken using the BLASTn algorithm to compare the 5' and 3' border sequences with sequences in the soybean genome from 'Williams 82' (for a general discussion of this type of analysis refer to Section 4.5.2). This indicated a 99% identity with a segment of Chromosome 6 of 'Williams 82' and therefore confirmed that the flanking sequences in event DAS-44406-6 are of *Glycine max* origin. Sequence alignment of the parental locus in 'Maverick' with the full sequence of the DAS-44406-6 insert and flanking regions, indicated that there has been a deletion of a 4,383 bp fragment from the parental locus during the T-DNA integration in DAS-44406-6. Within this deletion, an open reading frame (ORF) of 777 bp was identified; no analysis was done to ascertain whether this ORF was associated with a promoter and/or terminator. A BLASTp search of the ORF against sequences in GenBank (see Section 4.5.2) did not identify any similarity to known proteins i.e. it does not appear that there has been disruption to any known endogenous genes as a result of the deletion.

3.4.3 Novel open reading frame analysis

Two separate *in silico* analyses of the whole insert and the flanking regions (5' -1494 bp and 3' - 1885 bp) were done to determine whether any novel ORFs had been created. Each analysis comprised a search of six-frame translations between stop codons regardless of the presence of a start codon or the number of amino acid residues coded by the nucleotide sequence.

A total of 651 ORFs within the T-DNA insert were identified. Twelve ORFs were identified in the flanking regions (see Table 2). Discussion of the analysis of the novel ORFs is given in Section 4.1.

Table 2: Location and characterisation of novel ORFs in the flanking regions

Location of flanking region	Nucleotide location of reading frame	Number of amino acids coded	Deduced amino acid sequence
5' border	1483 - 1554	24	GHHGGPNSLKLESSQLRSTGQIRS
	1307 -1507	67	RVGPDIVACYWGFLSVACVLHYCMGLAHPTIQCIF MCDNVMGFYCSCCFLFRNLHVNGKVIMEVRIV_
	1476 - 1502	9	TVRSSWRSE
	11943 - 11773	25	FQTIRTSMMTLPFTCKFLKRKQQEQ
	1553 - 491	21	ERIWPVDLNCELSNFKLFGPP
	1534 - 1469	22	TSIASFLISNYSDLHDDLTVYM
3' border	11737 - 11790	18	IYPAPASQQLDLQRTNVV
	11726 - 12019	98	FTIEYILPQPANSSIYRERMSCDMWNKATTTYMNL TIESGSPSCDVIHGMDMVADRKRKKKCMYMCENE SFFYPNNKKKLIIYPKNYLHDRYVHFFP_
	11682 - 11825	48	KRPQCVIKLSKRQFDLQLNISCPSQPTARFTENECRV ICGTRQRQQHT_
	11943 - 11773	57	KKLSFSHIYMHFFFLFRSATMSMPWITSQLGDPDSIV RFMYVVVVALFHISHDIRSL_
	11831 - 11730	34	DSCMLLSLPCSTYHTTFVLCKSSCWLAGAGYIQL
	11965 - 11720	82	FFFIIWIKKTLIFTHIHAFLFSFSIGHHVHAMDYITTR RPRLYCEIHVCCCRCLVPHITRHSFSVNRAVGWLGQ DIFNCKSN_

3.5 Stability of the genetic changes

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 F_2 generation of soybean 44406 generated by crossing T_2 plants with an elite non-GM line taken from the Applicant's soybean germplasm development programme. The F_1 plants were self-pollinated to obtain the F_2 generation (refer to Figure 3).

Leaves of 119 F_2 soybean 44406 plants were analysed by Lateral Flow Strip testing for expression of the PAT protein. A Chi squared (X^2) analysis of the results was done for the

3:1 segregation ratio of PAT positive versus negative plants. A total of 96 plants were positive for PAT while 23 were negative. The X^2 value of 2.042 (P>0.05) indicated that the segregation ratio was consistent with the Mendelian inheritance pattern of a single dominant trait.

All 119 plants were also tested by event-specific PCR for the presence/absence of the DAS-44406-6 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.

The genetic stability of event DAS-44406-6 in the soybean genome was established by the experimental work described in Section 3.4.1 in which the hybridisation pattern of the event was shown to be identical across the T_2, T_3,T_4, T_6 and F_2 generations.

3.6 Antibiotic resistance marker genes

No antibiotic marker genes are present in soybean 44406. Plasmid backbone analysis shows that no plasmid backbone has been integrated into the soybean genome during transformation, i.e. the *Spec* gene, which was used as a bacterial selectable marker gene, is not present in soybean 44406.

3.7 Conclusion

Comprehensive molecular analyses of soybean line DAS-44406-6 indicate that a single copy of T-DNA containing three expression cassettes for the genes *2m epsps, aad-12* and *pat* has been inserted at a single locus in Chromosome 6 of the soybean genome. No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred to the plant. As a result of the integration of the three expression cassettes, there is a 3 bp insertion at the 5' junction region and a deletion of 4,385 bp of host DNA. 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 consider 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 open reading frames during the introduction of the T-DNA of plasmid pDAB8264 (see Section 3.4.3)

- those that were expected to be produced as a result of the expression of the introduced genes. Soybean 44406 expresses three novel proteins, 2m EPSPS, AAD-12, and PAT.

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

Study submitted:

Guttikonda, S. (2012). Bioinformatics evaluation of the putative reading frames across the whole T-DNA insert and junctions in DAS-44406-6 soybean for potential protein allergenicity and toxicity. Study ID# 120481, Dow AgroSciences (unpublished).

Twelve novel ORFs were identified in the flanking regions and 651 in the T-DNA insert itself, in event DAS-44406-6 (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 celiac-induction sequences, residing in the FARRP (Food Allergy Research and Resource Program) dataset (Version 12) within AllergenOnline (University of Nebraska; http:www.allergenonline.org/).

Of the 12 ORFs identified in the flanking regions (see Table 2), only six were longer than 29 amino acids (the number required to satisfy >35% identity over at least 80 amino acids). These six underwent a similarity search using the FASTA algorithm and the BLOSUM50 scoring matrix (for general information of this type of analysis see Section 4.6.2). No similarities with known allergens were found. All of the ORFs were screened for any matches of eight contiguous amino acids to known allergens. No matches of eight or more contiguous amino acids were found in any of the sequences.

Similarly, for the T-DNA, no matches of eight or more contiguous amino acids were found for any of the sequences. A total of 441 of the ORFs were less than 29 amino acids; of the 210 remaining ORFs searched using the FASTA algorithm, no similarities with known allergens were found.

4.1.2 Toxicity assessment

The sequences corresponding to the 12 identified ORFs in the flanking regions were compared with protein sequences present in a number of large public reference databases including Uniprot_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). No significant similarities of the 12 ORFs to any sequences (including those of known toxins) in the databases were found.

The BLASTP search of the T-DNA insert sequences returned 10 ORFs that showed alignments with an *E*-value <1.0 (for explanation see Section 4.5.2). As expected, three of the alignments were with 2m EPSPS, AAD-12 and PAT. None of the other significant alignments were related to any known protein toxins.

4.1.3 Conclusion

It is concluded that, in the unlikely event that any of the identified 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 2m EPSPS, AAD-12 and PAT proteins

4.2.1 2m EPSPS protein

Glyphosate acts as a herbicide by inhibiting the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). This endogenous enzyme is involved in the shikimate pathway for aromatic amino acid biosynthesis which occurs exclusively in plants and microorganisms, including fungi. Inhibition of the wild type EPSPS enzyme by glyphosate leads to deficiencies in aromatic amino acids in plant cells and eventually to the death of the whole plant. The shikimate biochemical pathway is not present in animals. For this reason, enzymes of the shikimate pathway have been considered as potential targets for essentially non-toxic herbicides (such as glyphosate) and antimicrobial compounds.

Naturally occurring EPSPS proteins are widespread in nature and have been extensively studied over a period of more than thirty years. The modified 2mEPSPS protein present in soybean 44406 differs from the wild type maize enzyme by two amino acid substitutions – threonine replaced by isoleucine at position 102, and proline replaced by serine at position 106 (Lebrun *et al.*, 1997). These two amino acid changes result in a protein with greater than 99.5% identity to the native maize EPSPS protein. However the modification confers a decreased binding affinity for glyphosate thus allowing the protein to maintain an adequate level of enzymatic activity in the presence of the herbicide. Plants expressing the modified maize enzyme therefore are able to continue to function in the presence of the herbicide.

The commercial soybean cultivar 'Maverick', used as the parent for the genetic modification described in this application, contains a wild type EPSPS protein.

4.2.2 AAD-12 protein

The native AAD-12 (GenBank Accession AAP88277) is an α-ketoglutarate dependent dioxygenase that catalyses the breakdown of pyridyloxyacetate auxins and achiral[2] phenoxy auxins to an intermediate that, itself, is then spontaneously broken down to a herbicidally inactive phenol and glyoxylate (refer to Figure 4). In the case of soybean 44406, the active herbicide that would be applied is 2,4-D and the inactive phenol produced would be 2,4-dichlorophenol (DCP). Such side-chain degradation of 2,4-D has been observed in many conventional plants (IPCS, 1984) albeit to a limited degree not necessarily associated with tolerance to 2,4-D.

Native AAD-12 is also known as (S)-phenoxypropionate/α-ketoglutarate-dioxygenase (SdpA) and is one of two enantiospecific[3] enzymes that occurs in *Delftia acidovorans*. While RdpA is highly specific to the R enantiomer of 2-phenoxypropionates and shows weak activity towards phenoxyacetates, SdpA is enantioselective to the S enantiomers but can also convert certain phenoxyacetates such as 2,4-D and 4-chloro-2-methylphenoxyacetate

---

[2] a term used to describe a molecule which, in a given configuration, is superimposible on its mirror image

[3] Chiral versions of otherwise identical compounds are termed enantiomers. The common way of denoting the enantiomers is with an (*S*) (left handed) or (*R*) (right handed). Enantiospecificity describes the ability of an enzyme to distinguish between the enantiomers.

(MCPA) (Paulin *et al*., 2010). The term AAD-12 to describe SdpA was first used by Wright *et al.* (2007).

AAD-12 is closely related to AAD-1, a protein conferring tolerance to 2,4-D and quizalofop-P-ethyl, that was considered in Application A1042 (FSANZ, 2010). AAD-12 has significantly greater *in vitro* activity on 2,4-D than AAD-1 (Wright *et al*., 2010b).

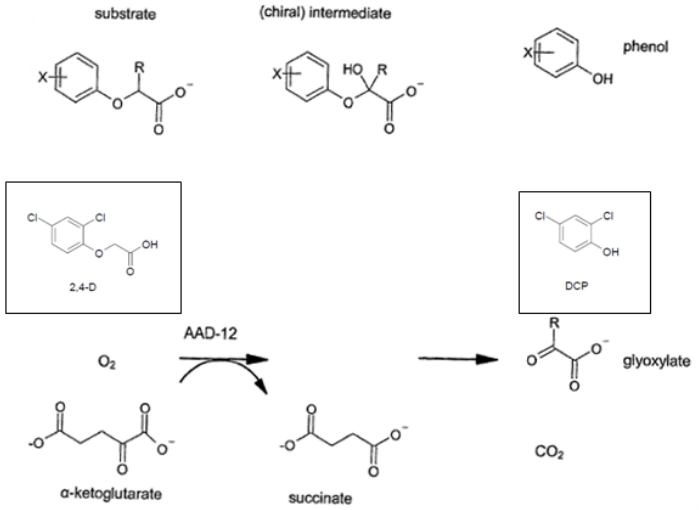


Figure 4: General representation of the conversion of pyridyloxyacetate and phenoxyacetate herbicides to an inactive phenol in the presence of AAD-12 (Diagram modified from Wright et al. (2007)). The structures of 2,4-D and its inactive phenol, DCP, are given in the rectangles.

4.2.3 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. hygroscopicus* 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 (L-PPT) (see Thompson *et al*., 1987) more recently known also as glufosinate ammonium. Free L-PPT 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 L-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 L-PPT to N-acetyl-L-PPT, a herbicidally-inactive compound. The kinetics and substrate specificity of the PAT enzyme are well characterised; it has a high specificity for L-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 L-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^{o}C, with maximum activity occurring between 40 and 45°C. Complete thermoinactivation occurs after 10 minutes at 60^{o}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

Studies submitted:

Maldonado, P.M. (2011a). Field expression of a transformed soybean cultivar containing aryloxyalkanoate dioxygenase (AAD-12), double mutant maize EPSPS gene (2mEPSPS), and phosphinothricin acetyltransferase (PAT) - event DAS-44406-6. Study ID# 101104.02, Dow AgroSciences (unpublished).

Maldonado, P.M. (2011b). Method validation for the determination of 5-enolpyruvylshikimate-3-phosphate synthase (2mEPSPS) protein in soybean tissues by enzyme-linked immunosorbent assay (ELISA). Study ID# 101768, Dow AgroSciences (unpublished).

Smith-Drake, J.S.; Sosa, M.J.; Shan, G. (2009a). Method validation for the determination of aryloxyalkanoate dioxygenase-12 (AAD-12) in soybean tissues using and enzyme-linked immunosorbent assay (ELISA). Study ID# 081008, Dow AgroSciences (unpublished)

Smith-Drake, J.S.; Sosa, M.J.; Shan, G. (2009b). Method validation for the determination of phosphinothricin acetyltransferase (PAT)) in soybean tissues using an enzyme-linked immunosorbent assay (ELISA). Study ID# 081022, Dow AgroSciences (unpublished)

The 2m EPSPS, AAD-12 and PAT proteins are expected to be expressed in all plant tissues since the *2m epsps, aad-12* and *pat* genes are driven by constitutive promoters (refer to Table 1). Ten sites in the U.S[4], representing regions of diverse agronomic practices and environmental conditions for soybean, were planted with soybean 44406 (generation T_4) and 'Maverick'. Five herbicide spraying treatments were applied to soybean 44406 namely, unsprayed, sprayed with 2,4-D (1120 g a.e./ha), sprayed with glufosinate (454 g a.i./ha), sprayed with glyphosate (1260 g a.e./ha) and sprayed with 2,4-D + glufosinate + glyphosate (1120 a.e/ha + 374 g a.e./ha + 1260 g a.e./ha). Samples for analysis of expression of 2m EPSPS, AAD-12 and PAT were taken from a number of plant parts at specific growth stages (refer to Table 2).

The 2m EPSPS, AAD-12 and PAT protein levels were determined by enzyme linked immunosorbent assay (ELISA) using commercial ELISA kits specific for each protein. A commercially available software programme was used to calculate the concentrations of immunoreactive 2m EPSPS, AAD-12 and PAT proteins from optical density values.

No 2m EPSPS, AAD-12 or PAT proteins were detected in samples taken from 'Maverick' plants. For soybean 44406 plants, 2m EPSPS, AAD-12 and PAT proteins were detected in all plant parts (Table 3).

2m EPSPS was lowest in the seed (approximately 22 µg/g dry weight) and highest in V10 – 12 leaves (approximately 2323 µg/g dry weight). AAD-12 was lowest in the roots (approximately 25 µg/g dry weight) and highest in V10 – 12 leaves (approximately 116 µg/g dry weight). PAT protein concentrations were much lower than those for AAD-12 but similarly, the V10 – 12 leaves contained the highest levels (approximately 10 µg/g dry weight) and the roots contained the lowest levels (approximately 2 µg/g dry weight).

[4] Iowa (2 sites), Illinois (2 sites), Indiana, Michigan, Missouri and Nebraska (2 sites)

Table 3: Average concentration of 2m EPSPS, AAD-12 and PAT proteins in various plant parts from soybean 44406

Sample source	Growth Stage	Treatment	Average (rounded) protein content in µg/g dry weight ±SD		
			2m EPSPS	AAD-12	PAT
Leaf (8 trifoliate leaves)	V5	unsprayed	2368±973	112±34	8±4
		+ 2,4-D	2261±1009	111±27	9±3
		+ glufosinate	2062±962	107±29	8±4
		+ glyphosate	1846±975	101±29	8±3
		3 herbicides	2100±784	103±34	8±3
Leaf (8 trifoliate leaves)	V10-12	unsprayed	2583±825	118±36	10±2
		+ 2,4-D	2203±584	121±36	9±3
		+ glufosinate	2188±543	109±25	10±2
		+ glyphosate	2512±1259	114±27	9±3
		3 herbicides	2131±726	119±46	10±3
Root (3 plants)	R3	unsprayed	89±32	23±10	1±0.6
		+ 2,4-D	93±20	24±10	1±0.6
		+ glufosinate	103±47	24±11	1±0.7
		+ glyphosate	112±30	29±7	1±0.4
		3 herbicides	104±43	27±9	1±0.6
Forage (3 plants)	R3	unsprayed	357±146	73±20	6±1
		+ 2,4-D	330±109	72±22	5±1
		+ glufosinate	321±74	73±20	6±1
		+ glyphosate	400±140	76±19	6±1
		3 herbicides	367±125	70±21	6±1
Seed (500 g)	R8 - maturity	unsprayed	21±6	27±9	2±0.4
		+ 2,4-D	22±6	27±10	2±0.3
		+ glufosinate	22±7	27±10	2±0.4
		+ glyphosate	22±6	25±6	2±0.3
		3 herbicides	21±6	25±6	2±0.3

For information on soybean growth stages see e.g. NDSU (2004).

4.4 Protein characterisation studies

Studies submitted:

Embrey, S.K.; Schafer, B.W. (2009). Certificate of analysis of the test/reference/control substance: phosphinothricin acetyltransferase (PAT – TSN031116-0001). Study ID# BIOT09-203839, Dow AgroSciences (unpublished).

Embrey, S.K. (2011). Certificate of analysis of the test/reference/control substance: aryloxyalkanoate dioxygenase (AAD-12) – TSN030732. Study ID# BIOT10-227507, Dow AgroSciences (unpublished).

Karnoup, A.; Kuppannan, K. (2008). Characterization of AAD-12: Batch TSN030732-002. Study ID# ML-AL MD-2008-003833. The Dow Chemical Company (ubpublished).

Karnoup, A.; Kuppannan, K. (2010). Analytical characterization of 2m-EPSPS (5-enolpyruvylshikimate-3-phosphate synthase) containing TIPS mutation. The Dow Chemical Company (unpublished).

Lin, G.; Allen, J.; Chew, L.; Shields, J.; Chiu, Y.; Greenwalt, S.; Xu, X.; Walsh, T. (2006). Production, purification, and characterization of recombinant AAD-12 expressed in *Pseudomonas fluorescens* for submission on supporting regulatory toxicology and eco-toxicology study. Study ID# DERBI 259733, Dow AgroSciences (unpublished).

Lin, G.; Shan, G.; Xu, X.;Frey, M (2011). Production and characterization of 2mEPSPS (DMMG) protein for supporting regulatory toxicology study. Study ID# DAI 1006, Dow AgroSciences (unpublished).

Schafer, B.W. (2010). Certificate of analysis of the test/reference/control substance 2mEPSPS protein (TSN033171-0001). Study ID# BIOT10-255698, Dow AgroSciences (unpublished).

Schafer, B.W.; Juba, A.A. (2011). Characterization of the phosphinothricin acetyltransferase (PAT) protein derived from transgenic soybean event DAS-44406-6. Study ID# 102098, Dow AgroSciences (unpublished).

Schafer, B.W.; Juba, A.N.; Harpham, N.J.; Mayes, M.R. (2011). Characterization of the aryloxyalkanoate dioxygenase-12 (AAD-12) and double mutant 5-enolpyruvylshikimate-3-phosphate synthase (2mEPSPS) proteins derived from transgenic soybean event DAS-44406-6. Study ID# 101707, Dow AgroSciences (unpublished).

Site directed mutagenesis of the wild type *epsps* gene from maize produced the double mutant enzyme 2m EPSPS which carries two amino acid changes. When fused to a chimaeric optimised chloroplast transit peptide, the 2m EPSPS enzyme is reported to generate optimal glyphosate tolerance in crops (Lebrun *et al.,* 1997). The N-terminal methionine (Met) is cleaved with the transit peptide. The mutations, known collectively as the TIPS mutation (Funke *et al*., 2009), are at positions 102 (Thr to Ile) and 106 (Pro to Ser) of the mature protein which comprises 444 amino acids (excluding the terminal Met) and has a calculated molecular weight of approximately 47 kDa.

The AAD-12 protein produced by soybean 44406 has an amino acid sequence that is 99% homologous with the native AAD-12, differing only in that an alanine has been added at position 2. The Applicant claims that this addition serves the dual purpose of facilitating cloning operations and optimising translation initiation. The AAD-12 protein comprises 293 amino acids and has an approximate molecular weight of 32 kDa.

The PAT protein produced by soybean 44406 is identical to the native protein (Uniprot Accession No. Q57146). It comprises 183 amino acids and has an approximate molecular weight of 21 kDa.

None of the proteins are produced in sufficient quantity in soybean 44406 to isolate enough for the toxicological and biochemical studies required for a safety assessment. A standard procedure to overcome this type of 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 for the toxicological and biochemical studies. The AAD-12 and 2m EPSPS proteins were therefore expressed in recombinant *Pseudomonas fluorescens* while the PAT protein was expressed in recombinant *Escherichia coli.* Characterisation tests were done to confirm the identity and equivalence of these bacterially-produced proteins to those produced in soybean 44406. For the PAT protein, only the identity of the protein as produced in soybean 44406 and compared with a bacterially-produced protein was considered, since other characteristics of this protein have already been extensively studied (see e.g. Thompson *et al*., 1987; Wehrmann *et al*., 1996; Hérouet *et al*., 2005)

4.4.1 Comparison of plant and microbial proteins

Testing of the equivalence of the microbial- and plant-derived 2m EPSPS, AAD-12 and PAT was done as follows:

- For all three proteins, SDS-PAGE and Western blotting was done. For the PAT analysis on SDS-PAGE/Western blotting, a total plant protein extract was compared with a purified PAT microbial extract while for the AAD-12 and 2m EPSPS samples, purified protein from both sources was used. Proteins on SDS were detected by Coomassie staining; immunoreactivity was detected on Western blots using
  - for 2m EPSPS, an anti-2m EPSPS monoclonal mouse primary antibody with a goat anti-mouse horseradish peroxidise-linked secondary antibody.
  - for AAD-12, an anti-AAD-12 polyclonal rabbit primary antibody with a goat anti-rabbit horseradish peroxidise-linked secondary antibody
  - for PAT, an anti-PAT polyclonal rabbit primary antibody with a goat anti-rabbit horseradish peroxidise-linked secondary antibody AND

a monoclonal mouse primary antibody with a goat anti-mouse horseradish peroxidise-linked secondary antibody.

- For the PAT protein, a commercially available ELISA plate assay (EnviroLogix) was also used for immunochemical detection.
- For the AAD-12 and 2m EPSPS proteins, peptide mass fingerprinting and sequence analysis was done using matrix-assisted laser desorption/ionization – time of flight (MALDI-TOF) and electrospray ionization LC/MS (ESI-LC/MS) mass spectrometry techniques in order to obtain information on the amino acid sequence of the proteins. N- and C-terminal peptides were then further analysed by tandem mass spectrometry to confirm their amino acid sequences.
- For AAD-12 and 2m EPSPS, a glycosylation analysis was undertaken. N-glycosylated proteins are glycosylated on an asparagine residue and commonly contain an asparagine-X-serine/threonine sequence (N-X-S/T), where X is 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). The occurrence of these motifs does not, however, indicate that the protein will necessarily be glycosylated and since *P. fluorescens*, like most prokaryotes, lacks the capacity for protein glycosylation (Wacker *et al*., 2002; Abu-Qarn *et al*., 2008) it would be unusual to find that N-glycosylation had occurred in the bacterially-derived proteins. Detection of post-translational glycosylation in 2m EPSPS and AAD-12 was done using a GelCode Glycoprotein Staining Kit (Thermo-Pierce).
- For the 2m EPSPS proteins the enzyme activity was determined by a spectrophotometric method using malachite green (Lanzetta *et al*., 1979). Studies supporting the biological equivalence of AAD-12 from microbial and plant sources were considered by FSANZ in Application A1046 (FSANZ, 2011a).

Molecular weight and immunoreactivity of AAD-12, 2m EPSPS and PAT

Material for SDS-PAGE and Western blotting analysis of the plant-produced proteins was taken from tissue of the T_4 generation of soybean 44406. The presence of the proteins was confirmed by either lateral flow strip assay (AAD-12 and 2m EPSPS) or ELISA assay (PAT).

The SDS-PAGE analyses of the AAD-12 and 2m EPSPS purified proteins indicated that each protein from both sources produced a single major band that ran at the expected approximate molecular weight calculated from the known amino acid sequence of each protein i.e. the estimated MW of AAD-12 was approximately 32 kDa and that of 2m EPSPS was 47 kDa.

Western blots using both polyclonal and monoclonal antibodies to PAT showed an immunoreactive band at approximately 20 kDa in both crude extract of soybean 44406 and bacterially-derived PAT which compared with the calculated molecular weight of 21 kDa. Western blots using antibodies to AAD-12 and 2m EPSPS showed an immunoreactive band at the expected approximate molecular weight for each protein derived from both sources.

MALDI-TOF and ESI-LC mass spectrometry of AAD-12 and 2mE EPSPS

Purified AAD-12 and 2m EPSPS proteins from both sources were run on SDS-PAGE and the corresponding protein bands were excised and subjected to in-gel digestion by trypsin and chymotrypsin followed by MALDI -TOF and ESI-LC mass spectrometry. The masses of the detected peptides were compared to those deduced from potential trypsin or chymotrypsin cleavage sites.

For AAD-12, MALDI-TOF sequence coverage was approximately 84% of the theoretical peptide sequences and included both the N- and C-termini. The analysis confirmed that the sequence of the soybean 44406-derived protein matched that of the microbially-derived protein. ESI-LC/MS analysis showed that, as for the microbially-derived protein, the N-terminal Met of the soybean 44406-derived AAD-12 protein had been removed and the alanine at position 2 of the soybean 44406-derived protein had been acetylated. Cleavage of the N-terminal Met and amino-terminal acetylation is common in eukaryotes (Polevoda and Sherman, 2002). The C-terminal sequences of the soybean 44406- and microbially-derived AAD-12 were as expected and were indistinguishable.

For 2m EPSPS, MALDI-TOF sequence coverage was approximately 86% of the theoretical sequences and included both N- and C- termini. The analysis confirmed the equivalence of the protein from both sources. For both the N- and C-termini, the sequence of the protein from both sources was indistinguishable and was identical to the expected.

Glycosylation of AAD-12 and 2m EPSPS

Glycosylation analysis of the purified AAD-12 and 2m EPSPS proteins from plant and microbial sources derived from soybean 44406 did not detect any covalently-linked carbohydrates. Further confirmation of the non-glycosylated status of both proteins from both sources can be deduced from the molecular weight analyses in which the proteins ran at the expected molecular weight in the SDS-PAGE/western blot analyses; glycosylation would add to the apparent overall protein molecular weight.

Enzyme activity of 2m EPSPS

Both the bacterially-derived and 44406-derived 2m EPSPS were shown to have the expected catalytic activity in the presence of PEP substrate.

4.4.2 Conclusion

A range of characterisation methods submitted by the Applicant and in previous studies confirmed the identity and equivalence of the 2m EPSPS, AAD-12 and PAT proteins produced in both a bacterial expression system and in soybean 44406. Based on weight-of-evidence, it is concluded that microbially-derived 2m EPSPS, AAD-12 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 protein will behave like any other dietary protein. 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

The PAT protein will not be addressed further in this assessment as it has been sufficiently characterised from a food safety perspective (see e.g. Delaney *et al*., 2008). A bioinformatics analysis was submitted comparing the PAT protein sequence with sequences in an updated (February 2011) protein database but none of the significant sequence alignments were associated with any known protein toxins. The focus is therefore on whether 2m EPSPS and AAD-12 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 that may confer resistance to heat or processing and/or digestion.

Studies submitted:

Guttikonda, S. (2011). Similarity assessment of 2m EPSPS protein to known toxins by bioinformatics analysis (updated March, 2011). Study ID# 110329, Dow AgroSciences (unpublished).
Song, P. (2011). Sequence similarity assessment of AAD-12 protein to known toxins by bioinformatics analysis (Updated February, 2011). Study ID# 110327, Dow AgroSciences (unpublished).
Embrey, S.K. (2011). Heat lability of double mutant 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) protein. Study ID# 110461, Dow AgroSciences (unpublished).
Schafer, B.W. (2008). Effect of heat treatment on a recombinant aryloxyalkanoate dioxygenase-12 (AAD-12). Study ID# 080140, Dow AgroSciences (unpublished).

Song, P. (2011). Sequence similarity assessment of PAT protein to known toxins by bioinformatics analysis (updated February 2011). Study ID# 110331, Dow AgroSciences (unpublished).
Jeong, Y.C.; Golden, R.M. (2011). 2mEPSPS: Acute oral toxicity study in Crl:CD1 (ICR) mice. Study ID# 101168, Dow AgroSciences (unpublished).
Wiescinski, C.M.; Golden, R.M. (2008). AAD-12: Acute oral toxicity study in Crl:CD1 (ICR) mice. Study ID# 081037, Dow AgroSciences (unpublished).

4.5.1 History of human consumption

As outlined in Section 4.2.1, homologues of the EPSPS protein are found in plants and a range of other organisms and would therefore be routinely consumed by humans as a normal part of the diet. In addition, the *epsps* gene, most usually sourced from *Agrobacterium tumefaciens*, has been widely used in the genetic modification of commercialised food crops over the last 10 years and there have been no safety concerns raised with the consumption of the protein (Delaney *et al*., 2008).

As stated in Section 4.2.2, the AAD-12 protein is also known as SdpA. While the TfdA enzyme, to which the SdpA and RdpA enantiospecific enzymes are related, is widespread in soil bacteria, few bacterial strains have been isolated that contain the *sdpA* and *rdpA* genes (refer to Section 3.2.2). There have also been few studies on the presence, natural abundance and activation of the *rdpA* and *sdpA* genes. However, it is known that the *sdpA* gene is present in bacteria in unsprayed agricultural soil (Paulin *et al*., 2010). In crops that are sprayed with phenoxyalkanoic herbicides, a significant amount of spray reaches the ground and can lead to an increase in the abundance of phenoxyalkanoic herbicide degraders (Gazitúa *et al*., 2010; Paulin *et al*., 2010). Therefore, humans would be exposed to the AAD-12 (SdpA) protein, particularly in those areas where phenoxyalkanoic herbicides are sprayed.

4.5.2 Amino acid sequence similarity to known protein toxins

Similarity searches were done for the 2m EPSPS and AAD-12 proteins, using the BLASTP[5] (Basic Local Alignment Search Tool Protein) algorithm (Altschul *et al*., 1997) against protein sequences present in a number of large public reference databases including 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. This detects more

[5] BLASTP is used to compare a protein sequence with a database of protein sequences.

similarities that 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. Although a statistically significant sequence similarity generally requires a match with an *E*-value of less than 0.01 (Pearson, 2000), setting a threshold *E*-value of 1.0 ensures that proteins with even limited similarity will not be excluded. 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).

The AAD-12 similarity search identified 1,322 alignments with an E value of < 1.0, of which 1,267 had E values < 0.01. The most predominant classes were taurine dioxygenases (382), ketoglutarate dehydrogenase (220), taurine catabolism dioxygenases (170), dioxygenases (166), 2,4-D/ α-ketoglutarate dioxygenases (70) and α-ketoglutarate dioxygenases (30). Taurine dioxygenase is related to TfdA and has been documented to share homology with AAD-12 (Westendorf *et al*., 2002). None of these protein classes or any other of the identified proteins is associated with toxicity to humans or other animals.

The 2m EPSPS similarity search identified 3,491 alignments with an E value of < 1.0, of which 2,565 had E values < 0.01. The most predominant classes were 3-phosphoshikimate 1-carboxyvinyltransferase (1,864), UDP-N-acetylglucosamine 1-carboxyvinyltransferase (348), and 5-enolpyruvyl shikimate 3-phosphate synthase (141). None of these protein classes or any other of the identified proteins is associated with toxicity to humans or other animals.

4.5.3 In vitro digestibility

See Section 4.6.3.

4.5.4 Stability to heat

The thermolability of a protein provides an indication of the stability of the protein under cooking/processing conditions. The study provided for the AAD-12 protein (Schafer 2008) has already been considered previously by FSANZ under Application A1046 (FSANZ, 2011a). This showed that while no significant changes to protein structure were indicated by SDS-PAGE after any of the heat treatments (up to 95^{o} C) except autoclaving, the enzymatic activity of the protein was eliminated by any heating to 50^{o} C or above.

Bacterially-derived 2m EPSPS protein was solubilised in buffer and then either kept on ice or incubated at 25, 37, 55, 75, and 95°C for 30 min. The intactness of the protein was detected by gel electrophoresis (SDS-PAGE) coupled with a Coomassie blue staining, and by ELISA. Enzyme activity of 2mEPSPS was determined based upon the presence or absence of inorganic phosphate as measured by the malachite green dye assay (see Section 4.4.1).

SDS-PAGE showed that as the temperature increased additional multimeric aggregates became apparent and there was also some degradation of the protein's primary structure ,as indicated by a fading in band intensity with increasing temperature. The ELISA analysis indicated that the epitopes for the polyclonal antibody were significantly altered once the protein was heated to or above 55°C. Enzyme activity was significantly diminished after heating. At 55 °C, approximately 73% of the activity was lost and at temperatures above 75 °C, approximately 94% of the activity was lost. Taken together, the results indicate that

heating above 55^{o} C denatures the 2m EPSPS protein resulting in a loss of immunoreactivity and enzymatic activity.

4.5.5 Acute oral toxicity study

Acute oral toxicity studies in mice using bacterially-produced 2m EPSPS and AAD-12 proteins were submitted by the Applicant but are not included in this safety assessment since no safety concerns were identified in any of the other studies.[6]

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.

As with the assessment of potential toxicity, the PAT protein is not considered in this assessment of allergenicity as it has been adequately characterised elsewhere (see e.g. Delaney *et al*., 2008). A bioinformatics analysis was submitted by the Applicant comparing the PAT protein sequence with sequences in an updated (February 2011) protein database but none of the significant sequence alignments were associated with any known protein allergens.

Studies submitted:

Guttikonda, S. (2011). Similarity assessment of 2mEPSPS protein to known allergens by bioinformatics analysis (updated March, 2011). Study ID# 110328, Dow AgroSciences (unpublished).

Song, P. (2011). Sequence similarity assessment of AAD-12 protein to known allergens by bioinformatics analysis (updated February, 2011). Study ID# 110326, Dow AgroSciences (unpublished).

Embrey, S.K. (2011). *In vitro* simulated gastric fluid digestibility study of double mutant 5-enolpyruvylshikimate-3-phosphate synthase (2mEPSPS) protein. Study ID# 102106, Dow AgroSciences (unpublished).

Embrey, S.K.; Schafer, B.W. (2008). *In vitro* simulated gastric fluid digestibility of aryloxyalkanoate dioxygenase-12 (abbreviation AAD-12). Study ID# 080064, Dow AgroSciences (unpublished).

Schafer, B.W.; Embrey, S.K. (2011). *In vitro* simulated gastric fluid digestibility of aryloxyalkanoate dioxygenase-12 (AAD-12) in soybean leaf extracts. Study ID# 110265, Dow AgroSciences (unpublished).

[6] Acute oral toxicity studies are available at http://www.foodstandards.gov.au/foodstandards/applications/applicationa1073food5541.cfm

Song, P. (2011). Sequence similarity assessment of PAT protein to known allergens by bioinformatics analysis (updated February, 2011). Study ID# 110330, Dow AgroSciences (unpublished).

4.6.1 Source of the protein

The 2m EPSPS protein is derived from *Zea mays*. Few food-induced allergic reactions to corn have been reported and corn tends not be regarded as a commonly allergenic food (OECD, 2002). Nevertheless, there are allergenicity concerns associated with certain proteins expressed in the pollen of corn, e.g. Zm13 protein (Heiss *et al*., 1996), and a number of allergens have also been isolated from the seed, although most have not been clinically evaluated for their allergenic potential. The most significant of these is Zm14 (Pastorello *et al*., 2000), a lipid transfer protein that can maintain its IgE binding capacity even after heat treatment (Pastorello *et al*., 2003). Zm14 has been associated with anaphylactic reaction in susceptible individuals. The EPSPS proteins, from corn or other species, have not themselves been implicated in any food-related allergic reactions.

The AAD-12 protein is derived from a common soil bacterium *Delftia acidovorans* which has not been reported to be a source of allergenic proteins. Humans would be inadvertently exposed to this organism through the ingestion of fresh produce.

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 2m EPSPS and AAD-12 sequences were compared with all known allergen sequences contained in a reference allergen database (Food Allergy Research and Resource Program - FARRP, version 11 – released in January 2011) using the FASTA algorithm and BLOSUM50 scoring matrix. The criteria used to indicate potential allergenicity in both searches were a minimum of 8-contiguous amino acid identity or 35% identity on a window of 80 amino acids within the sequence of an allergenic protein.

No matches were found in either search with known allergenic proteins or with known allergenic epitopes.

Neither protein shared a sequence of eight or more consecutive identical amino acids or 35% or greater identity over 80 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.

A pepsin digestibility assay (Thomas *et al*., 2004) was conducted to determine the digestive stability of the 2m EPSPS and AAD-12 proteins purified from a bacterial system and AAD-12 contained in lyophilized leaf tissue from soybean 44406, using simulated gastric fluid (SGF) (U.S.Pharmacopeia, 1995). For AAD-12 the test material was evaluated following incubation in SGF at 37º for 0.5, 1, 2, 4, 8, and 16 minutes). For 2m EPSPS, the incubation times were 1, 2, 4, 8, and 16 minutes The equivalent of a zero time point was prepared by neutralizing the SGF with sodium carbonate prior to adding the test protein sample. All samples were then run on SDS-PAGE. Proteins were visualised by GelCode Blue staining of the resulting gel. Two control proteins were treated in parallel: bovine serum albumin (BSA) is known to hydrolyse readily in pepsin and served as a positive control; β-lactoglobulin is known to persist in pepsin and was used as a negative control. For the soybean 44406 leaf extract, a further control was tested – leaf extract from 'Maverick' spiked with microbial-derived AAD-12. Western blotting of the SDS gels was also performed using an appropriate rabbit polyclonal primary antibody and horseradish peroxidase-linked secondary antibody.

The AAD-12 protein was not detectable after 30 s, while the 2m EPSPS protein was not detectable after 1 min. This indicates that both proteins are readily digested by pepsin under simulated gastric conditions.

4.6.4 Stability to heat

See Section 4.5.4.

4.7 Conclusion

Soybean line DAS-44406-6 expresses three novel proteins, 2m EPSPS, AAD-12 and PAT. Expression analyses of the three proteins showed that all three were detected in the plant parts tested. The 2m EPSPS protein was lowest in the seed (approximately 22 µg/g dry weight) and highest in V10 – 12 leaves (approximately 2323 µg/g dry weight). AAD-12 was lowest in the roots (approximately 25 µg/g dry weight) and highest in V10 – 12 leaves (approximately 116 µg/g dry weight). PAT protein concentrations were much lower than those for AAD-12 but similarly, the V10 – 12 leaves contained the highest levels (approximately 10 µg/g dry weight) and the roots contained the lowest levels (approximately 2 µg/g dry weight).

A number of studies were used to confirm the identity and physicochemical properties of the plant-derived 2m EPSPS, AAD-12 and PAT proteins. These studies demonstrated that all three proteins conform in size and amino acid sequence to that expected, and do not exhibit any post-translational modification including glycosylation.

In relation to potential toxicity and allergenicity, the PAT protein has previously been demonstrated to be non-toxic to mammals and also does not exhibit any potential to be allergenic to humans.

For the 2m EPSPS and AAD-12 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 showed that both proteins are inactivated by heating. Taken together, the evidence indicates that neither 2m EPSPS nor AAD-12 are likely to be toxic or allergenic to humans.

5. Other novel substances

As part of the safety assessment it is important to establish whether the expression of a novel protein(s) is likely to result in the accumulation of other novel substances. If such substances are found to occur as a result of the genetic modification, then it is important to determine their potential toxicity.

In the case of soybean DAS-44406-6, the metabolic profiles resulting from the novel protein x herbicide interactions have been assessed in previous applications (for 2m EPSPS/glyphosate, see e.g. Application A1051 (FSANZ, 2011b); for AAD-12/2,4-D see Application A1046 (FSANZ, 2011a)) or, in the case of PAT/glufosinate ammonium, have significant history of use. There are no concerns that the spraying of soybean 44406 with glyphosate, 2,4-D or glufosinate ammonium would result in the production of metabolites that are not also produced in crops sprayed with the same herbicides and already used in the food supply.

6. 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).

6.1 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 done by the Applicant were based on this previous version.

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 yielded comparable results to those obtained for seed.

Studies submitted:

Lepping, M. (2011). Field production and agronomic characetristics of a transformed soybean cultivar containing aryloxyalkanoate dioxygenase-12 (AAD-12) double mutant maize EPSPS gene (2mEPSPS) and phosphinothricin acetyltransferase (PAT) – event DAS-44406-6. Study ID# 101104.01, Dow AgroSciences (unpublished).

Lepping, M. (2011). Nutrient composition of a transformed soybean cultivar containing aryloxyalkanoate dioxygenase-12 (AAD-12) double mutant maize EPSPS gene (2mEPSPS) and phosphinothricin acetyltransferase (PAT) – event DAS-44406-6. Study ID# 101104.03, Dow AgroSciences (unpublished).

6.2 Study design and conduct

The test (DAS-44406-6, seed of T4 lineage), and control ('Maverick') lines were grown under similar conditions at 10 field sites across North America[7] during the 2010 growing season. 'Maverick' is the original transformed line and therefore represents the isogenic control line for the purposes of the comparative analyses. Six different non-GM soybean lines were also grown under the same conditions in order to generate a reference range for each analyte. The purity and identity of the test, control and reference lines was verified.

Five herbicide treatments, at target rates already described in Section 4.3, were applied to soybean 44406, namely: unsprayed; sprayed with 2,4-D (Weedar 64); sprayed with glufosinate (Liberty); sprayed with glyphosate (Durango DMA) and sprayed with 2,4-D + glufosinate + glyphosate. Seeds of the test and control soybean lines were planted in a randomised complete block design, with four replicated plots at each of the 10 sites. The reference lines were randomised across sites in a balanced incomplete-block design.

Seed and forage from soybean 44406 and 'Maverick' were harvested from all replicated plots and analysed for composition. Forage was collected at the R3 plant growth stage, and seed (R_4) was harvested at physiological maturity (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 Software[8] (SAS) data sets and analysed using SAS version 9.2. 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 between the unsprayed and each of the 3 sprayed treatments using t-tests. Probability values were adjusted using False Discovery Rate 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 44406 and 'Maverick' a 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

[7] The 10 sites were: Sycamore, GA; Richland, IA ; Bagley, IA; Carlyle, IL; Wyoming, IL; Sheridan, IN; Deerfield, MI; Fisk, MO; Brunswick, NE; York, NE.

[8] SAS website - http://www.sas.com/technologies/analytics/statistics/stat/index.html

- 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.
- A combined literature range for each analyte, compiled from published literature[9]. Any mean value for a soybean 44406 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.

6.3 Seed composition

6.3.1 Proximates and fibre

Results of the proximate and fibre analysis are shown in Table 4. Significant overall treatment effects were found for protein, carbohydrate and NDF. For carbohydrate, the means for all 44406 treatments were lower than the 'Maverick' mean, while for protein and NDF the main contributor to the overall effect came largely from a single treatment. All mean results were within both the reference range and the range reported in the literature.

Table 4: Mean percentage of proximates and fibre in seed from 'Maverick' and DAS-44406-6.

Analyte	'Mav'	44406 no spray	44406 + 2,4-D	44406 + glufos	44406 + glyphos	44406 + 3 herb	Overall treat effect (P-value)	Reference range	Combined literature range
Protein (%dw)	37.8	38.0	38.5	38.2	38.2	38.6[2]	0.039	35.1 – 44.9	32 – 48.4
Fat (%dw)	18.9	19.5	19.2	19.3	19.2	19.1	NS	15.3 – 22.9	8.1 – 24.7
Ash (%dw)	5.15	5.23	5.24	5.24	5.21	5.22	NS	4.45 – 6.3	3.8 – 6.9
Moisture (%fw)	10.6	10.2	10.0	9.9	9.9	9.9	NS	7.26 – 17.2	4.7 – 34.4
Carbohydrate (%dw)[1]	38.13	37.2	37.1	37.21	37.38	37.04	0.002	28.7 – 43.0	29.3 – 50.2
Total dietary fibre (%dw)	22.4	21.6	21.7	21.5	21.9	21.4	NS	16.2 – 27.7	Not reported
ADF (%dw)	15.5	15.2	15.5	15.0	15.6	14.9	NS	8.02 – 20.9	7.81 – 26.6
NDF (%dw)	17.7	17.0	17.1	16.7	17.3	16.5	0.03	13.3 – 22.2	8.53 – 23.9

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

[2] Mauve shading represents 44406 means where a pairwise contrast t-test showed a significantly lower value than for the control mean (using an adjusted P value), while orange shading represents 44406 means that were significantly higher than the control.

[9] References included Iskander (1987); Hartwig & Kilen (1991); Padgette et al (Padgette *et al*., 1996); Taylor et al (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)

6.3.2 Fatty acids

The levels of 22 fatty acids were measured. Of these, the following had > 50% of samples below the limit of quantitation (LOQ) 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:

- There was no significant difference in the combined site analyses between the control and soybean 44406 in terms of the levels of stearic, arachidic and eicosenoic acids.
- The mean levels of palmitic and oleic acids were significantly lower in all soybean 44406 treatments than in 'Maverick' but all means were within both the reference range and the combined literature range.
- The mean levels of linoleic and linolenic acids were significantly higher in all soybean 44406 treatments than in 'Maverick' but all means were within both the reference range and the combined literature range.
- Overall, the levels of behenic acid in the soybean 44406 treatments were higher than in 'Maverick' but in a pairwise contrast (t-test) only the three-herbicide treatment mean was significantly higher than the control mean. All means were within both the reference range and the combined literature range.

Table 5: Mean percentage composition, relative to total fat, of major fatty acids in seed from 'Maverick' and DAS-44406-6

Fatty Acid	'Mav' %total	44406 no spray %total	44406 + 2,4-D %total	44406 + glufos %total	44406 + glyphos %total	44406 + 3 herb %total	Overall treat effect (P-value)	Reference range %total	Combined literature range %total
Palmitic (16:0)	10.9	10.7[1]	10.7	10.7	10.7	10.6	<0.001	9.5 – 11.3	9.5 – 15.7
Stearic (18:0)	4.51	4.47	4.47	4.48	4.48	4.52	NS	3.28 – 4.98	2.59 – 5.88
Oleic (18:1)	23.5	21.4	21.5	21.8	21.7	21.9	<0.001	18.1 – 27.9	14.3 – 45.6
Linoleic (18:2)	53.0	54.8	54.6	54.4	54.5	54.2	<0.001	50.1 – 56.7	35.36 – 58.8
Linolenic (18:3)	7.32	7.77	7.79	7.86	7.76	7.92	<0.001	4.83 – 9.82	3 – 12.52
Arachidic (20:0)	0.328	0.323	0.323	0.325	0.323	0.327	NS	0.254 – 0.427	0.163 – 0.57
Eicosenoic (20:1)	0.169	0.171	0.171	0.171	0.172	0.168	NS	<LOQ - 0.272	<LOQ – 0.35.
Behenic (22:0)	0.326	0.332	0.331	0.332	0.328	0.335	0.009	0.29 – 0.454	0.277 – 0.595

[1] Mauve shading represents 44406 means where a pairwise contrast t-test showed a significantly lower value than for the control mean (using an adjusted P value), while orange shading represents 44406 means that were significantly higher than the control.

6.3.3 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. Overall, the only significant difference was for cysteine where the pairwise analysis showed three of the 44406 treatment means were

significantly higher than the control mean. All means were within both the reference range and the combined literature range.

Table 6: Mean percentage dry weight (dw), relative to total dry weight, of amino acids in seed from 'Maverick' and DAS-44406-6

Amino Acid	'Mav' %dw	44406 no spray %dw	44406 + 2,4-D %dw	44406 + glufos %dw	44406 + glyphos %dw	44406 + 3 herb %dw	Overall treat effect (P-value)	Reference range %dw	Combined literature range %dw
Alanine	1.68	1.67	1.67	1.68	1.67	1.68	NS	1.55 – 1.9	1.43 – 2.10
Arginine	2.74	2.71	2.75	2.74	2.74	2.75	NS	2.59 – 3.45	2.15 – 3.46
Aspartate	4.18	4.19	4.23	4.22	4.22	4.24	NS	3.58 – 4.94	3.81 – 6.04
Cysteine	0.532	0.556[1]	0.547	0.56	0.549	0.56	0.002	0.429 – 0.71	0.37 – 0.81
Glutamate	6.26	6.22	6.26	6.27	6.26	6.28	NS	5.85 – 7.77	5.84 – 9.15
Glycine	1.64	1.63	1.65	1.64	1.65	1.65	NS	1.55 – 1.89	1.41 – 2.00
Histidine	1.02	1.02	1.04	1.02	1.03	1.03	NS	0.197 – 1.18	0.86 – 1.24
Isoleucine	1.81	1.81	1.83	1.81	1.83	1.82	NS	1.68 – 2.18	1.49 – 2.08
Leucine	2.84	2.83	2.87	2.85	2.86	2.86	NS	2.68 – 3.32	2.2 – 4.0
Lysine	2.46	2.48	2.53	2.48	2.49	2.49	NS	2 – 3.04	2.19 – 3.32
Methionine	0.504	0.502	0.505	0.501	0.499	0.507	NS	0.418 – 0.596	0.39 – 0.68
Phenylalanine	1.9	1.89	1.91	1.9	1.91	1.91	NS	1.8 – 2.28	1.6 – 2.44
Proline	1.96	1.95	1.95	1.95	1.96	1.95	NS	1.78 – 2.41	1.63 – 2.28
Serine	1.79	1.80	1.79	1.82	1.82	1.83	NS	1.56 – 2.2	1.11 – 2.48
Threonine	1.52	1.51	1.52	1.52	1.53	1.53	NS	1.4 – 1.75	1.14 – 1.89
Tryptophan	0.574	0.588	0.589	0.575	0.583	0.583	NS	0.495 – 0.704	0.30 – 0.67
Tyrosine	1.44	1.44	1.46	1.45	1.45	1.45	NS	1.34 – 1.64	0.79 – 1.61
Valine	1.85	1.85	1.86	1.85	1.86	1.85	NS	1.71 – 2.16	1.5 – 2.44

[1]Orange shading represents 44406 means where a pairwise contrast t-test showed a significantly higher value than for the control mean (using an adjusted P value)

6.3.4 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. Levels of the three parent isoflavones were below the limit of quantitation and no data are presented. Levels for the conjugates (comprising glucoside esters and β-glucosides and reported as aglycon equivalents) are given in Table 7 and show that the mean levels of

total daidzein, total genistein and total glycitein in soybean 44406 were not significantly different from those in 'Maverick'.

Table 7: Mean weight (µg/g dry weight expressed as aglycon equivalents) of isoflavones in DAS- 44406-6 and 'Maverick' seed

Isoflavone	'Mav' (µg/g dw)	44406 no spray (µg/g dw)	44406 + 2,4-D (µg/g dw)	44406 + glufos (µg/g dw)	44406 + glyphos (µg/g dw)	44406 + 3 herb (µg/g dw)	Overall treat effect (P-value)	Reference range (µg/g dw)	Combined literature range (µg/g dw)
Total Daidzein (aglycon equivalents)	809	777	799	800	771	781	NS	153 - 1710	25 - 2453
Total Genistein (aglycon equivalents)	890	863	870	877	848	855	NS	205 - 1980	28 - 2837
Total Glycitein (aglycon equivalents)	453	459	465	452	448	443	NS	85 - 1630	15 - 349

6.3.5 Anti-nutrients

Levels of key anti-nutrients are given in Table 8. Overall, there were no significant differences between the 44406 means and the control mean for phytic acid, raffinose and stachyose. For lectin and trypsin inhibitor some of the 44406 treatment means were significantly higher than the control mean. All means were within both the reference range and the combined literature range.

Table 8: Mean levels of anti-nutrients in DAS-44406-6 and 'Maverick' seed.

Anti-nutrient	'Mav' (µg/g dw)	44406 no spray (µg/g dw)	44406 + 2,4-D (µg/g dw)	44406 + glufos (µg/g dw)	44406 + glyphos (µg/g dw)	44406 + 3 herb (µg/g dw)	Overall treat effect (P-value)	Reference range (µg/g dw)	Combined literature range (µg/g dw)
Lectin (Haemagglutinat Units/mg protein dw)	79	107[1]	94	94	92	99	0.004	18.5 - 144	37 - 323
Phytic acid (%dw)	1.19	1.18	1.19	1.18	1.18	1.19	NS	0.55 – 1.54	0.41 – 2.74
Raffinose (%dw)	0.82	0.8	0.8	0.8	0.8	0.79	NS	0.57 – 1.4	0.21 – 1.62
Stachyose (%dw)	3.88	3.88	3.87	3.86	3.83	3.89	NS	2.92 – 4.48	1.21 – 6.1
Trypsin inhibitor (trysin inhibitor units/mg)	30.8	35	35.5	36.6	33.6	34.2	0.025	15.6 – 59.7	18.14 – 118.68

[1] Orange shading represents 44406 means where a pairwise contrast t-test showed a significantly higher value than for the control mean (using an adjusted P value)

6.3.6 Minerals

Levels of 10 minerals were measured. Sodium was below the level of quantitation. The means for the remaining nine minerals are given in Table 9 and show:

- The mean calcium levels in three of the DAS-44406-6 treatments were significantly higher than the mean for 'Maverick'.
- For potassium, the means for all 44406 treatments were significantly higher than the mean for 'Maverick'. Additionally, the mean potassium level in 'Maverick' as well as in all the 44406 treatments was outside (lower than) the literature range.

- The copper results for all DAS-44406-6 means and the 'Maverick' mean were higher than the combined literature range but within the reference range.

These results do not raise any safety concerns.

Table 9: Mean values for mineral levels in seed from 'Maverick' and DAS-44406-6.

Mineral	'Mav'	44406 no spray	44406 + 2,4-D	44406 + glufos	44406 + glyphos	44406 + 3 herb	Overall treat effect (P-value)	Reference range	Combined literature range
Calcium (mg/100g dw)	301	324[1]	318	304	320	306	<0.001	174 - 383	116 - 510
Copper (mg/100g dw)	1.32	1.34	1.35	1.32	1.35	1.33	NS	0.91 – 1.77	0.63 – 1.09
Iron (mg/100g dw)	8.2	8.3	7.7	8.6	7.8	8.5	NS	5.35 – 87.9	3.73 – 10.95
Magnesium (mg/100g dw)	229	231	230	227	230	229	NS	195 - 317	219 - 312
Manganese (mg/100g dw)	2.99	3.1	3.27	3.09	3.18	3.14	NS	1.9 – 9.53	2.52 – 3.87
Phosphorus (mg/100g dw)	557	561	558	554	558	557	NS	360 - 659	506 - 935
Potassium (mg/100g dw)	1730	1780	1790	1770	1770	1770	<0.001	1530 - 2030	1868 - 2510
Selenium (ppb)	451	438	489	389	420	469	NS	59.5 - 3380	Not reported
Zinc (mg/100g dw)	4.17	4.34	4.34	4.25	4.3	4.3	NS	3.34 – 5.82	4.98 – 7.57

[1] Orange shading represents 44406 means where a pairwise contrast t-test showed a significantly higher value than for the control mean (using an adjusted P value)

6.3.7 Vitamins

Levels of 13 vitamins were measured. Vitamin A and β-tocopherol were below the LOQ. The means for the remaining 11 are given in Table 10.

- The mean levels of Vitamin B1 in all the 44406 treatments and in the control were all outside (higher than) the literature range but were within the reference range. No significant differences from the control were identified.
- The mean level of vitamin B9 in the 44406 treatment sprayed with all three herbicides was significantly lower than the mean level in 'Maverick but within both the reference range and literature range.
- The mean levels of ɣ-tocopherol in three of the 44406 treatments were significantly higher than in the control. This difference was also reflected in the mean levels of total tocopherols. All means were within the reference range.

Table 10: Mean weight (µg/g dry weight) of vitamins in seed from 'Maverick' and DAS-44406-6

Vitamiin	'Mav' (µg/g dw)	44406 no spray (µg/g dw)	44406 + 2,4-D (µg/g dw)	44406 + glufos (µg/g dw)	44406 + glyphos (µg/g dw)	44406 + 3 herb (µg/g dw)	Overall treat effect (P-value)	Reference range (µg/g dw)	Combined literature range (µg/g dw)
Vitamin B1 (thiamin)	3.64	3.42	3.46	3.32	3.38	3.55	NS	1.65 – 5.48	1.01 – 2.54
Vitamin B2 (riboflavin)	3.99	3.90	3.88	3.99	3.88	3.77	NS	2.72 – 4.76	1.9 – 3.21
Vitamin B3 (niacin)	26.5	26.2	26.5	26.3	26.6	26.4	NS	20.1 - 33	Not reported
Vitamin B5 (pantothenate)	15.4	15.9	15.6	15.3	15.6	15.8	NS	9.55 – 18.1	Not reported
Vitamin B6 (pyridoxine)	4.89	4.85	4.86	4.79	4.94	4.93	NS	2.77 – 6.2	Not reported
Vitamin B9 (Folic acid)	4.29	4.09	4.03	4.02	4.07	3.92[1]	0.029	2.35 – 5.98	2.38 – 4.71
Vitamin C	121.4	107.8	111.1	112.1	112.8	113.6	NS	0 - 141	Not reported
Vitamin E (α-tocopherol)	18.6	22.2	22.4	22.2	21.9	22.2	NS	6.43 – 49.9	0.108 – 61.7
γ-tocopherol	174	185	184	179	183	181	0.005	116 - 215	Not reported
δ-tocopherol	73.3	72	71.5	73.8	72.1	72.2	NS	40 - 114	Not reported
Total tocopherol	266	279	278	275	277	276	0.035	199 - 321	Not reported

[1] Mauve shading represents 44406 means where a pairwise contrast t-test showed a significantly lower value than for the control mean (using an adjusted P value), while orange shading represents 44406 means that were significantly higher than the control.

6.3.8 Summary of analysis of key components

Statistically significant differences in the analyte levels found in seed of soybean 44406 and 'Maverick' are summarised in Table 11. These statistically significant differences do not raise safety concerns, given that:

- there are no trends in the results,
- for the majority of analytes there were no significant differences, and
- for all the analytes where there was a significant difference, the soybean 44406 means fall within both the reference range and (where it exists) the literature range.

Table 11: Summary of analyte means found in seed of DAS-44406-6 treatments that are significantly (adj. P<0.05) different from those found in seed of the control line 'Maverick'

Analyte	'Mav' %dw	44406 no spray %dw	44406 + 2,4-D %dw	44406 + glufos %dw	44406 + glyphos %dw	44406 + 3 herb %dw	DAS-44406-6 within reference range?	DAS-44406-6 within lit range?
Protein (%dw)	37.8	38.0	38.5	38.2	38.2	38.6[1]	yes	yes
Fat (%dw)	18.9	19.5	19.2	19.3	19.2	19.1	yes	yes
Carbohydrate (%dw)[1]	38.13	37.2	37.1	37.21	37.38	37.04	yes	yes
NDF (%dw)	17.7	17.0	17.1	16.7	17.3	16.5	yes	yes
Palmitic Acid (% total fat)	10.9	10.7[1]	10.7	10.7	10.7	10.6	yes	yes
Oleic Acid (% total fat)	23.5	21.4	21.5	21.8	21.7	21.9	yes	yes
Linoleic Acid (% total fat)	53.0	54.8	54.6	54.4	54.5	54.2	yes	yes
Linolenic Acid (% total fat)	7.32	7.77	7.79	7.86	7.76	7.92	yes	yes
Behenic Acid (% total fat)	0.326	0.332	0.331	0.332	0.328	0.335	yes	yes
Cystine (%dw)	0.532	0.556	0.547	0.56	0.549	0.56	yes	yes
Lectin (Haemagglutinat Units/mg protein dw)	79	107	94	94	92	99	yes	yes
Trypsin inhibitor (trysin inhibitor units/mg)	30.8	35	35.5	36.6	33.6	34.2	yes	yes
Calcium (mg/100g dw)	301	324	318	304	320	306	yes	yes
Potassium (mg/100g dw)	1730	1780	1790	1770	1770	1770	yes	yes
Vitamin B9 (Folic acid)	4.29	4.09	4.03	4.02	4.07	3.92	yes	yes
γ-tocopherol	174	185	184	179	183	181	yes	Not reported
γ-tocopherol	174	185	184	179	183	181	yes	Not reported

[1] Mauve shading represents 44406 means where a pairwise contrast t-test showed a significantly lower value than for the control mean (using an adjusted P value), while orange shading represents 44406 means that were significantly higher than the control.

6.4 Conclusion

The compositional analyses do not indicate any differences of biological significance between the seed from soybean DAS-44406-6 and the non-GM control 'Maverick'. Significant differences were noted in a number of constituents. However the differences were typically small and all mean values were within both the tolerance 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. The spraying of soybean line DAS-44406-6 with 2,4-D, glufosinate ammonium and glyphosate, either alone or in combination, did not have a significant effect on seed composition.

7. 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 44406 is the result of a simple genetic modification to confer herbicide tolerance 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 44406, indicate it is equivalent in composition to conventional soybean cultivars. The introduction of soybean line DAS-44406-6 into the food supply is therefore expected to have little nutritional impact.

References[10]

Abranches, R., Shultz, R.W., Thompson, W.F. and Allen, G.C. (2005) Matrix attachment regions and regulated transcription increase and stabilize transgene expression. *Plant Biotechnology Journal* 3(5):535-543.

Abu-Qarn, M., Eichler, J. and Sharon, N. (2008) Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Current Opinion in Structural Biology* 18:544-550.

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.

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.

Baelum, J., Nicolaisen, M.H., Holben, W.E., Strobel, B.W., Sorensen, J. and Jacobsen, S. (2008) Direct analysis of *tdfA* gene expression by indigenous bacteria in phenoxy acid amended agricultural soil. *The ISME Journal* 2:677-687.

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.

[10] All website references were current as at 22 January 2013

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* :

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.

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.

Chaboute, M.-E., Chaubet, N., Philipps, G., Ehling, M. and Gigot, C. (1987) Genomic organization and nucleotide sequences of two histone H3 and two histone H4 genes of *Arabidopsis thaliana*. *Plant Molecular Biology* 8:179-191.

Chun, J., Lee, J., Bae, J., Kim, M., Lee, J.-G., Shin, S.-Y., Kim, Y.R. and Lee, K.-H. (2009) Delftia acidovorans from the drainage in an immunocompetent patient with empyema. *Tuberculosis and Respiratory Diseases* 67(3):239-243.

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. (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.

El-Banna, N.M. (2007) Antifungal activity of *Comamonas acidovorans* isolated from water pond in south Jordan. *African Journal of Biotechnology* 6(19):2216-2219.

FAOSTAT. (2011) Online database of the Food and Agriculture Organization of the United Nations. http://faostat.fao.org/.

FSANZ (2010) *Application A1042 - Food derived from herbicide-tolerant corn line DAS-40278-9*. Report prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/foodstandards/applications/applicationa1042food4758.cfm.

FSANZ (2011a) *Application A1046 - Food derived from herbicide-tolerant soybean line DAS-68416-4*. Report prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/foodstandards/applications/applicationa1046food4807.cfm.

FSANZ (2011b) *Application A1051 - Food derived from Herbicide-tolerant Soybean Event FG72*. Report prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/foodstandards/applications/applicationa1051food4902.cfm.

Funke, T., Yang, Y., Healy-Fried, M.L., Olesen, S., Becker, A. and Schönbrunn, E. (2009) Structural basis of glyphosate resistance resulting from the double mutation Thr973Ile and Pro1013Ser in 5-enolpyruvylshikimate-3-phosphate synthase from *Escherichia coli*. *The Journal of Biological Chemistry* 284(15):9854-9860.

Gazitúa, M.C., Slater, A.W., Melo, F. and González, B. (2010) Novel a-ketoglutarate dioxygenase *tfdA*-related genes are found in soil DNA after exposure to phenoxyalkanoic herbicides. *Environmental Microbiology* e-pub, doi: 10.1111/j.1462-2920.2010.02215.x:

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.

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.

Hall, G., Jr., Allen, G.C., Loer, D.S., Thompson, W.F. and Spiker, S. (1991) Nuclear scaffolds and scaffold-attachment regions in higher plants. *Proceedings of the National Academy of Sciences* 88(20):9320-9324.

Han, J., Sun, L., Dong, X., Cai, Z., Sun, X., Yang, H., Wang, Y. and Song, W. (2005) Characterization of a novel plant growth-promoting bacteria strain *Delftia tsuruhatensis* HR4 both as a diazotroph and a potential biocontrol agent against various plant pathogens. *Systematic and Applied Microbiology* 28:66-76.

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.

Heiss, S., Flicker, S., Hamilton, D.A., Kraft, D., Mascarenhas, J.P. and Valenta, R. (1996) Expression of Zm13, a pollen specific maize protein in *Escherichia coli* reveals IgE-binding capacity and allergenic potential. *FEBS Letters* 381:217-221.

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.

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.

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.

IPCS (1984) *2,4-Dichlorophenoxyacetic acid (2,4-D)*. Report No. Environmental Health Criteria 29, International Programme on Chemical Safety INCHEM. http://www.inchem.org/documents/ehc/ehc/ehc29.htm.

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

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.cropscience.org.au/icsc2004/poster/3/4/6/327_jamesat.htm#TopOfPage.

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.

Labuda, I.M., Goers, S.K. and Keon, K.A. (1992) Bioconversion process for the production of vanillin. (US Patent 5,128,253): http://www.freepatentsonline.com/5279950.html.

Lanzetta, P.A., Alvarez, L.J., Reinach, P.S. and Candia, O.A. (1979) An improved assay for nanomole amounts of inorganic phosphate. *Analytical Biochemistry* 100:95-97.

Lebrun, M., Leroux, B. and Sailland, A. (1996) Chimeric gene for the transformation of plants. (US5510471): http://www.patentlens.net/imageserver/getimage/US_5510471.pdf;jsessionid=7F9CD99DF5B39E5D227B1FA1D7BBEB4C?id=8737286&page=all.

Lebrun, M., Sailland, A. and Freyssinet, G. (1997) Mutated 5-enol-pyruvylshikimate-3-phosphate synthase, gene coding for said protein and transformed plants containing said gene. (WO/9704103 A2): http://www.patentlens.net/imageserver/getimage/WO_1997_004103_A2.pdf;jsessionid=88679D1B0F706258C54B42EE59A4C287?id=20537688&page=all.

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.

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) b Protein C is not glycosylated at asparagine 329. *The Journal of Biological Chemistry* 265:11397-11404.

Müller, R.H., Jorks, S., Kleinsteuber, S. and Babel, W. (1999) Comamonas acidovorans strain MC1: a new isolate capable of degrading the chiral herbicides dichlorprop and mecoprop and the herbicides 2,4-D and MCPA. *Microbiological Research* 154(3):241-246.

NDSU. (2004) Soybean Growth and Management Quick Guide: Reproductive Stages. Publication A-1174, North Dakota State University Agriculture and University Extension, http://www.ag.ndsu.edu/pubs/plantsci/rowcrops/a1174/a1174-2.htm.

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.

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 (2002) *Consensus document on compositional considerations for new varieties of maize (Zea mays): Key food and feed nutrients, antinutrients and secondard plant metabolites*. Series on the safety of novel foods and feeds, No. 6. Organisation for Economic Co-operation and Development, Paris. http://www.oecd.org/officialdocuments/displaydocumentpdf?cote=env/jm/mono(2002)25&doclanguage=en.

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.

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.

OGTR (2008) *The biology of Zea mays L. ssp mays (maize or corn)*. Prepared by the Office of the Gene Technology Regulator, Australia. http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/riskassessments-1.

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.

Padgette, 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.

Pastorello, E.A., Farioli, L., Pravettoni, V., Ispano, M., Scibola, E., Trambaioli, C., Giuffrida, M.G., Ansaloni, R., Godovac-Zimmermann, J., Conti, A., Fortunato, D. and Ortolani, C. (2000) The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. *Journal of Allergy and Clinical Immunology* 106(4):744-751.

Pastorello, E.A., Pompei, C., Pravettoni, V., Farioli, L., Calamari, A.M., Scibilia, J., Robino, A.M., Conti, A., Iametti, S., Fortunato, D., Bonomi, S. and Ortolani, C. (2003) Lipid-transfer protein is the major maize allergen maintaining IgE-binding activity after cooking at 100 degrees C, as demonstrated in anaphylactic patients and patients with positive double-blind, placebo-controlled food challenge results. *Journal of Allergy and Clinical Immunology* 112(4):775-783.

Patel, D.S., Desai, A.J. and Desai, J.D. (1998) Biodegradation of dimethylterephthalate by *Comamonas acidovorans* D-4. *Indian Journal of Experimental Biology* 36(3):321-324.

Paulin, M.M., Nicolaisen, M.H. and Sorensen, J. (2010) Abundance and expression of enantioselective *rdpA* and *sdpA* dioxygenase genes during degradation of the racemic herbicide (R,S)-2-(2,4-dichlorophenoxy)propionate in soil. *Applied and Environmental Microbiology* 76(9):2873-2883.

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.

Polevoda, B. and Sherman, F. (2002) The diversity of acetylated proteins. *Genome Biology* 3(5):reviews0006.

Schleinizt, K.M., Kleinsteuber, S., Vallaeys, T. and Babel, W. (2004) Localization and characterization of two novel genes encoding stereospecific dioxygenases catalyzing 2(2,4-dichlorophenoxy) propionate cleavage in *Delftia acidovorans* MC1. *Applied and Environmental Microbiology* 70(9):5357-5365.

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.

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.

Strauch, E., Wohlleben, W. and Pühler, A. (1988) Cloning of a phosphinothricin N-acetyltransferase gene from *Streptomyces viridochromogenes* Tü494 and its expression in *Streptomyces lividans* and *Escherichia coli*. *Gene* 63(1):65-74.

Streber, W.R., Timmis, K.N. and Zenk, M.H. (1987) Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. *Journal of Bacteriology* 169(7):2950-2955.

Sudesh, K. (2004) Microbial polyhydroxyalkanoates (PHAs): an emerging biomaterial for tissue engineering and therapeutic applications. *Medical Journal of Malaysia* 59 Suppl B:55-56.

Tamaoka, J., Ha, D.-M. and Komagata, K. (1987) Reclassification of *Pseudomonas acidovorans* den Dooren de Jong 1926 and *Pseudomonas testosteroni* Marcus and Talalay 1956 as *Comamonas acidovorans* comb. nov. and *Comamonas testosteroni* comb. nov., with an emended description of the genus *Comamonas*. *International Journal of Systematic Bacteriology* 37(1):52-59.

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

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., Crameri, 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.

U.S.Pharmacopeia. (1995) *The United States Pharmacopeia 23. The National Formulary 18*. United States Pharmacopeia Convention, Inc, Rockville, MD.

Urata, M., Uchida, E., Nojiri, H., Omori, T., Obo, R., Miyaura, N. and Ouchiyama, N. (2004) Genes involved in aniline degradation by *Delftia acidovorans* Strain 7N and its distribution in the natural environment. *Bioscience, Biotechnology and Biochemistry* 68(12):2457-2465.

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.

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 mosiac virus (CVMV) promoter. *Plant Molecular Biology* 31:1129-1139.

Verma, D., Verma, M., Dey, M., Jain, R.K. and Wu, R. (2005) Molecular dissection of the tobacco Rb7 matrix attachment region (MAR): Effect of 5' half on gene expression in rice. *Plant Science* 169(4):704-711.

Wacker, M., Linton, D., Hitchen, P.G., Nita-Lazar, M., Haslam, S.M., North, S.J., Panico, M., Morris, H.R., Dell, A., Wren, B.W. and Aebi, M. (2002) N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* 298:1790-1793.

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.

Wen, A., Fegan, M., Hayward, C., Chakraborty, S. and Sly, L.I. (1999) Phylogenetic relationships among members of the *Comamonadaceae*, and the description of *Delftia acidovorans* (den Dooren de Jong 1926 and Tamaoka et al. 1987) gen.nov., comb. nov. *International Journal of Systematic Bacteriology* 49:567-576.

Westendorf, A., Benndorf, D., Müller, R.H. and Babel, W. (2002) The two enantiospecific dichlorprop/a-ketoglutarate-dioxygenases from *Delftia acidovorans* MC1 - protein and sequence data of RdpA and SdpA. *Microbiological Research* 157(4):317-322.

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.

Wood, D.W., Setubal, J.C., Kaul, R., Monks, D.E., Kitajima, J.P., Okura, V.K., Zhou, Y., Chen, L., Wood, G.E., Almeida Jr, N.F., Woo, L., Chen, Y., Paulsen, I.T., Eisen, J.A., Karp, P.D., Bovee Sr, D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M.-J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Romero, P., Gordon, D., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z.Y., Dolan, M., Chumley, F., Tingey, S.V., Tomb, J.F., Gordon, M.P., Ison, M.V. and Ester, E.W. (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* 294(5550):2317-2323.

Wright, T.R., Lira, J.M., Walsh, T., Merlo, D.J., Jayakumar, P. and Lin, G. (2007) Novel herbicide resistance genes. (World Patent 2007/053482 A2): http://www.patentlens.net/imageserver/getimage/WO_2007_053482_A2.pdf?id=14882617&page=all.

Wright, T.R., Shan, G., Walsh, T.A., Lira, J.M., Cui, C., Song, P., Zhuang, M., Arnold, N.L., Lin, G., Yau, K., Russell, S.M., Cicchillo, R.M., Peterson, M.A., Simpson, D.M., Zhou, N., Ponsamuel, J. and Zhang, Z. (2010a) Robust crop resistance to broadleaf and grass herbicides provided by aryloxyalkanoate dioxygenase transgenes. *Proceedings of the National Academy of Sciences* ePub, doi:101073/pnas.1013154107:

Wright, T.R., Shan, G., Walsh, T.A., Lira, J.M., Cui, C., Song, P., Zhuang, M., Arnold, N.L., Lin, G., Yau, K., Russell, S.M., Cicchillo, R.M., Peterson, M.A., Simpson, D.M., Zhou, N., Ponsamuel, J. and Zhang, Z. (2010b) Robust crop resistance to broadleaf and grass herbicides provided by aryloxyalkanoate dioxygenase transgenes. *Proceedings of the National Academy of Sciences* ePub, doi:101073/pnas.1013154107:

Yoon, S.-H., Li, C., Lee, Y.-M., Lee, S.-H., Kim, S.-H., Choi, M.-S., Seo, W.-T., Yang, J.-K., Kim, J.-Y. and Kim, S.-W. (2005) Production of vanillin from ferulic acid using recombinant strains of *Escherichia coli*. *Biotechnology and Bioprocess Engineering* 10:378-384.

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

Zambryski, P., Depicker, A., Kruger, K. and Goodman, H.M. (1982) Tumor induction by *Agrobacterium tumefaciens*: analysis of the boundaries of T-DNA. *Journal of Molecular and Applied Genetics* 1(4):361-370.

Zeng, P., Vadnais, 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.