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

SUPPORTING DOCUMENT 1

APPLICATION A1051 – FOOD DERIVED FROM HERBICIDE-TOLERANT SOYBEAN LINE FG72

SAFETY ASSESSMENT REPORT (APPROVAL)

SUMMARY AND CONCLUSIONS

Background

Bayer CropScience Pty Ltd has developed a genetically modified (GM) soybean line known as FG72 that is tolerant to two herbicides, glyphosate and isoxaflutole. Tolerance to glyphosate is achieved through expression of a 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) encoded by the *2mepsps* gene derived from *Zea mays* (corn). The *epsps* gene has been widely used in the genetic modification of a number of crop species. Tolerance to isoxaflutole is achieved through expression of a modified p-hydroxyphenylpyruvate dioxygenase (HPPD) encoded by the *hppdPF W336* gene originally derived from the soil bacterium *Pseudomonas fluorescens*.

In conducting a safety assessment of food derived from herbicide-tolerant soybean line FG72, a number of criteria have been addressed including: a characterisation of the transferred genes, their origin, function and stability in the soybean genome; the changes at the level of DNA, protein and in the whole food; compositional analyses; evaluation of intended and unintended changes; and the potential for the newly expressed proteins to be either allergenic or toxic in humans.

This safety assessment report addresses only food safety and nutritional issues. It therefore does not address: 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; or the safety of food derived from the non-GM (conventional) plant.

History of Use

Soybean (*Glycine max*), the host organism 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 accounts for 94% of the soybean products consumed by humans and is used mainly as a salad and cooking oil, bakery shortening, and frying fat as well as being incorporated into processed products such as margarine.

Molecular Characterisation

Soybean line FG72 contains two novel gene cassettes, one containing the *hppdPf W36* gene and the other containing the *2mepsps* gene. There are no antibiotic resistance marker genes present in soybean line FG72.

Genetic modification of the parent line 'Jack' was achieved using a biolistic technique in which the two novel gene cassettes were contained within a single linear fragment. Comprehensive molecular analyses of soybean line FG72 indicate that there is a single insertion site comprising two partial sequences in a head to head orientation, followed by two complete copies of the linear fragment, arranged in a head to tail orientation. In addition, a genomic fragment from 'Jack' has translocated to a new position and is flanked at one end by 158 base pairs of a promoter sequence from the introduced linear fragment.

The introduced genetic elements are stably inherited from one generation to the next. Plasmid backbone analysis shows that no plasmid backbone has been incorporated into the transgenic locus. A number of unexpected ORFs are present at the junctions associated with the insertion site but lack the necessary elements for expression of a protein. No known endogenous genes have been interrupted by insertion of the new genetic material.

Characterisation of Novel Protein

Soybean line FG72 expresses two novel proteins, HPPDPf W336 and 2mEPSPS both of which were detected in all plant parts analysed. The concentration of HPPDPf W336 was lowest in the seed (approximately 1.5 µg/g dry weight) and highest in younger leaves (approximately 38 µg/g dry weight). Overall the 2mEPSPS protein concentrations were much higher than those for HPPDPf W336. The leaves contained the highest levels (older leaves contained approximately 660 µg/g dry weight) while roots contained the lowest levels (approximately 40 µg/g dry weight). The level of 2mEPSPS in the seed was approximately 150 µg/g dry weight. During processing of the seed, the HPPDPf W336 protein may be concentrated to a small degree in hulls and protein isolate, and is undetectable in other processed fractions. Levels of the 2mEPSPS protein are reduced in all fractions during processing, being undetectable in toasted meal, crude lecithin and all forms of oil.

Several studies were done to confirm the identity and physicochemical properties of the plant-derived HPPDPf W336 and 2mEPSPS proteins, and demonstrated that they both conform in size and amino acid sequence to that expected, and do not exhibit any post-translational modification including glycosylation.

For both proteins, bioinformatic studies have confirmed their lack of any significant amino acid sequence similarity to known protein toxins or allergens and digestibility studies have demonstrated that the proteins would be rapidly degraded in the stomach following ingestion. Acute toxicity studies in mice have also confirmed their absence of toxicity in animals. Both proteins exhibit a degree of heat stability however, given their digestive lability combined with their lack of similarity to known protein toxins or allergens and the loss of enzyme activity with heating, this does not raise any safety concerns.

Taken together, the evidence indicates that HPPDPf W336 and 2mEPSPS are unlikely to be toxic or allergenic to humans.

Herbicide Metabolites

The residues generated on soybean line FG72 as a result of spraying with isoxaflutole are the same as those found on conventional crops sprayed with isoxaflutole. Residue data derived from supervised trials indicate that the residue levels in seed are below the limit of quantitation and that there is some concentration of residue in meal and aspirated grain fractions, both of which are used in animal feed, but not in other processed commodities. In the absence of any measurable exposure to either parent herbicide or metabolites the risk to public health and safety is likely to be negligible.

Compositional Analyses

Detailed compositional analyses were done to establish the nutritional adequacy of seed from soybean line FG72 under both herbicide sprayed and unsprayed conditions. Analyses were done of 77 analytes encompassing proximates (crude fat/protein, carbohydrate and ash), acid detergent fibre, neutral detergent fibre, fatty acids, amino acids, isoflavones, anti-nutrients, minerals, and vitamins. The levels were compared to levels in the seeds of the non-GM parent 'Jack'.

These analyses indicated that the seeds of soybean line FG72 are compositionally equivalent to those of the parental line. Out of all the analytes tested, there were significant differences between the non-GM control and soybean FG72 in only 19 analytes. In all of these, the mean levels observed in seeds of soybean FG72 were within the range of natural variation reported in the literature. There were no consistent trends in the effect that herbicide spraying of soybean FG72 had on mean analyte levels.

Mean levels of a range of analytes were also obtained for processed products derived from soybean. There were no meaningful differences between the control and the GM line for any analyte measured in processed products used for human consumption.

In addition, no difference between seeds of soybean line FG72 and 'Jack' were found, in terms of presence and mean level of endogenous allergens.

The compositional data are consistent with the conclusion that there are no biologically significant differences in the levels of key components in seed and processed products derived from soybean line FG72 when compared with the non-GM control or with the range of levels found in non-GM commercial soybean cultivars.

Nutritional Impact

The extensive compositional analyses of seed from soybean line FG72 indicate it is equivalent in composition to conventional soybean cultivars. The introduction of soybean line FG72 into the food supply is therefore expected to have little nutritional impact. In addition, a feeding study in chickens with toasted soybean meal from soybean line FG72 indicates it is equivalent to non-GM soybean in its ability to support typical growth and well-being.

Conclusion

No potential public health and safety concerns have been identified in the assessment of soybean line FG72. On the basis of the data provided in the present Application, and other available information, food derived from soybean line FG72 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 FIGURES **1**
LIST OF TABLES **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 **6**
3.1 Method used in the genetic modification 7
3.2 Function and regulation of introduced genes 8
3.3 Breeding of soybean plants containing transformation event FG72 9
3.4 Characterisation of the genes in the plant 10
3.5 Stability of the genetic changes in soybean line FG72 16
3.6 Antibiotic resistance marker genes 17
3.7 Conclusion 17
4. CHARACTERISATION OF NOVEL PROTEINS **17**
4.1 Potential toxicity/allergenicity of ORFs created by the transformation procedure 18
4.2 Function and phenotypic effects of the HPPDPf W36 and 2mEPSPS proteins 18
4.3 Protein expression analysis 20
4.4 Protein characterisation and equivalence 24
4.5 Potential toxicity of the novel proteins 28
4.6 Potential allergenicity of the novel proteins 32
4.7 Conclusion 35
5. HERBICIDE METABOLITES **36**
5.1 Metabolism of isoxaflutole 37
5.2 Isoxaflutole residue chemistry studies 38
5.3 ADI for isoxaflutole 39
5.4 Conclusion 40
6. COMPOSITIONAL ANALYSIS **40**
6.1 Key components 40
6.2 Study design and conduct for key components 41
6.3 Analyses of key components 41
6.4 Compositional data for processed commodities 48
6.5 Assessment of endogenous allergenic potential 49
6.6 Conclusion 50
7. NUTRITIONAL IMPACT **50**
7.1 Feeding study 51
REFERENCES **52**

LIST OF FIGURES

Figure 1: Genes and regulatory elements contained in plasmid pSF10 (A) and the linear fragment (B) derived from it. 7
Figure 2: Breeding strategy for plants containing event FG72 10
Figure 3: Inserted transgenic sequences in event FG72 12
Figure 4: Representation of the arrangement of translocated genetic material (as a result of the transformation event) between A) the 5' junction and B) the 3' junction 13
Figure 5: The nine junctions identified in event FG72 14
Figure 6: Insertion sites in the 'Jack' genome 15

Figure 7: Schematic diagram of the action of HPPD in the presence and absence of isoxaflutole 19

LIST OF TABLES

Table 1: Description of the genetic elements contained in the introduced DNA fragment . 8

Table 2: Segregation data for F_2 progeny 17

Table 3: Novel protein content in FG72 soybean parts at different growth stages.. 22

Table 4: Levels of novel proteins in FG72 RAC 23

Table 5: Levels of novel proteins in processed fractions of soybean FG72 24

Table 6: Study design for acute oral toxicity testing..... 32

Table 8: Metabolites in raw agricultural commodities (forage, hay or seed) of soybean FG72 following a single spray application of isoxaflutole at either pre-emergent or full flowering stage. 37

Table 9: Levels of total isoxaflutole (sum of isoxaflutole and diketonitrile), expressed as ppm, in seed of soybean line FG72 sprayed with Balance® Pro 480 SC 38

Table 10: Levels of total isoxaflutole (sum of isoxaflutole and diketonitrile), expressed as ppm, in seed and processed fractions of soybean line FG 72 sprayed with Balance® Pro 480 SC at 5X the maximum recommended rate. 39

Table 11: Study design for each trial site 41

Table 12: Mean (±standard deviation) percentage dry weight (%dw) of proximates and fibre in seed from 'Jack' and FG72..... 42

Table 13: Mean (±standard deviation) percentage composition, relative to total fat, of major fatty acids in seed from 'Jack' and FG72..... 43

Table 14: Mean percentage dry weight (dw), relative to total dry weight, of amino acids in seed from 'Jack' and FG72..... 44

Table 15 Mean weight (µg/g dry weight) of isoflavones in soybean FG72 and 'Jack' seed..... 45

Table 16 Mean levels of anti-nutrients in soybean FG72 and 'Jack' seed. 45

Table 17: Mean values for mineral levels in seed from 'Jack' and FG72. 46

Table 18: Mean weight (µg/g dry weight) of vitamins in seed from 'Jack' and FG72 47

Table 19: Summary of analyte levels found in seed of soybean FG72 that are significantly ($P < 0.05$) different from those found in seed of 'Jack'. 47

Table 20: Summary of differences in allergens found in FG72 and 'Jack'..... 50

LIST OF ABBREVIATIONS

ADF	acid detergent fibre
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BSA	bovine serum albumin
bw	body weight
CCI	Confidential Commercial Information
DKN	diketonitrile
DNA	deoxyribonucleic acid
EPSPS	5-enolpyruvylshikimate-3-phosphate synthase
dw	dry weight
ELISA	enzyme linked immunosorbent assay
FASTA	Fast Alignment Search Tool - All
FSANZ	Food Standards Australia New Zealand
GM	genetically modified
HPLC	high performance liquid chromatography
HPP	hydroxyphenylpyruvate
HPPD	p-hydroxyphenylpyruvate dioxygenase (same as 4-hydroxyphenylpyruvate)
ILSI	International Life Sciences Institute
kDa	kilo Dalton
LCMS	liquid chromatography mass spectrometry
LOQ	limit of quantitation
NDF	neutral detergent fibre
OECD	Organisation for Economic Co-operation and Development
ORF	open reading frame
PCR	polymerase chain reaction
Z-PCR	zygosity PCR
RuBisCo	Ribulose bisphosphate carboxylase
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
U.S.	United States of America

1. Introduction

A genetically modified (GM) soybean line, FG72, has been developed that provides tolerance to the herbicides glyphosate and isoxaflutole.

Tolerance to glyphosate is achieved through expression of a 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) encoded by the *2mepsps* gene derived from *Zea mays* (corn), a common crop plant. The *epsps* gene has been widely used in the genetic modification of a number of crop species. Tolerance to isoxaflutole is achieved through expression of a modified p-hydroxyphenylpyruvate dioxygenase (HPPD) encoded by the *hppdPf W336* gene originally derived from the soil bacterium *Pseudomonas fluorescens*.

It is anticipated that soybean plants containing event FG72 may be grown in the United States of America (U.S)., Argentina, Brazil, China and India subject to approval. There is no plan to grow the 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 'Jack' was used as the parent 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 FG72. 'Jack' was released in the U.S. in August 1989 for its resistance to soybean cyst nematode and higher yield when compared to cultivars of similar maturity (Nickell *et al*., 1990).

Soybean is grown as a commercial food and feed crop in over 35 countries worldwide (OECD, 2000) 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:

- 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. This crude oil is further refined to produce cooking oil, shortening and lecithins 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 e.g. livestock, pet and poultry food), protein concentrate and isolate (for use in both edible and

[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

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

Soybean 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.

Soybean FG72 is intended primarily for use as a broad-acre commodity (field soybean) to produce products derived from cracked soybeans, and is not intended for vegetable or garden purposes where food-grade products may include tofu, soybean sprouts, soy milk, and green soybean (e.g. edamame). This latter type of soybean generally has a different size, flavour and texture to field soybean.

2.2 Donor organisms

2.2.1 Zea mays

Corn, *Zea mays*, is the source of the *epsps* gene that was modified to produce the *2mepsps* gene in soybean line FG72 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 700 million tonnes (http://faostat.fao.org/default.aspx) with the top producers in 2009/2010 being the United States (41% of world production) and China (19% of world production) (http://www.dailyfutures.com/grains/).

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.

2.2.2 Pseudomonas fluorescens

The source of the *hppdPF W336* gene is the bacterial species *Pseudomonas fluorescens* strain A32. Members of this species are gram negative, rod-shaped bacteria that are non-pathogenic saprophytes. They are ubiquitous inhabitants of soil and water and have a commensal[2] association with plants (Paulsen *et al*., 2005). They are widely distributed among fresh foods, especially vegetables, meats, poultry, and seafood products (Jay *et al*., 2005) and in food-related environments such as dairies and poultry farms (Sillankorva *et al*., 2008). They can be associated with food spoilage particularly where they have the opportunity to form biofilms.

2.2.3 Other organisms

Genetic elements from several other organisms have been used in the genetic modification of soybean FG72 (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 FG72.

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:

Criel, I. (2009). Description of vector pSF10. ID# M-201720-03-1, Bayer CropScience (unpublished).
Verhaeghe, S. (2009a). Detailed insert characterization of *Glycine max* transformation event FG72 by Southern blot analysis. ID# M-358354-02-1, Bayer CropScience (unpublished).
Verhaeghe, S. (2009b). Full DNA sequence of event insert and integration site of *Glycine max* transformation event FG72. ID# M-356895-02-1, Bayer CropScience (unpublished).
Verhaeghe, S. (2009c). Confirmation of the absence of vector backbone sequences in *Glycine max* transformation event FG72. ID# M-359801-02-1, Bayer CropScience (unpublished).
Verhaeghe, S. (2009d). Bioinformatics analysis of *Glycine max* transformation event FG72. ID# M-356899-02-1, Bayer CropScience (unpublished).
Verhaeghe, S. (2009e). Structural stability analysis of *Glycine max* event FG72 in different generations, in different backgrounds and when grown in different environments. ID# M-358391-01-1, Bayer CropScience (unpublished).

[2] Commensalism is a symbiotic relationship in which one organism derives benefit and the other is neither benefitted not harmed.

3.1 Method used in the genetic modification

Soybean cultivar 'Jack' was transformed with a linear DNA fragment derived from a *SalI* digestion of plasmid pSF10 (Figure 1) using a biolistic method (Klein *et al*., 1987).

In summary, cells from an embryogenic suspension culture were bombarded with gold particles coated with the purified fragment. The cells were placed in a medium containing isoxaflutole to select putative transformants. Following a round of multiplication cycles in the presence of isoxaflutole, transformed cells were regenerated and the resulting plants (T_0) were transferred to the glasshouse. Glyphosate and isoxaflutole were then used as selection agents and surviving plants were allowed to flower, self-pollinate and set seed (T_1). During further selfing rounds, characterisation and selection were continued.

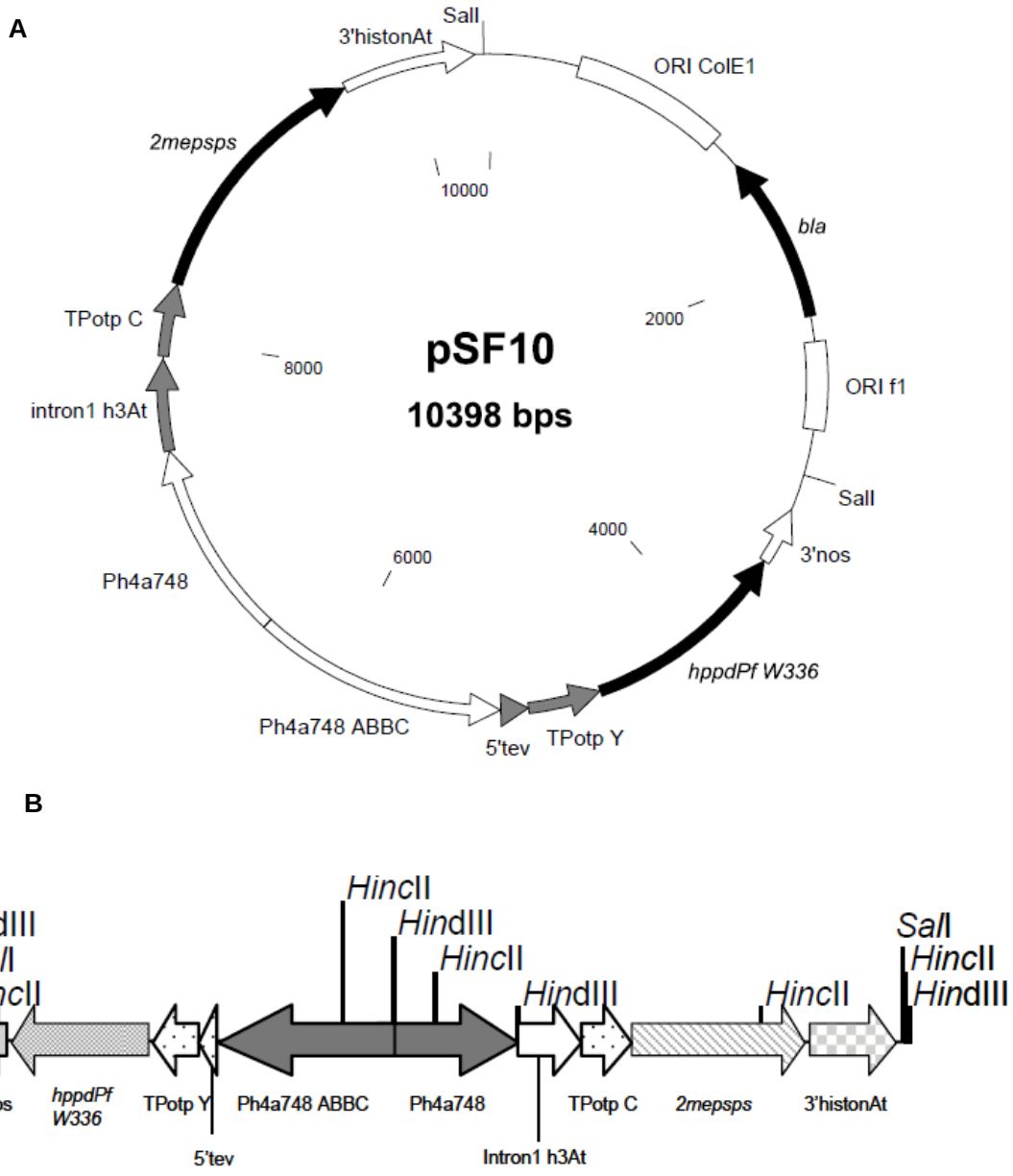

Figure 1: Genes and regulatory elements contained in plasmid pSF10 (A) and the linear fragment (B) derived from it.

3.2 Function and regulation of introduced genes

Information on the genetic elements in the introduced DNA fragment is summarised in Table 1.

Table 1: Description of the genetic elements contained in the introduced DNA fragment

Genetic element	bp location on pSF10	Size (bp)	Source	Orient.	Description & Function	References
3'*nos*	3262 - 3553	291	*Agrobacterium tumefaciens*	Anti-clockwise	• Sequence including the 3'UTR of the nopaline synthase gene from the T-DNA of pTiT37. • Terminates *hppdPf W336* gene expression and directs polyadenylation.	Depicker *et al*. (1982)
hppdPf W336	3554 - 4630	1076	*Pseudomonas fluorescens* strain A32	Anti-clockwise	• Coding sequence of 4-hydroxyphenylpyruvate dioxygenase, modified by the replacement of glycine (336) with tryptophan.	Boudec *et al.* (2001)
TPotp Y	4631 - 5002	371	*Zea mays* and *Helianthus annuus*	Anti-clockwise	• Coding sequences of an optimized transit peptide derivative of the RuBisCo small subunit genes from both species • Targets the HPPDPf W336 protein to the plastids	Lebrun *et al*. (1996)
5'*tev*	5003 - 5143	140	Tobacco etch virus	Anti-clockwise	• Leader sequence • Enhances translation of *hppdPf W336* gene	Carrington & Freed (1990)
Ph4a748 ABBC	5144 - 6433	1289	*Arabidopsis thaliana*	Anti-clockwise	• Sequence including the promoter region of the histone *H4* gene containing an internal duplication • Drives high level constitutive expression of *hppdPf W336* gene	Chaboute *et al*. (1987)
Ph4a748	6434 - 7448	1014	*Arabidopsis thaliana*	Clockwise	• Sequence including the promoter region of the histone *H4* gene Drives high level constitutive expression of *2mepsps* gene	Chaboute *et al*. (1987)
intron *h3At*	7449 - 7929	480	*Arabidopsis thaliana*	Clockwise	• First intron of gene II of the histone H3.III variant • Enhances expression of *2mepsps* gene	Chaboute *et al*. (1987)
TPotp C	7930 - 8301	371	*Zea mays* and *Helianthus annuus*		• Coding sequences of an optimized transit peptide derivative of the RuBisCo small subunit genes from both species • Targets the 2mEPSPS protein to the plastids	Lebrun *et al*. (1996)
2mepsps	8302 - 9639	1337	*Zea mays*	Clockwise	• Coding sequence of the double mutant 5-enolpyruvylshikimate-3-phosphate synthase gene	Lebrun et al. (1997)
3'*histonAt*	9640 - 10326	686	*Arabidopsis thaliana*	Clockwise	• Sequence including the 3' UTR of the histone H4 gene • Terminates *2mepsps* gene expression	Chaboute *et al*. (1987)

3.2.1 hppdPf W336 *expression cassette*

The gene encoding HPPD occurs in all five kingdoms – bacteria, fungi, algae, plants and animals including humans (Wada *et al*., 1975; Roche *et al*., 1982; Awata *et al*., 1994; Hamer *et al*., 2001; Galvez-Valdivieso *et al*., 2010). The protein plays a central role in aromatic amino acid catabolism (phenylalanine and tyrosine) in mammals and plants, and in plastidic quinione synthesis in plants (refer to Section 4.2.1).

The *hppd* (*hppdPf*) gene for the genetic modification described for soybean FG72 was initially isolated and cloned from the bacterium *Pseudomonas fluorescens* A32. The substitution of glycine (G) with tryptophan (W) at position 336 in the expressed HPPDPf protein resulted in the *hppdPf W336* gene (Boudec *et al.*, 2001). This substitution, which was achieved by random mutagenesis, was found to result in improved tolerance to isoxaflutole in plants carrying the substitution compared with those expressing the wild type *hppdPf* gene.

The *hppdPf W336* coding region in the linear transformation fragment is 1,076 bp in length and is driven by the constitutive, duplicated histone H4 promoter from the plant *Arabidopsis thaliana* followed by the enhancer sequence of the Tobacco etch virus. An optimised transit peptide derived from elements from the ribulose bisphosphate carboxylase (RuBisCo) subunits of both corn (*Zea mays*) and sunflower (*Helianthus annuus*) targets the HPPDPf W336 protein to the chloroplasts where endogenous HPPD is involved in the catabolism of the aromatic amino acids phenylalanine and tyrosine. The *hppdPf W336* coding region is terminated by a sequence from the 3' end of the *nopaline synthase (nos)* gene from *Agrobacterium tumefaciens.*

3.2.2 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 kDa protein with normal enzyme function and reduced affinity for the herbicide glyphosate.

The Ph4a748At promoter and h3At intron are regulatory elements used to control expression of the *2mepsps* gene and are derived from the histone H4 gene of *A. thaliana*. The use of these elements 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. The 3'histonAt terminator from *A. thaliana* corresponds to the polyadenylation signal which is essential to end transcription of the introduced gene.

3.3 Breeding of soybean plants containing transformation event FG72

A breeding programme was undertaken for the purposes of:

- obtaining generations suitable for analysing the characteristics of soybean line FG72
- ensuring that the FG72 event is incorporated into elite breeding line(s) for commercialisation of isoxaflutole- and glyphosate-tolerant soybean.

The breeding pedigree for the various generations is given in Figure 2.

Following selection of T_0 plants (see Section 3.1) a series of selfing and seed bulking together with selection for tolerance to glyphosate and isoxaflutole proceeded up to generation T_9. At the T_6 generation, plants were crossed with a number of elite lines to produce an F1 generation which was then selfed to produce further F generations.

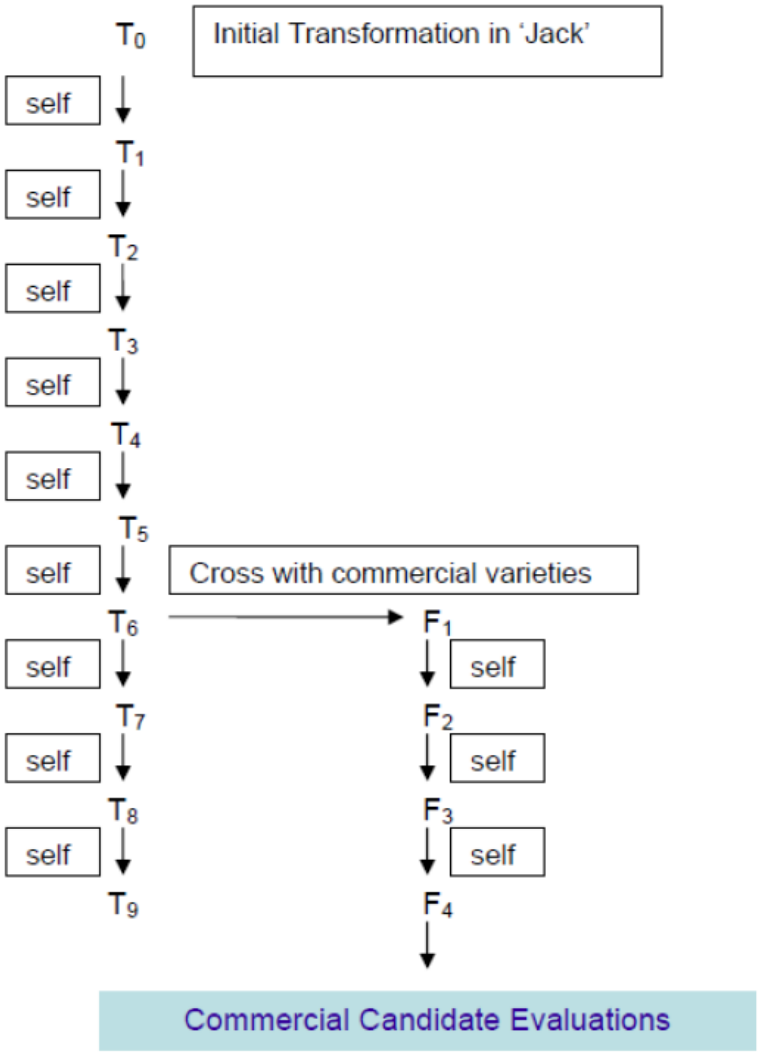

Figure 2: Breeding strategy for plants containing event FG72

3.4 Characterisation of the genes in the plant

A range of analyses were undertaken in order to characterise the genetic modification in soybean line FG72. These analyses focussed on the nature of the insertion of the introduced genetic elements and whether any unintended genetic re-arrangements may have occurred as a consequence of the transformation procedure.

3.4.1 Transgene copy number and insertion integrity

Total genomic DNA from leaf tissue of certified soybean FG72 (T_7 generation) and 'Jack' (negative control) seedlings was used for Southern blot analyses. A positive control (DNA from 'Jack' spiked with HindIII-digested DNA from the pSF10 plasmid) was also included in the Southern blot analyses. The DNA from soybean FG72 seedlings was digested with one of 10 restriction enzymes, while DNA from 'Jack' was digested with HindIII. The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with eight different radiolabelled probes that represented various sections, including the complete fragment, of the linear fragment used for transformation. The lengths of all

hybridisation fragments were estimated using GeneTools software. The negative control showed no hybridisation with any of the probes.

The Southern blot analyses indicated that there is a single insert in event FG72 comprising two partial 3'histoneAt sequences in a head to head orientation, followed by 2 complete copies of the linear fragment used for transformation, arranged in a head to tail orientation (refer to Figure 3). In addition, a genomic fragment has translocated to a new position and is flanked at the 3' end by 158 base pairs of the Ph4a748 promoter sequence.

3.4.2 Full DNA sequence of insert

Genomic DNA was obtained from leaf tissue of T_7 generation soybean FG72 plants and a negative control (cultivar 'Jack'). These samples were used to characterise the DNA sequence in the transgene insertion and its flanking border regions. Seven PCR fragments (FG72-TR1 to FG72-TR7 in Figure 3) spanning the inserted sequences in event FG72 were amplified to determine the sequence of the FG72 locus; two PCR fragments (FG72-TL1 and FG72-TL2) were amplified to determine the sequence of the newly-created junctions due to translocation of genomic sequences; four PCR fragments (JACK) were amplified to determine the sequence of the 'Jack' regions affected by the inserted genetic material. DNA sequencing analysis was done using commercially available software (Sequencher®) and alignment of the sequences was managed using Clone Manager software.

The sequences at the insertion site in event FG72 were shown to comprise:

- 1,452 bp of 5' flanking sequences of which at least the last 1,166 are verified as identical to sequences in the 'Jack' integration site
- 15,163 bp of sequences derived from the linear fragment used for transformation
- 24 bp of filler DNA that does not correspond to either plasmid or 'Jack' DNA
- 1,168 bp of 3' flanking sequences verified as identical to sequences in the 'Jack' integration site

At the region in which DNA from 'Jack' has translocated, the sequences (with the exception of the 157 bp corresponding to Ph4A748, a 25 bp deletion from the JACK-WT2 fragment, and a 2 bp deletion from the JACK-WT3 fragment) all derive from and are identical to 'Jack' genomic material. The (re) arrangement is summarised in Figure 3 and Figure 4. The translocation of genomic sequences resulted in the generation of two additional junctions (fragments FG72-TL1 and FG72-TL2), of which the 3' junction is joined by 158 bp of Ph4a748 promoter sequences (fragment FG72-TL2).

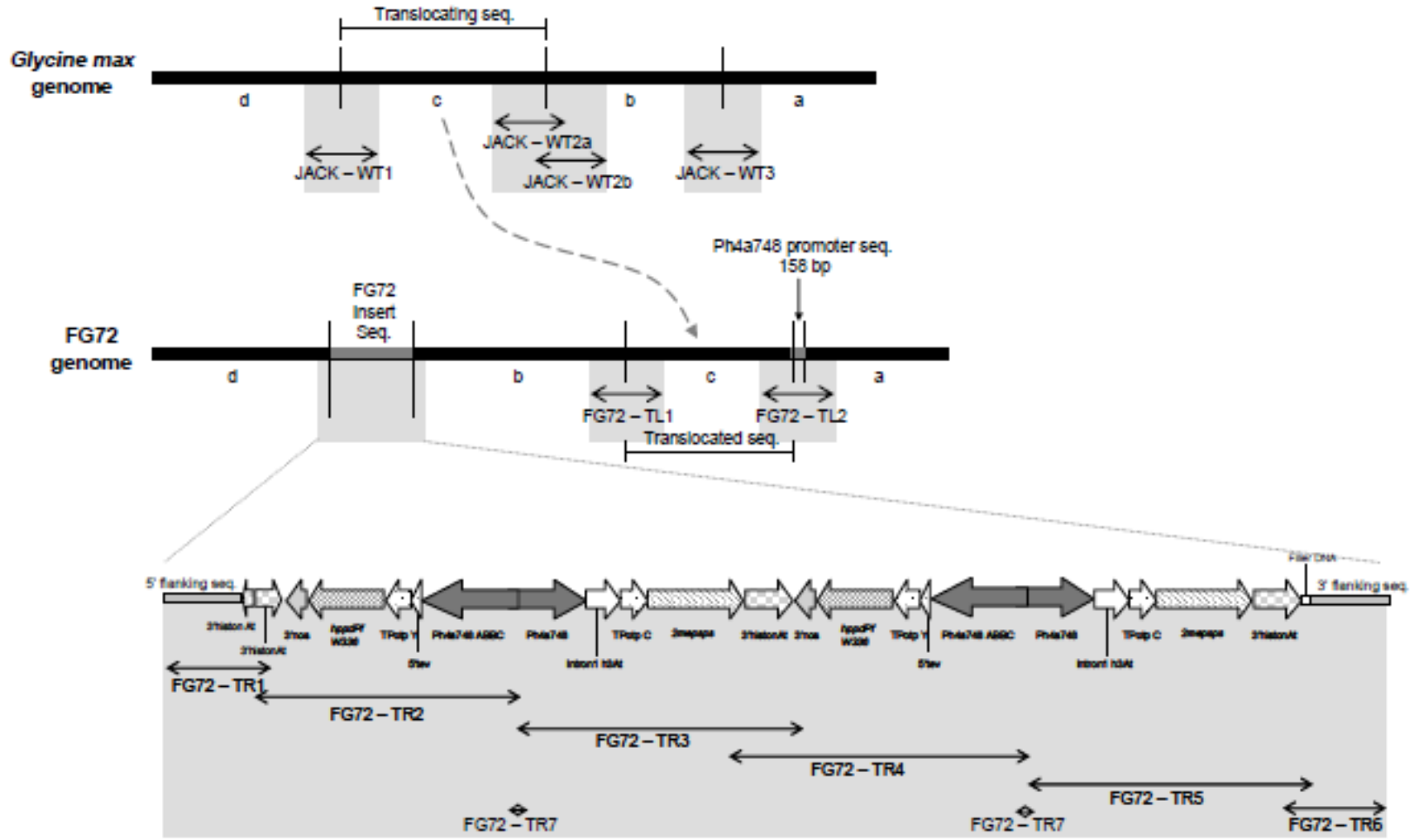

Figure 3: Inserted transgenic sequences in event FG72

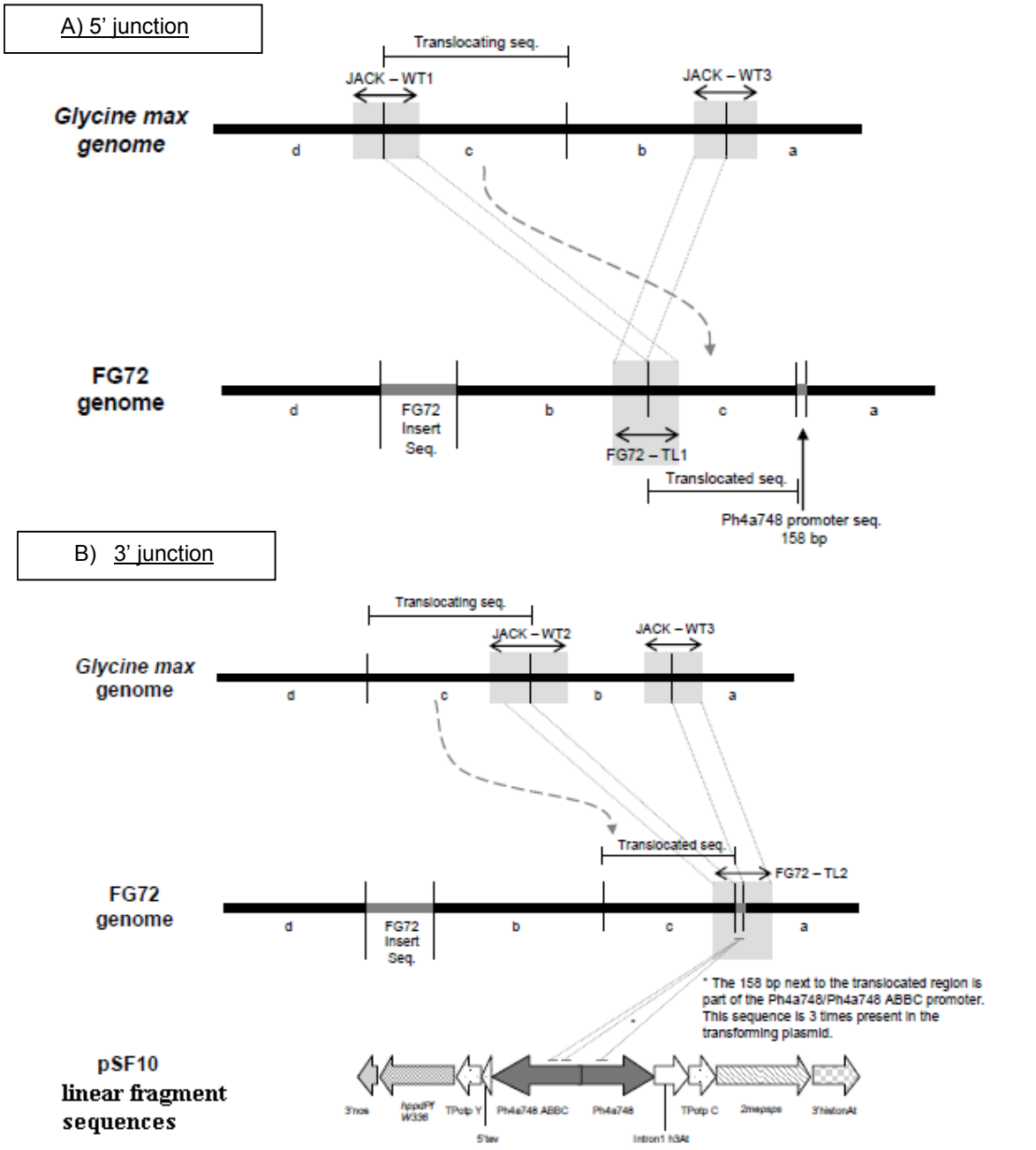

Figure 4: Representation of the arrangement of translocated 'Jack' genetic material and fragment copies in the FG72 genome.

3.4.3 Plasmid backbone DNA analysis

Southern blot analysis was done to determine whether any plasmid pSF10 backbone had been included along with the inserted DNA in soybean FG72. Leaf tissue from confirmed seedlings of FG72 (generation T_7) and a negative control ('Jack') was used for this analysis. A positive control (DNA from 'Jack' spiked with vector DNA from plasmid pSF10 digested with HindIII) was also included in the Southern blot analyses. The DNA from T_7 was digested with HindIII and HincII restriction enzymes while control DNA was digested with HindIII enzyme. The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with two overlapping radiolabelled probes covering most of the

vector backbone sequence of pSF10, and an *hppdPf W336* probe covering the linear fragment used for transformation.

When hybridising with the vector backbone probes, no hybridisation fragments were observed in either the 'Jack' samples or FG72 samples, while the expected fragments were shown in the spiked 'Jack' samples. The expected Southern blot profile was obtained in the FG72 samples after hybridization with the *hppdPf W336* probe therefore demonstrating that an adequate amount of a sufficient quality of FG72 genomic DNA was loaded on the gels to be able to detect the vector backbone sequences in the FG72 event, had they been present.

PCR reactions were performed using FG72 genomic DNA, WT JACK genomic DNA and pSF10 plasmid DNA as template DNA, to assess the integration in FG72 of the vector backbone fragments that were not covered by the two vector backbone probes. No PCR fragments of correct size were obtained in the FG72 genomic DNA samples, while the expected fragments were obtained in pSF10 plasmid DNA samples.

Taken together, the Southern blot and PCR analyses demonstrate the absence of vector backbone sequences in the inserted DNA.

3.4.4 Open reading frame (ORF) analysis

A bioinformatics analysis was performed to determine whether any new ORFs that may lead to the production of proteins had been created in the junctions either between the FG72 inserted linear fragment insert and the host genomic DNA, or in the areas bordering the translocated sequence. A total of nine junctions were identified (see Figure 5)

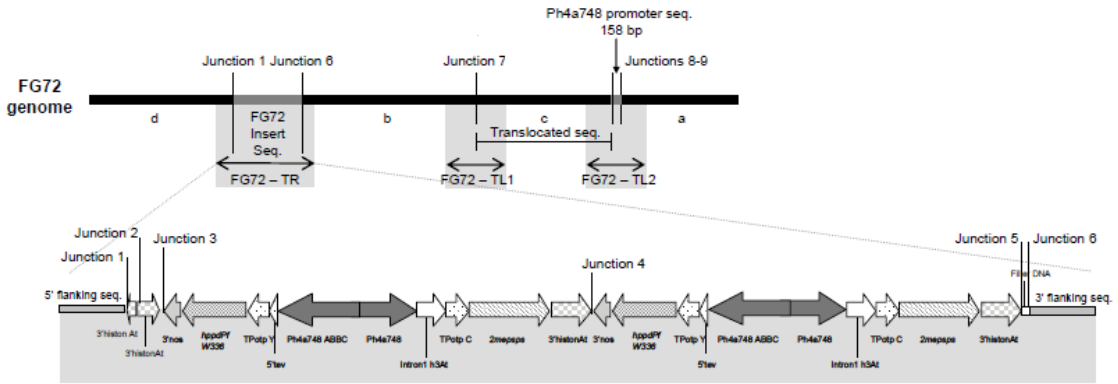

Figure 5: The nine junctions identified in event FG72

The *in silico* analysis was done through the search programme, GetORF from the European Molecular Biology Open Software Suite. In a first analysis, an ORF was defined as a region between two translation stop codons (TAA, TAG, TGA) with a minimum size coding for three amino acids. Such ORFs may have a coding potential but the absence of a start codon means this potential cannot be realised. In a second analysis, an ORF was defined as a region between a start (ATG) and standard stop (TAA, TAG, TGA) codon with a minimum size coding for three amino acids. These ORFs (a sub-set of ORFs bounded by stop codons) theoretically would have a coding potential that could be realised because of the presence of a start codon.

Another programme, FGENESH (Softberry Inc. version 2.4), which predicts introns and exons by statistical sequence analysis and polyA signals by homology search with known plant consensus sequences, was used to identify any new, potentially expressed genes

spanning any of the junction regions.

In addition to the above, a homology search was done using the pattern-finding tool TSSP (Softberry Inc.) to identify any genetic elements in the junctions that play a role in the regulation of gene expression.

A comparison was also made of the sequence surrounding the first ATG codon of any putative ORFs with a consensus sequence for the ribosome binding site (RBS) of plant genomes (Joshi *et al*., 1997) in order to determine if it may be a putative start of translation.

FGENESH did not identify any newly created putative genes in the junction regions; it did predict a gene in the 5' end sequence of the translocated region which corresponds with the putative cysteine protease described in Section 3.4.5 and is an endogenous gene in the 'Jack' genome.

Across the nine junctions, the GetORF analysis identified 46 newly created ORFs defined between two stop codons but of these only eight included a start codon. TSSP identified seven putative promoters associated with the eight ORFs. In most cases the promoters were either not in the correct orientation to initiate transcription or were too far upstream. Information on the ATG context of these ORFs indicate that most of the essential nucleotides were absent suggesting any potential promoters would be biologically inactive.

Taken together, these analyses suggest there is low likelihood of translation of any of the identified unexpected ORFs (see also relevant discussion in Section 4.1).

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

An analysis of the pre-insertion locus was done in order to ascertain whether regulatory sequences, endogenous soybean genes and/or ORFs may have been disrupted by the insertion of the transgenic sequences or the translocation of genomic sequences. Five insertion sites were identified (see Figure 6).

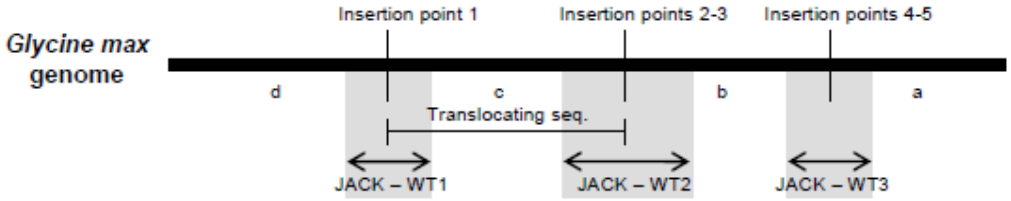

Figure 6: Insertion sites in the 'Jack' genome

A BLASTx (Basic Local Alignment Search Tool)[3] search (Altschul *et al*., 1997) was done to search for similarities between the pre-insertion locus and known soybean genomic nucleotide sequences in the DAD (DNA Data Bank of Japan Aminoacid Database), UniProt (Universal Protein Resource) and GenPept databases (for further information on bioinformatic analysis see Section 4.5.2.). Additionally, GetORF and TSSP searches were carried out.
The BLASTx homology search identified homology of the 5' end sequence of the translocated region at Insertion point 1 with part of a putative cysteine protease. The interruption of the function of the gene was not predicted. No other genes interrupted as a result of the transformation procedure were identified.

[3] BLASTx takes a nucleotide sequence, translates it, and then queries it against a protein sequence (Gish and States, 1993)

GetORF identified 5 interrupted ORFs, defined between a start and a stop codon, at Insertion points 2 -3. The TSSP search identified three putative promoter regions (two in Insertion point 1 and one in Insertion points 4-5), that are interrupted by the transformation event.

Taken together, the analyses indicate that no known genes were interrupted by the insertion of transgenic DNA in the pre-insertion locus.

3.5 Stability of the genetic changes in soybean line FG72

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).

3.5.1 Genetic stability

The genetic stability of event FG72 was evaluated in leaf tissue from individual plants of:

- three different generations: T_2, T_7 and T_9
- two different genetic backgrounds: 21 x F_4 plants obtained by crossing line FG72 with conventional lines 3068115-48 and 3066617-48 followed by 3 rounds of selfing.
- four different harvest locations in the U.S.: 22 x T_9 plants from each of Adel (Iowa), Osborn (Missouri), Fithian (Illinois), and Sharpsville (Indiana).

For each FG72 plant the identity and zygosity was confirmed by zygosity-PCR (zPCR)[4].

The non-GM cultivar 'Jack', grown at all locations, was used as a negative control, and cultivar 'Jack' spiked with DNA from plasmid pSF10 was used as a positive control. Genomic DNA isolated from leaf tissue and the plasmid DNA was digested with HindIII restriction enzyme. The resulting DNA fragments were separated and transferred to a membrane for hybridisation with two radiolabelled probes, one for the entire linear fragment in pSF10 and one for Ph4a748B. The fragment sizes were determined using GeneTools software.

There was no hybridisation in the 'Jack' negative control. The Southern blot analysis confirmed the presence of the expected hybridisation fragments in all tested transgenic DNA samples and in the positive control and therefore confirmed the genetic stability of the modification in FG72 over different generations, in different genetic backgrounds and across different environments.

3.5.2 Phenotypic stability

As outlined in Section 3.3, successive rounds of selfing and selection were carried out and at T_6, plants were crossed with conventional soybean breeding lines and then selfed in order to select commercial candidate lines. At the F_2 generation, leaf samples were collected from 901 plants and analysed using zPCR. The results (presented in Table 2) showed that there was no significant difference ($p = 0.05$; X^2 less than 3.84 with 1 degree of freedom) between

[4] zPCR is a polymerase chain reaction-based analysis to determine whether an individual plant carries a particular GM event at a certain insertion locus at a homozygous, hemizygous or azygous state. In this instance, the technique used three oligonucleotide primers.

the expected and observed ratios and therefore, that the FG72 phenotype is stably inherited under Mendelian principles.

Table 2: Segregation data for F_2 progeny

Zygosity of the FG72 locus	Total plants	Observed ratio	Expected ratio
Azygous (nul/nul)	212	0.24	0.25
Heterozygous (FG72/nul)	471	0.52	0.5
Homozygous (FG72/FG72)	218	0.25	0.25
	901		$X^2 = 0.172$

3.6 Antibiotic resistance marker genes

No antibiotic resistance marker genes are present in soybean line FG72. Plasmid backbone analysis shows that no plasmid backbone has been integrated into the soybean genome during transformation, i.e. the *bla* gene, which was used as a bacterial selectable marker gene, is not present in soybean FG72.

3.7 Conclusion

Soybean line FG72 contains two novel gene cassettes. One contains the *hppdPf W36* gene conferring tolerance to isoxazole herbicides such as isoxaflutole and the other contains the *2mepsps* gene conferring tolerance to glyphosate herbicides. There are no antibiotic resistance marker genes present in soybean line FG72.

Comprehensive molecular analyses of soybean line FG72 indicate there is a single insertion site comprising two partial 3'histoneAt sequences in a head to head orientation, followed by two complete copies of the linear fragment used for transformation, arranged in a head to tail orientation. In addition, a genomic fragment has translocated to a new position and is flanked at the 3' end by 158 base pairs of Ph4a748 promoter sequence.

The introduced genetic elements are stably inherited from one generation to the next. Plasmid backbone analysis shows that no plasmid backbone has been incorporated into the transgenic locus. A number of unexpected ORFs are present at the junctions associated with the insertion site but are most likely biologically inactive as they lack the necessary elements for expression of a protein. No known endogenous genes have been interrupted by insertion of the new genetic material.

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 ORFs during the transformation process (see Section 3.4.4).
- Those that were expected to be directly produced as a result of the translation of the introduced genes. Soybean FG72 expresses two novel proteins, an HPPDPf W36 protein and an EPSPS protein. A number of different analyses were done to determine the identity, physiochemical properties, *in planta* expression, bioactivity and potential toxicity and allergenicity of the two proteins.

Because the expression of proteins *in planta* is usually too low to allow purification of sufficient quantities for safety assessment studies, a bacterial expression system was used to generate large quantities of both proteins. The equivalence of the bacterial-produced proteins to the plant-produced proteins was determined as part of the protein characterisation.

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

Study submitted:

Capt, A. (2009). *In silico* analysis of putative reading frame (ORF) sequences for identifying potential homologies to known toxins and allergens. ID# M-357579-01-1, Bayer CropScience (unpublished).

ORF analysis identified the formation of eight unexpected ORFs between start and stop codons in event FG72 (Section 3.4.4). Of these, one ORF coded for less than eight amino acids and was therefore excluded from further bioinformatic analysis because it contained less than the minimum number of contiguous amino acids.

To evaluate the allergenicity of proteins that might potentially be produced from translation of the seven ORFs, an epitope search was carried out to identify any short sequences of amino acids that might represent an isolated shared allergenic epitope. This search compared the sequences with known allergens in the AllergenOnline database using the FindPatterns algorithm. The criterion indicating allergenicity was a 100% identity, on a window of eight amino acids, with an allergenic protein. No significant similarities were identified.

The sequences corresponding to the seven ORFs were also compared with protein sequences present in a number of large public reference databases: Uniprot_Swissprot, Uniprot_TrEMBL, PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept. The similarity searches used the BLASTP (Basic Local Alignment Search Tool Protein) algorithm (refer to Section 4.5.2.1 for an explanation). No significant similarities of the seven ORFs to any sequences (including those of known toxins) in the databases were found.

It is concluded that, in the unlikely event that any of the eight ORFs were expressed, the encoded proteins are unlikley to be allergenic or toxic to humans.

4.2 Function and phenotypic effects of the HPPDPf W36 and 2mEPSPS proteins

4.2.1 HPPDPf W336 protein

As indicated in Section 3.2.1, p-hydroxyphenylpyruvate dioxygenase (HPPD) occurs in all five kingdoms – bacteria, fungi, plants and animals including humans. In plants, it is involved

in the catabolism of the aromatic amino acids phenylalanine and tyrosine (see Figure 7) and leads to the formation of vitamin E (an antioxidant) and plastoquinones (which are elements of the electron transfer chain of photosynthesis). It is a mononuclear, non-heme, iron-containing enzyme which is a member of the family of 2-oxoacid dependent dioxygenases (Ryle and Hausinger, 2002).

HPPD is a target for bleaching herbicides such as isoxaflutole (Pallett *et al*., 1998; Viviani *et al*., 1998). The bioactive constituent of isoxaflutole, namely diketonitrile (DKN) is formed in the plant by enzymically opening the isoxazole ring in isoxaflutole. DKN is then further metabolised to herbicidally inactive benzoates. DKN is able to mimic the α-keto acid group of hydroxyphenylpyruvate (HPP), the usual HPPD substrate and compete for binding to HPPD. Competitive, though partially irreversible, binding to the active site of the HPPD then inhibits the activity of the HPPD. This inhibition leads to tyrosine accumulation, plastoquinone and vitamin E depletion and an accumulation of phytoene, a noncoloured precursor of carotenoids. The inhibition of carotenoid biosynthesis results in the bleaching and subsequent death of treated plants.

HPPDPf W336 is an enzyme having a lower specificity for DKN (higher k_m and higher k_{on}) than the wild type HPPD. This reduced specificity confers a higher level of tolerance against isoxaflutole. HPPDPf W336 was generated through mutagenesis (Matringe *et al*., 2005). HPPDPf W336 possesses greater than 99.5% homology to the native HPPD protein from *P. fluorescens*. The *Pseudomonas* HPPD has 53% conserved amino acids and 27% identical amino acids with human liver HPPD (Rüetschi *et al*., 1993).

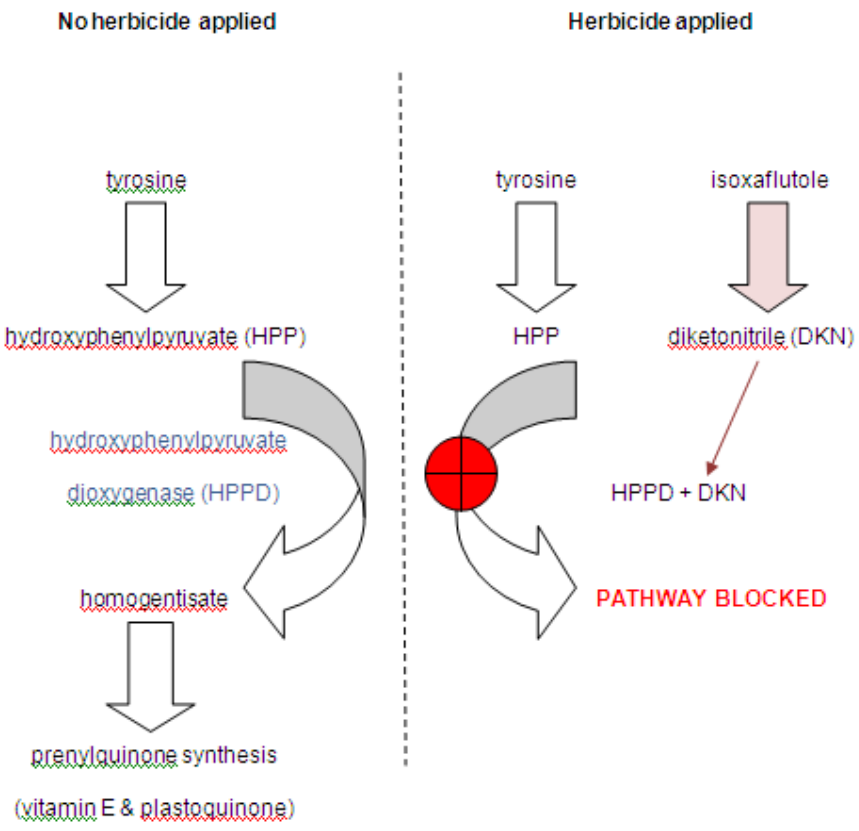

Figure 7: Schematic diagram of the action of HPPD in the presence and absence of isoxaflutole

4.2.2 2mEPSPS 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 FG72 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 'Jack', used as the parent for the genetic modification described in this application, contains a wild type EPSPS protein.

4.3 Protein expression analysis

During the course of ongoing method development, a proprietary extraction buffer was developed for the 2mEPSPS protein. In earlier studies done in 2009, the plant material was extracted with a buffer known as PBST, while in later studies completed in 2011, a buffer known as SDI was used. A comparison of the two buffers indicated that both gave rise to similar results for leaf, stem and root samples but that the SDI buffer resulted in significantly enhanced detection of 2mEPSPS in seed and processed products derived from the seed.

4.3.1 Novel protein expression in plant tissues

Study submitted:

Habex, V.; Debaveye, J. (2009). Expression analysis of HPPD W336 and 2mEPSPS in transgenic soybean event FG72. ID# M-358352-02-1, Bayer CropScience (unpublished).

HPPDPf W336 and 2mEPSPS are expected to be expressed in all plant tissues since the *hppdPf W336* and *2mepsps* genes are driven by constitutive promoters (refer to Table 1). Plants of FG72 (generation T_7) and 'Jack' were grown in a single greenhouse trial in Belgium.

HPPDPf W336 and 2mEPSPS protein levels were determined by enzyme linked immunosorbent assay (ELISA) using commercial ELISA kits specific for each protein. The plant parts (at particular growth stages) sampled from 'Jack' control plants and glyphosate-sprayed FG72 plants are given in Table 3. Ten samples of each component were measured; figures in Table 3 are averages. These results in Table 3 do not include levels of 2mEPSPS in the seed since the PBST extraction buffer had been used. The results for the level of 2mEPSPS in SDI-extracted seed were determined in separate experiments and have been included in Sections 4.3.2. and 4.3.3

For all FG72 plants, identity and zygosity was confirmed by PCR. All ‘Jack’ plant material was analysed using a discriminating PCR to test for adventitious presence of various regulatory elements and genes.

No HPPDPf W336 or 2mEPSPS proteins were detected in samples taken from ‘Jack’ plants. For soybean FG72 plants, HPPDPf W336 and 2mEPSPS proteins were detected in all parts (Table 3). HPPDPf W336 was lowest in the seed (approximately 1.5 µg/g dry weight) and highest in younger leaves (approximately 38 µg/g dry weight). In vegetative plant parts, 2mEPSPS protein concentrations were much higher than those for HPPDPf W336. The leaves contained the highest levels (older leaves contained approximately 660 µg/g dry weight) while roots contained the lowest levels (approximately 40 µg/g dry weight).

Table 3: Novel protein content in FG72 soybean parts at different growth stages. NOTE: The level of 2mEPSPS in seed has not been included because a different extraction buffer needed to be used for seed.

Matrix	Growth stage	HPPD W336 protein content				2mEPSPS protein content			
		µg/g fresh weight		µg/g dry weight		µg/g fresh weight		µg/g dry weight	
		Average ± SD	Range	Average ± SD	Range	Average ± SD	Range	Average ± SD	Range
Leaf	V4	6.10 ± 2.78	2.65 – 10.4	38.4 ± 17.5	16.7 – 65.7	90.4 ± 26.1	44.9 – 152	569 ± 164	283 – 958
	V6	6.48 ± 4.08	2.31 – 17.4	35.8 ± 22.5	12.8 – 96.0	79.1 ± 29.6	39.2 – 136	437 ± 163	216 – 753
	V8	4.69 ± 1.87	2.00 – 8.91	27.2 ± 10.9	11.6 – 51.8	115 ± 38.2	60.5 – 203	668 ± 222	351 – 1180
Stem	V4	1.48 ± 0.42	0.74 – 2.20	16.6 ± 4.65	8.29 – 24.6	18.8 ± 6.16	6.08 – 31.3	211 ± 68.9	68.0 – 350
	V8	0.69 ± 0.35	0.29 – 1.49	6.04 ± 3.10	2.49 – 13.0	13.4 ± 2.62	8.71 – 17.3	117 ± 22.9	76.1 – 151
Root	V4	0.87 ± 0.35	0.45 – 1.66	5.81 ± 2.30	2.98 – 11.0	4.89 ± 1.99	1.63 – 8.21	32.5 ± 13.2	10.8 – 54.5
	V8	0.84 ± 0.50	0.20 – 1.64	6.42 ± 3.82	1.51 – 12.4	5.75 ± 2.31	2.62 – 10.7	43.7 ± 17.6	19.9 – 81.2
Seed	NA	1.27 ± 0.42	0.71 – 2.68	1.41 ± 0.47	0.79 – 2.96				

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

4.3.2 Novel protein content in raw agricultural commodity (RAC)

Studies submitted:

Kowite, W.J. (2009). Production of raw agricultural commodities (grain) of transgenic event FG72 soybenas from multiple field trials, USA, 2008. ID# M-353435-02-1, Bayer CropScience (unpublished).

Poe, M.R. (2011). Analyses of the raw agricultural commodity of soybean event FG72 for HPPD W336 and 2mEPSPS proteins. USA 2009. ID# M-357840-02-1, Bayer CropScience (unpublished).

Currier, T.C.; Harbin, A.M. (2011). Analyses of the raw agricultural commodity of soybean event FG72 2mEPSPS protein.USA 2010. ID# M-403249-01-1, Bayer CropScience (unpublished).

Seed samples of sprayed (glyphosate + isoxaflutole) and unsprayed FG72 plants from generation T_8 were obtained from field trials in 10 locations in the U.S. Details of the sites and growing conditions are given in Section 6.2 as the same plants were used for the RAC and compositional analyses. The identity and purity of the FG72 seed was confirmed by PCR analysis. ELISA analysis in conjunction with SoftMax Pro™ software was used to derive the concentrations of the two proteins.

The experimental work, although initially using 2mEPSPS data from seed extracted with PBST buffer, was amended to reanalyze samples from the same batch using the optimized SDI buffer.

The results for the two proteins are given in Table 4.

Table 4: Levels of novel proteins in FG72 RAC

Protein	µg/g fresh weight (mean +/-SD)		µg/g dry weight (mean)		% crude protein	
	unsprayed	sprayed	unsprayed	sprayed	unsprayed	sprayed
HPPDPf W336	0.846±0.199	0.802±0.223	0.936	0.887	0.00024	0.00023
2mEPSPS	130±22	140±33	150	150	0.039	0.041

These results indicate that the level of 2mEPSPS in the seed is approximately 150 µg/g dry weight.

4.3.3 Novel protein content in processed fractions

Studies submitted:

Kowite, W.J. (2009). Production of processed commodities from transgenic event FG72 soybeans and the non-transgenic counterpart (2009). ID# M- 357014-01-1, Bayer CropScience (unpublished).

Robinson, T.D. (2009). Analyses of processed commodities from transgenic event FG72 soybeans for HPPD W336 and 2mEPSPS proteins. USA 2009. ID# M-358233-01-1, Bayer CropScience (unpublished).

Massengill, J.A. (2011). Analyses of processed commodities from transgenic event FG72 soybeans, USA 2011. ID# 11-BM99L222, Bayer CropScience (unpublished).

The Applicant provided data for the levels of HPPDPf W336 and 2mEPSPS in various products derived from seeds from line FG72 which were collected from plants grown at a field site in Adel (Iowa) under typical soybean production conditions. The FG72 plants were either unsprayed or were sprayed with glyphosate and isoxaflutole. The seed was harvested and processed and the novel protein levels were quantified using ELISA. The analyses (refer to Table 5) showed that the two proteins were not detected in meal, oil or lecithin samples.

The experimental work, although initially using 2mEPSPS data from product extracted with PBST buffer (Robinson, 2009 study), was amended to reanalyze samples from the same batches using the optimized SDI buffer (Massengill, 2011 study).

Table 5: Levels of novel proteins in processed fractions of soybean FG72

Protein	Processed fraction	µg/g fresh weight (mean +/-SD)		µg/g dry weight (mean)	
		unsprayed	sprayed	unsprayed	sprayed
HPPDPf W336	Whole soybean	0.75±0.03	0.91±0.05	0.83	1.01
	Hull	0.94±0.06	0.95±0.04	1.06	1.07
	Protein isolate	0.627±0.05	1.07±0.01	0.64	1.1
	Untoasted meal	ND*	ND	ND	ND
	Crude oil	ND	ND	ND	ND
	Crude lecithin	ND	ND	ND	ND
2mEPSPS	Whole soybean	112±4.64	111±1.95	127±5.24	125±2.19
	Hull	50.6±6.39	77.5±7.81	58.4±7.38	88.5±8.92
	Protein isolate	0.977±0.047	1.56±0.022	1.02±0.049	1.62±0.023
	Untoasted meal	0.041±0.006	0.048±0.005	0.046±0.006	0.054±0.005
	Toasted meal	ND	ND	ND	ND
	Crude oil	ND	ND	ND	ND
	Refined oil	ND	ND	ND	ND
	RBD oil	ND	ND	ND	ND
	Crude lecithin	ND	ND	ND	ND

*ND = not detectable

The results indicate that the HPPDPf W336 protein may be concentrated to a small degree in hulls and protein isolate, and is undetectable in other processed fractions. Levels of the 2mEPSPS protein are reduced in all fractions during processing, being undetectable in toasted meal, crude lecithin and all forms of oil.

4.4 Protein characterisation and equivalence

Studies submitted:

Habex, V.; (2009). Certificate of analysis for the HPPD W336 protein produced in *E. coli* batch n° LB020309. Report No. BBS09-001, Bayer CropScience (unpublished).

Bautsoens, N.; Hendrickx, K. (2006). Certificate of analysis for the 2mEPSPS protein produced in E. coli batch N° LEJ5837. Report No. BBS06-003, Bayer CropScience (unpublished).

Martone, A. (2009). Structural and functional equivalence of the 2mEPSPS protein produced in *Escherichia coli* to 2mEPSPS in FG72 soybean, *Glycine max,* USA, 2009. ID# M357733-01-1, Bayer CropScience (unpublished).

Martone, A. (2009). Structural and functional equivalence of the HPPD W336 protein produced in *Escherichia coli* to HPPD W336 in FG72 soybean, *Glycine max,* USA, 2009. ID# M357737-01-1, Bayer CropScience (unpublished).

Capt, A. (2008). 2mEPSPS protein epitope homology and N-glycosylation searches. ID# M-273851-02-1, Bayer CropScience (unpublished).

Capt, A. (2009). HPPD W336 protein amino acid sequence homology search with known allergens. ID# M-355637-01-1, Bayer CropScience (unpublished).

HPPDPf W336 comprises 358 amino acids and has a calculated molecular weight of 40 kDa. The protein encoded by the transformation vector is a fusion protein between HPPDPf W336 and the transit peptide TPotp Y (see Table 1). The theoretical molecular weight of the fusion protein is 48 kDa but it would be expected that the transit peptide would be cleaved from HPPDPf W336 during transfer to the chloroplast. As discussed in Section 4.2.1,

HPPDPf W336 shows less inhibition to isoxaflutole compared with the endogenous soybean HPPD protein.

Site directed mutagenesis of the wild type *epsps* gene from maize produced the double mutant enzyme 2mEPSPS which carries two amino acid changes. When fused to a chimaeric optimized chloroplast transit peptide, the 2mEPSPS enzyme is reported to generate optimal glyphosate tolerance in crops (Lebrun *et al.,* 1997). A methionine codon was added to the amino-terminal end of the mature 2mEPSPS protein sequence to restore the cleavage site of the transit peptide. With the addition of the methionine residue, the mutations are at positions 103 (Thr to Ile) and 107 (Pro to Ser) of the mature protein which comprises 445 amino acids and has a calculated molecular weight of 47 kDa

HPPDPf W336 and 2mEPSPS are not produced in sufficient quantity in soybean FG72 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. HPPDPf W336 and 2mEPSPS were therefore expressed in recombinant *Escherichia coli* and characterisation tests were done to confirm the identity and equivalence of these bacterially-produced proteins to those produced in soybean FG72.

4.4.1 Microbially expressed proteins

HPPDPf W336 and 2mEPSPS protein were obtained from *E. coli* bacterial expression systems using standard methods. In order to characterise the bacterially-produced proteins, three analyses were undertaken for both proteins:

- molecular weight (sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE))
- immunoreactivity (western blotting)
- activity assay

The purity and concentration of each protein were also determined. Several different studies of each protein were done over time as required by the need to generate protein for safety assessment tests. Results of only one study (as referenced in 'Studies submitted') of each protein are given here but all studies gave similar results.

The molecular weights of the proteins were calculated using a regression analysis derived from plotting migration of marker proteins in the SDS-PAGE gel against molecular weight. An apparent molecular weight of 40.7 kDa was obtained for the HPPDPf W336 protein and of 48 kDa for the 2mEPSPS protein. This is considered to be good experimental agreement with the calculated molecular weights of 40 kDa for HPPDPf W336 and 47 kDa for 2mEPSPS.

Immunoreactivity was tested by incubating blotted polyvinylidene fluoride membranes separately as follows:

- For HPPDPf W336: monoclonal mouse anti-HPPD Pf W336 followed by commercially available (Sigma) rabbit anti-mouse alkaline phosphatase secondary antibody.
- For 2mEPSPS: polyclonal rabbit anti-2mEPSPS followed by commercially available (Sigma) goat anti-rabbit alkaline phosphatase secondary antibody.

For both proteins there was staining of a single band at the expected molecular weight.

For both activity assays, a colorimetric method was used:

- HPPDPf W336 catalyses the transformation of HPP to homogentisate (refer to Figure 7). Activity can therefore be measured by determining the amount of HPP remaining in the assay mixture at the end of the incubation period after derivatisation with 2,4-dinitrophenylhydrazine. The activity of the bacterially-derived protein was confirmed.
- EPSPS catalyses the transfer of the enolpyruvyl moiety of phosphoenolpyruvate to the 5-OH hydroxyl group of shikimate 3-phosphate and releases inorganic phosphate. The inorganic phosphate is measured using a malachite green dye method. This method confirmed the activity of the bacterially-derived protein.

Taken together, the above three analyses confirmed the identity of the proteins obtained from the bacterial expression system as HPPDPf W336 and 2mEPSPS.

The concentration and purity of the two proteins was calculated to be the following:

- HPPDPf W336: 1.01 mg/mL; 96% purity
- 2mEPSPS: 0.93 mg/mL; 99% purity

4.4.2 Plant-produced proteins

4.4.2.1 Glycosylation analysis

Many eukaryotic proteins are glycoproteins that have been post-translationally modified by the addition of carbohydrate moieties (glycans) covalently linked to the polypeptide backbone. The non-virulent *E. coli* strains used in the laboratory for cloning and expression of novel proteins lack the necessary biochemical machinery for protein glycosylation (Wacker *et al*., 2002; Abu-Qarn *et al*., 2008).

N-glycosylated proteins are glycosylated on an asparagine residue and commonly contain an asparagine-X-serine/threonine sequence (N-X~(P)-[S/T), where X~(P) indicates any amino acid except proline (Orlando and Yang, 1998). Although rare, the sequence asparagine-X-Cysteine (N-X-C) can also be an N-glycosylation site (Miletich and Broze Jr., 1990). The consensus sequences searched were therefore of the following type:

N-X~(P)-[S,T] or N-X-C

The occurrence of these motifs does not, however, indicate that the protein will necessarily be glycosylated.

An *in silico* approach was used to search the sequences of the HPPDPf W336 and 2mEPSPS plant-produced proteins for the occurrence of the two motifs given above. No potential N-glycosylation site was identified on the HPPD W336 amino acid sequence.

The search found two potential sites in the 2mEPSPS protein on the basis of the N-X~(P)-[S,T] consensus sequence. However, evidence of the actual non-glycosylated status of the 2mEPSPS plant-derived protein is provided from two sources. Firstly, the protein ran at the expected molecular weight in SDS-PAGE/western blot analyses; glycosylation would increase the apparent molecular weight. Secondly, electrospray liquid chromatography/mass spectrometry (LC/MS) did not indicate any difference between the tryptic peptides from bacterial and plant sources. It would not be expected that the plant-derived proteins would be glycosylated since chloroplastic proteins targeted directly to the chloroplast do not transit through the endoplasmic reticulum where glycosylation occurs in eukaryotes (Pattison and Amtmann, 2009).

4.4.3 Equivalence of plant- and microbially-produced proteins

Having established the authenticity and characteristics of the bacterially-derived HPPDPf W336 and 2mEPSPS proteins, it was then necessary to confirm whether they were equivalent to the plant-derived proteins. A number of analyses were included as follows.

4.4.3.1 Molecular weight

The molecular weights of the FG72- and *E. coli*-derived 2mEPSPS and HPPDPf W336 protein preparations were estimated on SDS-PAGE. The electrophoretic mobility of the *E. coli* produced 2mEPSPS protein and the plant produced 2mEPSPS protein were similar; the molecular weights subsequently calculated were 50.3 kDa for the microbial protein and 49.7 kDa for the plant protein. For HPPDPf W336 the electrophoretic mobility and the calculated molecular weight (42.5 KDa) were the same for both the bacterially- and plant-derived proteins.

4.4.3.2 Immunoreactivity

The identity and immunoreactivity of the plant- and *E. coli*-derived 2mEPSPS and HPPDPf W336 proteins were investigated by Western blot analysis. Immunoreactivity was tested by incubating blotted polyvinylidene fluoride membranes as follows:

- For HPPDPf W336: monoclonal mouse anti-HPPD Pf W336 followed by a commercially available (BioRad) goat anti-mouse enzyme-linked (horseradish peroxidase) secondary antibody.
- For 2mEPSPS: polyclonal rabbit anti-2mEPSPS followed by a commercially available (BioRad) goat anti-rabbit enzyme linked (horseradish peroxidase) secondary antibody.

The observed immunoreactivities of each protein, whether sourced from FG72 or *E. coli* were the same.

4.4.3.3 HPLC/Electrospray Mass Spectrometry (LC/MS)

Peptide maps were produced from the selected ion chromatograms obtained for tryptic digests of the plant- and *E. coli*-derived 2mEPSPS and HPPDPf W336 proteins.

Based on the theoretical peptide sequences, it was estimated that the microbially-produced 2mEPSPS protein provided 98% coverage while the plant-produced protein provided 71% coverage. The reduced coverage for the plant derived protein is most likely a reflection of the limited amount of enzyme available for analysis. However, it is considered to be adequate to provide convincing evidence of identity. The mass spectrometry analysis also indicated that the N-terminal methionine, as expected, had been cleaved from the bacterial- and plant-produced protein (see discussion in Section 4.4.3.4).

For the HPPDPf W336 protein, there was 95.2% coverage of the bacterial protein and 70.1% coverage of the plant protein. As for the 2mEPSPS protein the N-terminal methionine of HPPDPfW336 from both sources was not detected.

4.4.3.4 N-terminal sequence analysis

The first five amino acids of the proteins were analysed on a protein sequencer following Edman degradation. For the 2mEPSPS protein, the chromatogram showed that the N-terminal methionine in both microbially- and plant-derived proteins had been cleaved. This

result was anticipated since the terminal methionine is routinely cleaved from nascent proteins by methionine aminopeptidase (Polevoda and Sherman, 2000). For the remaining amino acids, technical difficulties due to the homogeneity of the samples and blocking (a common and often unavoidable problem encountered in protein sequencing) prevented a clear analysis of the 2mEPSPS protein from both sources.

For the HPPDPf W336 protein, there was insufficient yield of the plant protein to permit Edman degradation. The microbially-derived protein also had the terminal methionine missing; the sequence of the remaining amino acids was as predicted from the DNA sequence.

4.4.3.5 Enzymatic activity

Using the method described in Section 4.4.1, the *E. coli*-produced and plant- produced 2mEPSPS proteins were shown to be active by the release of phosphate during the reaction of shikimate-3-phosphate and phosphoenolpyruvate with 2mEPSPS to form 5-enolpyruvlshikmate-3-phospate. The non-GM 'Jack' soybean also was shown to be active for EPSPS. This is to be expected since soybean contains a wild type EPSPS, which should also react with the assay.

Insufficent HPPDPf W336 protein could be obtained from FG72 soybean to permit an activity study.

4.4.3.6 Conclusion

Multiple analyses confirmed the identity of the 2mEPSPS and HPPDPf W336 proteins in soybean line FG72 and established the equivalence of the plant-produced protein to the bacterially-produced in both cases. From this it is concluded that the 2mEPSPS and HPPDPf W336 proteins produced in *E. coli* are therefore valid surrogates for use in further safety evaluations of the novel proteins.

4.5 Potential toxicity of the novel proteins

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.

The assessment focuses on: whether the novel protein has a prior history of safe human consumption, or is sufficiently similar to proteins that have been safely consumed in food; amino acid sequence similarity with known protein toxins and anti-nutrients; structural properties of the novel protein including whether it is resistant to heat or processing and/or digestion. Appropriate oral toxicity studies in animals may also be considered, particularly where results from the biochemical, bioinformatic, digestibility or stability studies indicate a concern.

4.5.1 History of human consumption

As outlined in Section 4.2, homologues of both proteins are found in plants and a range of other organisms and would therefore be routinely consumed 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).

4.5.2 Similarities with known protein toxins

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

Studies submitted:

Capt, A. (2009). HPPD W336 protein amino acid sequence homology search with known toxins. ID# M-355651-01-1, Bayer CropScience (unpublished).

Capt, A. (2008). 2mEPSPS protein overall amino acid sequence homology search with known toxins and allergens. ID# M-273792-02-1, Bayer CropScience (unpublished).

4.5.2.1 HPPDPf W336

The amino acid sequences of HPPDPf W336 were compared with protein sequences present in a number of large public reference databases: eg Uniprot_Swissprot, Uniprot_TrEMBL, PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept. The similarity search used the BLASTP (Basic Local Alignment Search Tool Protein) algorithm (Altschul *et al*., 1997), Version 2.2.20 and the BLOSUM62 scoring matrix.

BLASTP is now frequently applied for searching for similarities in protein sequences by performing local alignments. This detects more similarities that would be found using the entire query sequence length. A parameter known as the *E* value (see eg Baxevanis, 2005) represents the probability that a particular alignment is due to random chance. 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).

As expected, the query sequence matched with HPPD proteins from a number of organisms. In addition, an identity of 54% (192 identical residues) was found with the VLLY protein, described as a hemolysin from *Vibrio vulnificus* a pathogenic bacterium and an identity of 50% was found between HPPD and legiolysins (LLY) produced by strains of *Legionella pneumophila.* Further bioinformatics analyses were done to investigate the similarities and it was concluded that both the VLLY and LLY proteins are HPPD-like proteins that share common domains with HPPD W336. Research has indicated that neither the VLLY nor LLY proteins themselves are directly involved in haemolytic activity but rather that it is oxidation or polymerisation of the homogentisate formed by the action of the two proteins that leads to toxicity (Heoedus and Nayak, 1994; Chang *et al*., 1997; Steinert *et al*., 2001) i.e. HPPD is not directly haemolytic.

4.5.2.2 2mEPSPS

The same reference databases were used as for the HPPDPf W366 toxin searches. Two kinds of searches were done:

- an overall identity search using the FASTA algorithm and BLOSUM62 scoring matrix.
- an 80-mer sliding window search using the FASTA algorithm and BLOSUM50 scoring matrix.

The criterion indicating potential toxicity was a 35% identity, over at least 80 consecutive amino acids, with a toxin.

The FASTA algorithm is essentially the same as the BLAST algorithm but uses a slightly different statistical approach that provides a Z-score. This describes the number of standard deviations from the mean score for the database search. The higher the Z-score for a reported match, the more significant the match, with a score of >15 indicating an extremely significant match (Xiong, 2006).

Similarities were found between 2mEPSPS and other EPSPS proteins or with the arom[5] protein from various origins, none of which are known toxins.

4.5.3 In vitro *digestibility*

See Section 4.6.3

4.5.4 Thermolability

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

Studies submitted:

Rascle, J.B. (2009). HPPD W336 protein heat stability study. ID# M-354574-01-2, Bayer CropScience (unpublished).
Rouquie, D. (2007). 2mEPSPS protein heat stability study. ID# M-293053-02-1, Bayer CropScience (unpublished).
Habex, V. (2011). The modified 4-hydroxyphenylpyruvate dioxygenase gene product (HPPD W336) description and characterization. Report ID BIO2-026_Protdescript_445, Bayer CropScience (unpublished).

HPPDPF W336 and 2mEPSPS proteins purified from bacterial expression systems (refer to Section 4.3.2) were solubilised in buffer and then incubated in a temperature-controlled hot-block at 60°, 75° or 90°C for 10, 30 or 60 minutes. The integrity of the proteins was detected by gel electrophoresis (SDS-PAGE) coupled with a Coomassie blue staining, and by a Western blot analysis.

Additionally, for the HPPDPf W336 protein, the protein was incubated at 45°C, 60°C and 95°C for 2.5, 5, 10, 20 and 60 minutes. Subsequently the activity of the protein was assessed under standard conditions (refer to Section 4.4.1).

[5] The arom protein catalyses steps in the shikimate pathway of microbial eukaryotes.

4.5.4.1 HPPDPf W336

No significant changes in apparent molecular weight were indicated by SDS-PAGE at any incubation conditions or times. Similarly, Western blot analysis (rabbit polyclonal anti-HPPDPf W336 primary antibody; polyclonal goat anti-rabbit, peroxidise-linked secondary antibody) indicated that maximum immunoreactivity was detectable in all temperature x time combinations. From this it was concluded that the HPPDPf W336 protein is structurally stable at 90° C for 60 minutes. However, as anticipated the enzymatic activity dropped below 50% after the protein was incubated at 45°C for 20 minutes. At more elevated temperatures (60°C and 95°C), the HPPDPf W336 activity was completely destroyed after 2.5 minutes.

4.5.4.2 2mEPSPS

No significant changes in apparent molecular weight were indicated by SDS-PAGE after heat treatment at 60° or 75° for any of the times. Marked but not complete denaturation of the protein was noted after incubation at 90° for 60 min.

Western blot analysis (rabbit polyclonal anti-2mEPSPS primary antibody; polyclonal goat anti-rabbit alkaline phosphatase-linked secondary antibody) indicated there was no loss of immunoreactivity at the highest temperature x time combination.

In a previous application in which the 2mEPSPS protein was considered (Application A614 - Food derived from Glyphosate-Tolerant Cotton Line GHB614), the Applicant tested the activity of the enzyme at increasing temperature. The activity increased linearly to approximately 60° C then decreased sharply above 60° C. The enzyme appeared to be inactive at 75°

It is concluded that the 2mEPSPS protein is heat-stable up to 90° C for 60 minutes but that activity of the enzyme is destroyed at 75° C or higher.

4.5.5 Acute toxicity studies

Acute toxicity is generally assessed using the oral route of exposure. In this instance, the Applicant also supplied an intravenous injection study for each protein. These latter studies were not considered by FSANZ as they are not relevant for dietary risk assessment.

Studies submitted:

Rascle, J.B. (2009). HPPD W336 protein acute toxicity by oral gavage in female mice. Study SA 09201, Bayer CropScience (unpublished).

Rouquie, D. (2006). 2mEPSPS protein acute oral toxicity by oral gavage in mice. ID# M-276952-01-1, Bayer CropScience (unpublished).

Rascle, J.B. (2009). HPPD W336 acute toxicity by intravenous injection in mice. ID# M-357601-01-1, Bayer CropScience (unpublished). NOTE: Study not considered by FSANZ.

Rouquie, D. (2008). 2mEPSPS protein acute toxicity by intravenous injection in mice. ID# M-297233-01-1, Bayer CropScience (unpublished). NOTE: Study not considered by FSANZ.

4.5.5.1 Acute toxicity – oral gavage

The study design for each protein is given in Table 6.

Table 6: Study design for acute oral toxicity testing

Test material	HPPDPf W336 produced in *E. coli*	2mEPSPS produced in *E. coli*
Vehicle	50 mM Tris pH 7.5	Storage buffer (0.1 M Tris/HCl, 2.7 mM KCl, 137 mM NaCl, 1 mM Dtt pH 7.5)
Test Species	Crl:OF1 mice (five females) – approx. 8 weeks old on day of treatment	Crl:OF1 mice (five females) – approx. 8 weeks old on day of treatment
Dose	2 x separate doses of test substance by oral gavage, within 4 h. Actual total dose of 2,000 mg/kg body weight HPPDPf W336**	2 x separate doses of test substance by oral gavage, within 4 h. Actual total dose of 2,000 mg/kg body weight 2mEPSPS**
Control	5 female mice administered 2,000 mg/kg BSA in Tris	5 female mice administered 2,000 mg/kg BSA in storage buffer
Length of study	15 d	15 d

** The dose of 2,000 mg/kg body weight is the maximum unexceptional dose recommended by the OECD for the testing of acute oral toxicity using the fixed dose procedure (OECD, 2001b).

Mice were observed for mortality, body weight gain and clinical signs over 15 days. At the end of the study all animals were killed and examined for organ or tissue damage or dysfunction. All mice in both control and test treatments survived for the duration of the studies. No clinical signs of systemic toxicity were observed in either test or control treatments. No macroscopic abnormalities attributable to the administration of the test proteins were present in the mice at necropsy on day 15. Under the conditions of the studies, oral administration of HPPDPf W336 or 2mEPSPS protein to female mice at the limit dose of 2,000 mg /kg bw produced no test substance-related clinical signs of toxicity, body weight losses, macroscopic abnormalities or mortality.

The oral toxicity studies support the conclusion that neither the HPPDPf W336 nor 2mEPSPS protein is acutely toxic.

4.6 Potential allergenicity of the novel proteins

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 eg 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 the novel protein to act as an allergen.

The allergenic potential of the HPPDPf W336 and 2mEPSPS proteins was assessed by:

- considering the source of the gene encoding each protein and the history of use or exposure

- comparing the alignment similarity of the amino acid sequences with known protein allergens
- evaluation of the stability, under simulated gastrointestinal conditions, of the microbially produced HPPDPf W336 and 2mEPSPS proteins.

4.6.1 Source of each protein

The HPPDPf W336 protein is derived from a common soil bacterium to which humans have been naturally exposed and which may have been inadvertently ingested on fresh produce without eliciting adverse effects. *Pseudomonas fluorescens* has not been reported to be a source of allergenic proteins (OECD, 1997).

The 2mEPSPS 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 elsewhere, have not themselves been implicated in any food-related allergic reactions.

4.6.2 Similarity to known allergens

Study submitted:

Capt, A. (2008). 2mEPSPS protein epitope homology and N-glycosylation searches. ID# M-273851-02-1, Bayer CropScience (unpublished).

Capt, A. (2008). 2mEPSPS protein overall amino acid sequence homology search with known toxins and allergens. ID# M-273792-02-1, Bayer CropScience (unpublished).

Capt, A. (2009). HPPD W336 protein amino acid sequence homology search with known allergens. ID# M-355637-01-1, Bayer CropScience (unpublished).

Bioinformatic analysis provides 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).

4.6.2.1 HPPDPF W336

Several bioinformatic approaches were used to evaluate the potential allergenicity of the HPPDPf W336 protein:

- A search to identify any short sequences of amino acids that might give rise to a shared allergenic epitope. This search compared the HPPDPf W336 sequence with known allergens in the AllergenOnline database using the FindPatterns algorithm. The criterion indicating allergenicity was a 100% identity, on a window of 8 consecutive amino acids, with an allergenic protein.
- An overall (i.e. complete amino acid sequence) identity search using the AllergenOnline database and the FASTA (refer to Section 4.5.2) algorithm with

BLOSUM62 scoring matrix. The criterion indicating allergenicity was a 35% identity, over 80 amino acids or more, with an allergenic protein.

- An 80-mer sliding window search using the AllergenOnline database and the FASTA algorithm with BLOSUM50 scoring matrix. This type of search subdivides the sequence of the query protein into 80-mer blocks (i.e. aa 1-80; then 2-81, then 3-82 etc). Only matches of at least 35% identity with an allergenic protein are considered.

None of the searches found any significant amino acid sequence homology with known allergens.

4.6.2.2 2mEPSPS

The same bioinformatic approaches were used as for the HPPDPf W336 protein:

As with the toxin search, similarities were found between 2mEPSPS and other EPSPS proteins or with the arom protein from various origins, none of which are known allergens. No matches were detected with known allergenic epitopes.

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, and this enables them to be exposed to the intestinal mucosa that leads to an allergic response (Astwood and Fuchs, 1996; Metcalfe *et al*., 1996; Kimber *et al*., 1999). Therefore a correlation exists between resistance to digestion by pepsin and potential allergenicity although this may be inconsistent (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 unlikely 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 HPPDPf W336 and 2mEPSPS. In addition to the pepsin protocol using simulated gastric fluid (SGF), a second digestibility study was done using simulated intestinal fluid (SIF) containing pancreatin, which is a mixture of enzymes including amylase, trypsin, lipase, ribonuclease and protease. The SIF study by itself, however, may not be entirely informative because ordinarily an ingested protein would first be exposed to pepsin-mediated hydrolysis in the acidic environment of the stomach before being subject to further digestion in the small intestine.

4.6.3.1 Simulated gastric fluid (SGF) studies

Studies submitted:

Rascle, J.B. (2009). HPPD W336 protein *in vitro* digestibility study in human simulated gastric fluid. ID# M-356196-01-1, Bayer CropScience (unpublished).

Rouquie, D. (2011). 2mEPSPS protein *in vitro* digestibility study in human simulated gastric fluid. ID# M-406126-01-1, Bayer CropScience (unpublished).

The *in vitro* digestibility of the *E. coli*-derived HPPDPf W336 and 2mEPSPS proteins in SGF (U.S.Pharmacopeia, 1990) containing pepsin at pH 1.2 was evaluated by incubating samples at 37º for selected times (0, 0.5, 2, 5, 10, 20, 30 and 60 minutes) and then running them on SDS-PAGE gels. Proteins were visualised by Coomassie staining of the resulting gel. Two control proteins were treated in parallel: horseradish peroxidase (HRP) is known to

hydrolyse readily in pepsin and served as a positive control; ovalbumin (OVA) is known to be only slowly hydrolysed in pepsin and was used as a negative control.

Western blotting of the SDS-PAGE gels were also performed using for both proteins, a polyclonal rabbit primary antibody and a polyclonal goat anti-rabbit peroxidise-linked secondary antibody.

Both the SDS-PAGE gels and Western blots indicated that more than 90% of the HPPDPf W336 and 2mEPSPS proteins were degraded within 30 s. The HRP positive control was rapidly hydrolysed (< 30 seconds) while the OVA negative control showed the expected slow rate of hydrolysis..

4.6.3.2 Simulated intestinal fluid (SIF) studies

Studies submitted:

Rascle, J.B. (2009). HPPD W336 protein *in vitro* digestibility study in human simulated intestinal fluid. ID# M-356198-01-1, Bayer CropScience (unpublished).
Rouquie, D. (2011). 2mEPSPS protein *in vitro* digestibility study in simulated intestinal fluid. ID#M-275371-02-1, Bayer CropScience (unpublished).

The *in vitro* digestibility of the *E. coli*-derived HPPDPf W336 and 2mEPSPS proteins in SIF (U.S.Pharmacopeia, 1990) containing pancreatin at pH 7.5 was assessed by incubating samples at 37^{o} C for specified time intervals (0, 0.5, 2, 5, 10, 20, 30 and 60 minutes), and analysing by SDS-PAGE with Coomassie staining, and Western blotting using appropriate antibodies (see 4.6.3.1).

For both proteins, the SDS-PAGE and Western blot analysis showed that there was rapid degradation in SIF within 30 s.

4.7 Conclusion

Soybean line FG72 expresses two novel proteins, HPPDPf W336 and 2mEPSPS, both of which were detected in all plant parts analysed. HPPDPf W336 was lowest in the seed (approximately 1.5 µg/g dry weight) and highest in younger leaves (approximately 38 µg/g dry weight). 2mEPSPS protein concentrations were much higher than those for HPPDPf W336. The leaves contained the highest levels (older leaves contained approximately 660 µg/g dry weight) while roots contained the lowest levels (approximately 40 µg/g dry weight). The level of 2mEPSPS in the seed was approximately 150 µg/g dry weight. During processing of the seed, the HPPDPf W336 protein may be concentrated to a small degree in hulls and protein isolate, and is undetectable in other processed fractions. Levels of the 2mEPSPS protein are reduced in all fractions during processing, being undetectable in toasted meal, crude lecithin and all forms of oil.

Several studies were done to confirm the identity and physicochemical properties of the plant-derived HPPDPf W336 and 2mEPSPS proteins, and demonstrated that they both conform in size and amino acid sequence to that expected, and do not exhibit any post-translational modification including glycosylation.

For both proteins, bioinformatic studies have confirmed their lack of any significant amino acid sequence similarity to known protein toxins or allergens and digestibility studies have demonstrated that the proteins would be rapidly degraded in the stomach following ingestion. Acute toxicity studies in mice have also confirmed their absence of toxicity in animals. Both proteins exhibit a degree of heat stability however, given their digestive lability

combined with their lack of similarity to known protein toxins or allergens and the loss of enzyme activity with heating, this does not raise any safety concerns.

Taken together, the evidence indicates that HPPDPf W336 and 2mEPSPS are unlikely to be toxic or allergenic to humans.

5. Herbicide metabolites

There are essentially three strategies available for making plants tolerant to herbicides:

- detoxifying the herbicide with an enzyme which transforms the herbicide, or its active metabolite, into biologically inactive products.
- inducing mutation(s) in the target enzyme so that the functional enzyme is less sensitive to the herbicide, or its active metabolite.
- inducing over-expression of the sensitive enzyme so that the concentration of target enzyme in the plant is sufficient in relation to the inhibiting herbicide so as to have enough functional enzyme available despite the presence of the herbicide.

In the case of herbicide-tolerant GM lines using the first strategy described above there is the possibility that novel metabolites are produced following application of the herbicide and these metabolites may be present in the final food. It is therefore necessary for those lines incorporating a herbicide/gene combination not previously assessed, to establish whether such metabolites occur. If they do, their toxicity needs to be determined in order to enable the establishment of an appropriate health-based guidance value (e.g. Acute Reference Dose – ARfD; Acceptable Daily Intake – ADI). Residue data also need to be considered to confirm the concentration of the novel GM trait-specific metabolites relative to the parent herbicide in the final food.

As discussed in Section 4.2.1, the way in which the HPPDPf W336 provides plant tolerance to isoxaflutole is by having a lower sensitivity to DKN (higher k_m and higher k_{on}) than the wild type HPPD. As there is no inactivation *per se* of isoxaflutole in FG72 plants there is no production of metabolites that are not already found in a non-GM plant.

The metabolism of isoxaflutole in nature follows the same pathway in plants and animals as well as undergoing the same degradation in soil and aqueous systems (see eg Pallett *et al*., 1998; Pallett *et al*., 2001; Beltrán *et al*., 2003; Rupprecht *et al*., 2004). The first step involves the opening of the isoxazole ring to form diketonitrile which is then further metabolised to herbicidally inactive benzoate derivatives.

Use of the *epsps* gene to confer tolerance to glyphosate has been considered in a wide range of food crops, including soybean, and therefore glyphosate residues will not be considered in this Assessment. The Applicant did supply a study (Beedle and Dallstream – see below) in which both isoxaflutole and glyphosate were sprayed on FG72 plants. The results from this study are as would be expected given what is known about the metabolism of both herbicides and do not raise any safety concerns.

Studies submitted:

Nguyen, T (2010). The metabolism of [phenyl-UL-^{14}C]-Isoxaflutole in soybean with pre-plant and post-emergent application. ID# M-368555-01-1, Bayer CropScience (unpublished).
Dallstream, K.A.; Fischer, D.R. (2010). Balance® Pro 480 SC - Magnitude of the residue in/on soybeans. ID# M-368661-01-1, Bayer CropScience (unpublished).
Fischer, D.R. (2010). Balance® Pro 480 SC - Magnitude of the residue in soybean processed commodities and aspirated grain fractions. ID# M-368662-01-1, Bayer CropScience (unpublished).
Beedle, E.C.; Dallstream, K.A. (2010). Balance® Pro 480 SC and Glyfos®- Magnitude of the residue in/on soybeans. ID# M-368669-01-1, Bayer CropScience (unpublished).

5.1 Metabolism of isoxaflutole

Seeds of FG72 were sown in containers (5 seeds/container) in a glasshouse and germinated plants were allowed to grow to harvest maturity. One of two [phenyl-UL-^{14}C]-isoxaflutole treatments was applied once via a hand held pump action sprayer to each of five containers:

- at pre-emergence (same day as sowing) when soil in the containers was sprayed with the equivalent of a field application rate of 330 g active ingredient/ha (ai/ha), which is 3.1X the anticipated maximum annual field application.
- at full flowering stage BBCH 65 (see Munger *et al*., 1997 for a description of the BBCH stages) also equivalent to a field application rate of 330 g ai/ha.

At stage BBCH75 (50% of pods have reached final length and continuation of pod filling), plants from both treatments were harvested for analysis of forage. The remaining plants were harvested at full maturity stage BBCH99 (above ground plant parts are dead and seeds are dry and hard) for analysis of hay and seeds. Identification and quantitation of the residues in forage, hay and seed extracts were accomplished by using reverse phase high performance liquid chromatography (HPLC) and liquid chromatography/mass spectrometry-mass spectrometry (LC/MS-MS).

The results presented in Table 7 show that the major residues after a single application of [phenyl-UL-^{14}C]-isoxaflutole were isoxaflutole (IFT), diketonitrile (DKN), a benzoate (IFT acid) and a benzamide (IFT amide), none of which is novel

Table 7: Metabolites in raw agricultural commodities (forage, hay or seed) of soybean FG72 following a single spray application of isoxaflutole at either pre-emergent or full flowering stage. Values represent the percentage of total radioactive residues. ND = not detectable.

Compound	Forage		Hay		Seed	
	Pre-emergent	Full flowering	Pre-emergent	Full flowering	Pre-emergent	Full flowering
IFT	ND	72	ND	25	ND	ND
DKN	13	18	13	21	17	24
IFT amide	53	ND	13	3	8	8
IFT acid	27	6	56	38	66	62
Total	93	96	82	87	92	94

IFT = isoxaflutole = (5-Cyclopropyl-4-isoxazolyl)[2-(methylsulfonyl)-4-(trifluoromethyl)phenyl] methanone
DKN = α-(Cyclopropylcarbonyl)-2-(methylsulfonyl)-β-oxo-4-(trifluoromethyl) benzenepropanenitrile
IFT amide = 2-(methylsulfonyl)-4-(trifluoromethyl) benzamide
IFT acid = 2-(Methylsulfonyl)-4-(trifluoromethyl) benzoic acid

5.2 Isoxaflutole residue chemistry studies

5.2.1 Soybean seed

Field trials were conducted in 2009/210 at 20 sites in the U.S. The trials were located in areas where soybean is commonly grown commercially. Four plots were included in each trial, one control plot and three treated plots (TRIPR, TR1V3 and TR1R1) where isoxaflutole was applied as Balance® Pro 480 SC as follows:

- For the TRIPR plot, a soil application was made in one of three application patterns; either a single application just prior to planting (pre-plant), a single application followed by incorporation just prior to planting (pre-plant with incorporation), or a single application following planting and just prior to emergence (pre-emergent). Seed was planted within three days of the application event.

- For the TRIV3 plot, a single foliar application was made at or just prior to stage BBCH 13 (trifoliate leaf on the 3rd node unfolded) to BBCH 15 (trifoliate leaf on the 5th node unfolded)

- For the TRIR1 plot, a single foliar application was made at or just prior to stage BBCH 51 (first flower buds visible) to BBCH 66 (about 60% of flowers open)

All applications were made, using ground based equipment, at a rate (105 g ai/ha) equivalent to the maximum seasonal rate. Seed was collected from the trials at commercial maturity (BBCH 89) for analysis of total isoxaflutole (sum of isoxaflutole + diketonitrile).

The residue levels in soybean seed harvested at commercial maturity are summarised in Table 8.

Table 8: Levels of total isoxaflutole (sum of isoxaflutole and diketonitrile), expressed as ppm, in seed of soybean line FG72 sprayed with Balance® Pro 480 SC (Limit of detection (LOD) = 0.0017 ppm)

Treatment	PHI[1] (days)	No. of samples	Residue Levels[2] (ppm)					
			Min	Max	HAFT[3]	Median[4]	Mean	SD
TP1PR	111 - 151	40	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
TR1V3	87 - 120	40	<0.01	0.014	0.014	<0.01	<0.01	<0.01
TR1R1	72 - 126	40	<0.01	0.029	0.027	<0.01	<0.01	<0.01

[1]PHI = Pre-harvest interval; days between last application of herbicide and collection of field sample
[2]Any analyte residue measured to be less than the Limit of Quantitation (LOQ) was reported as <0.01 ppm.
[3]HAFT = highest average field trial
[4]Median is the geometric median value of the residue values reported in the treated samples

Based on the proposed use of isoxaflutole in soybean line FG72, the residue levels in soybean seed are all essentially less than the limit of quantitation.

5.2.2 Processed soybean commodities

A field trial was undertaken at a single site in the U.S. (Springfield, NE). There were 3 plots in the trial: one control plot, and two treated plots where Balance® Pro 480 SC was applied to FG72 plants at BBCH 60 [first flowers opened (sporadically in population)]. Application was made via a single broadcast foliar spray to plants in each of the two treated plots at a rate of 0.308 kg ai/ha (plot TRT3X) and 0.525 kg ai/ha (plot TRT5X). These applications to the two treated plots represented three times (3X) and five times the maximum recommended rate.

Samples for analysis (total isoxaflutole = sum of isoxaflutole + diketonitrile) were taken only from plot TRTX5, when plants had reached earliest commercial harvest (stage BBCH 89). Seeds were processed into meal, hulls, oil, milk and aspirated grain fractions[6]. The residue levels are summarised in Table 9. Meal and aspirated grain fraction, both of which are used in animal feed, showed some concentration of isoxaflutole residue.

Table 9: Levels of total isoxaflutole (sum of isoxaflutole and diketonitrile), expressed as ppm, in seed and processed fractions of soybean line FG 72 sprayed with Balance® Pro 480 SC at 5X the maximum recommended rate.

Commodity	Average total isoxaflutole residue (ppm)[1]
Seed (Raw Agricultural Commodity)	0.043
Soybean meal	0.052
Soybean hulls	0.032
Soybean oil	<0.01[2]
Soymilk	<0.01
Soybean aspirated grain fraction	0.312

[1] Any analyte residue measured to be less than the Limit of Quantitation (LOQ) was reported as <0.01 ppm.
[2]The level of residue in soybean oil was less than the LOD (0.0017 ppm)

5.3 ADI for isoxaflutole

As no novel herbicide metabolites are present in treated soybean line FG72, the existing health-based guidance value (i.e. Acceptable Daily Intake - ADI) for isoxaflutole is appropriate and relevant for assessing dietary risk with soybean line FG72. In Australia the ADI for isoxaflutole is 0.02 mg/kg bw/day[7].

[6] Aspirated grain fractions are those plant parts obtained during normal aspiration of cereal and oil seed crops in the handling of the product and consist primarily of plant parts, including glumes and contain not more than 15 percent ash (dirt). [American Feed Control Officials definition 60.43]. They are used in animal feed.

[7]ADIs are established by the Office of Chemical Safety within the Department of Health and Ageing http://www.health.gov.au/internet/main/publishing.nsf/Content/E8F4D2F95D616584CA2573D700770C2A/$File/ADI-report-sept10.pdf

5.4 Conclusion

The residues generated on soybean line FG72 as a result of spraying with isoxaflutole are the same as those found on conventional crops sprayed with isoxaflutole. Residue data derived from supervised trials indicate that the residue levels in seed are below the limit of quantitation and that there is some concentration of residue in meal and aspirated grain fractions but not in other processed commodities. In the absence of any measurable exposure to either parent herbicide or metabolites the risk to public health and safety is likely to be negligible.

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

For soybean a number of components are considered to be important for compositional analysis (OECD, 2001a). As a minimum, the key nutrients of soybean seed appropriate for a comparative study include the proximates (crude protein, fat, ash, acid detergent fibre (ADF) and neutral detergent fibre (NDF)), amino acids and fatty acids. In addition, mineral and vitamin levels may be considered and international guidance also suggests that levels of the key anti-nutrients phytic acid, trypsin inhibitors, lectins, isoflavones and stachyose and raffinose should be determined for new varieties of soybean. The reasons for focussing on these particular anti-nutrients are:

- phytic acid causes chelation of mineral nutrients (including calcium, magnesium, potassium, iron and zinc) thereby making them unavailable to monogastric animals, including humans
- trypsin inhibitors interfere with digestion of protein; lectins are proteins that bind to carbohydrate-containing molecules. Both trypsin inhibitors and lectins can inhibit animal growth. The activity of trypsin inhibitors and lectins is heat-labile and they are inactivated during processing of soybean protein products and soybean meal so that the final edible soybean product should contain minimal levels of these anti-nutrients.
- isoflavones are reported to possess biochemical activity including estrogenic, anti-estrogenic and hypocholesterolaemic effects that have been implicated in adversely affecting animal reproduction. Major isoflavones in soybeans include daidzein, genistein, glycitein and coumestrol.
- stachyose and raffinose are low molecular weight carbohydrates (oligosaccharides) that are associated with production of intestinal gas and resulting flatulence when they are consumed.

6.2 Study design and conduct for key components

Studies submitted:

Mackie, S.J.W. (2009). Composition of seed from FG72 soybean and its non-transgenic counterpart. ID# M-355723-01-1, Bayer CropScience (unpublished).
Rattemeyer-Matschurat, V. (2009). Analysis of substantial equivalence of double-herbicide-tolerant soybean. ID# M-356445-03-1, Bayer CropScience (unpublished).
Oberdörfer, R. (2011). Nutritional impact assessment report for the double-herbicide-tolerant soybean (Transformation event FG72). ID# M-357280-03-1

The test (FG72, seed of T_8 lineage), and control ('Jack') were grown under typical production conditions at ten field sites across North America[8] during the 2008 growing season. 'Jack' is the original transformed line and therefore represents the isogenic control line for the purposes of the comparative analyses. The identity of the test and control lines was verified by Real-Time Discriminating PCR. Additionally, 3 non-GM commercial lines (Stine® lines 2686-6; 2788 and 3000-0) were grown at each site in order to generate tolerance ranges for each analyte.

At each site, there were 6 treatments as outlined in Table 10. In total there were 120 samples taken over all sites.

Table 10: Study design for each trial site

Treatment	Line	Herbicide treatment	Replicates
A	Jack	unsprayed	3
B	FG72	unsprayed	3
C	FG72	sprayed*	3
D	2686-6	unsprayed	1
E	2788	unsprayed	1
F	3000-0	unsprayed	1

*sprayed = one application of Balance® Pro, Roundup Original Max® and ammonium sulphate at post V4-V5 growth stage.

Seed was harvested at normal maturity by hand or mechanical means and samples were shipped frozen to Wisconsin for analysis. Samples were analysed for proximates, minerals, anti-nutrients, total amino acids, total fatty acids, vitamins and isoflavones. Methods of composition analysis were based on internationally recognised procedures (e.g. those of the Association of Official Analytical Chemists), methods specified by the manufacturer of the equipment used for analysis, or other published methods.

6.3 Analyses of key components

For each analyte 'descriptive statistics' were generated i.e. a mean and standard deviation averaged over all sites for each of Treatments A, B and C. The values thus calculated are presented in Tables 12 – 18.

For statistical analysis, data were transformed into Statistical Analysis Software[9] (SAS) data sets and analysed using SAS version 8.2. In assessing the significance of any difference between means, a P-value of 0.05 was used (i.e. a P- value of ≥0.05 was not significant). Analysis of Variance (ANOVA) was used for over-all analysis with the factors TREAT, SITE

[8] The ten sites were: Marcus, IA ; Iowa Falls , IA; Scranton/Glidden, IA; Perry, IA; Adel, IA; Winterset, IA; Osborn, MO; Fithian, IL; Sharpsville, IN; Mediapolis, IN.
[9] SAS website - http://www.sas.com/technologies/analytics/statistics/stat/index.html

and the respective interaction. Based on the ANOVA model, treatment differences (A vs B and A vs C) were estimated and presented together with 95% confidence intervals (data not shown). Where a significant difference was found in the ANOVA, a t-test was then carried out to generate a P-value for each of the comparisons A vs B and A vs C. In cases of significant TREAT x SITE interactions (where $p<0.05$), treatment comparisons are not valid and a by-site ANOVA was performed individually for each site with factor TREAT followed by t-tests to compare Treatment A vs Treatment B and Treatment A vs Treatment C.

The results for the three treatments were compared to a combined literature range for each analyte, compiled by the Applicant from published literature for commercially available soybean[10]. Any statistically significant differences between FG72 and the 'Jack' control were also compared to the tolerance range compiled from the results of the non-GM commercial cultivars, to assess whether the differences were likely to be biologically meaningful. 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 (Harrigan *et al*., 2010). Therefore, even if means fall outside the published range, this is not necessarily a concern.

6.3.1 Proximates and fibre

Results of the proximate and fibre analysis are shown in Table 11. There was no significant difference between the control and FG72 in terms of the levels of protein, fat, ADF and NDF. The mean level of ash was significantly lower in both the sprayed and unsprayed soybean FG72 (shaded cells) compared to the level in 'Jack' but the means for both soybean FG72 treatments and for 'Jack' were within the range reported in the literature as well as within the tolerance range of the 3 commercial cultivars.

For the carbohydrate levels, there was a significant TREAT x SITE interaction but the majority of the individual by-site analyses did not show significant differences between treatments (A vs B and A vs C).

Table 11: Mean (±standard deviation) percentage dry weight (%dw) of proximates and fibre in seed from 'Jack' and FG72.

Analyte	A 'Jack' (%dw)	B FG72-unsprayed (%dw)	C FG72-sprayed (%dw)	Overall treat effect (P-value)	Tolerance range (%dw)	Combined literature range (%dw)
Protein	38.2±1.1	38.2±0.8	38.1±0.9	0.799	35.8 – 40.1	32.0 – 45.5
Fat	19.3±0.9	18.9±1.2	19.2±1.1	0.064	15.1 – 21.4	8.10 – 24.7
Carbohydrate[1]	37.3±1.2	37.9±1.0	37.6±1.2	N/A[2]	34.8 - 41.6	29.6 – 50.2
ADF	17.8±1.9	18.1±2.0	17.9±1.8	0.832	13.6 – 23.5	7.81 – 18.6
NDF	19.8±2.0	20.3±2.1	20.0±1.5	0.5	16.1 – 24.8	5.0 – 21.3
Ash	5.24±0.31	5.07±0.3	5.06±0.28	<0.001	4.89 – 5.73	3.89 – 6.99

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

[2] NA = not applicable because of TREAT x SITE interaction

6.3.2 Fatty Acids

The levels of 28 fatty acids were measured. Of these, the following were below the limits of quantitation - C8:0 caprylic, C10:0 capric, C12:0 lauric, C14:0 myristic, C14:1 myristoleic, C15:0 pentadecanoic, C15:1 pentadecenoic,C16:1 palmitoleic, C17:0 heptadecanoic, C17:1 heptadecenoic, C18:3 gamma linolenic, C18:4 stearidonic, C20:2 eicosadienoic, C20:3

[10] Published literature for soybean included Codex (2001); Douglas (1996); ILSI (2010); Kakade et al. (1972); Liener (1994); Novak & Haslberger (2000); OECD (2001a); Vaidehi & Kadam (1989).

eicosatrienoic, C20:4 arachidonic, C20:5 eicosapentaenoic, C22:1 erucic, C22:5 docosapentenoic and C22:6 docosohexanoic acids. Results for the remaining 9 fatty acids are given in Table 12 and can be summarised as follows:

- The mean levels of palmitic, linoleic, and linolenic acid were significantly lower in seeds of both the unsprayed and sprayed soybean FG72 (shaded in Table 12) compared with seeds of 'Jack'. The means fell within both the tolerance range and the combined literature range except for the means for palmitic acid in FG72 seed (from both sprayed and unsprayed plants) which were lower than the tolerance range but within the literature range.
- The mean levels of stearic, oleic, arachidic, eicosenoic and behenic acids were significantly higher in seeds of both the unsprayed and sprayed soybean FG72 (shaded in Table 12) compared with seeds of 'Jack'. All means fell within the combined literature range; the means for all but oleic acid in the GM line fell within the tolerance range.
- For lignoceric acid, there was a significant TREAT x SITE interaction but the majority of the individual by-site analyses did not show significant differences between treatments (A vs B and A vs C).

Table 12: Mean (±standard deviation) percentage composition, relative to total fat, of major fatty acids in seed from 'Jack' and FG72.

Analyte	A 'Jack' (%total)	B FG72-unsprayed (%total)	C FG72-sprayed (%total)	Overall treat effect (P-value)	Tolerance range (%total)	Combined literature range (%total)
Palmitic acid (C16:0)	10.06±0.22	9.34±0.17	9.38±0.23	<0.001	9.78 – 11.40	7.00 - 15.8
Stearic acid (C18:0)	4.28±0.16	4.52±0.19	4.51±0.23	<0.001	3.49 – 4.81	2.00 - 5.88
Oleic acid (C18:1)	21.97±1.05	24.65±0.99	24.12±0.90	<0.001	21.1 – 24.10	14.3 – 34.0
Linoleic acid (C18:2)	54.56±0.90	52.65±0.95	53.08±0.82	<0.001	51.5 – 55.4	48 - 60.0
Linolenic acid (C18:3)	8.27±0.50	7.94±0.45	8.01±0.48	<0.001	7.59 – 10.30	2.00 - 10
Arachidic acid (C20:0)	0.312±0.015	0.324±0.017	0.324±0.019	<0.001	0.25 – 0.35	<0.1 – 0.48
Eicosenoic acid (C20:1)	0.161±0.011	0.165±0.019	0.166±0.012	0.003	<0.1 – 0.18	0.14 - 0.35
Behenic acid (C22:0)	0.319±0.009	0.339±0.012	0.327±0.017	0.001	0.25 – 0.35	0.277-0.595
Lignoceric acid (C24:0)	0.113±0.020	0.119±0.022	0.122±0.025	N/A[1]	<0.1 – 0.15	0.15

[1]NA = not applicable because of TREAT x SITE interaction

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. There was no significant difference between the control and soybean FG72 for any of the amino acids (refer to Table 13). For serine there was a significant TREAT x SITE interaction but the majority of the individual by-site analyses did not show significant differences between treatments (A vs B and A vs C).

Table 13: Mean percentage dry weight (dw), relative to total dry weight, of amino acids in seed from 'Jack' and FG72.

Analyte	A 'Jack' (%dw)	B FG72-unsprayed (%dw)	C FG72-sprayed (%dw)	Overall treat effect (P-value)	Tolerance range (%dw)	Combined literature range (%dw)
Alanine	1.68±0.04	1.68±0.04	11.68	0.90	1.55 – 1.78	1.51 - 2.10
Arginine	2.94±0.10	2.97±0.10	2.95±0.10	0.34	2.69 – 3.13	2.17 - 3.40
Aspartate	4.40±0.12	4.38±0.12	4.37±0.13	0.55	4.06 – 4.67	3.81 - 5.12
Cystine	0.58±0.03	0.58±0.02	0.59±0.03	0.47	0.5 – 0.63	0.37 - 0.81
Glutamate	6.75±0.21	6.77±0.23	6.74±0.22	0.81	6.32 – 7.23	5.84 - 8.2
Glycine	1.68±0.04	1.68±0.04	1.68±0.04	0.96	1.53 – 1.76	1.46 - 2.27
Histidine	1.05±0.03	1.05±0.03	1.05±0.03	0.96	0.93 – 1.07	0.84 - 1.22
Isoleucine	1.81±0.05	1.80±0.05	1.79±0.05	0.38	1.62 – 1.96	1.54 – 2.32
Leucine	2.99±0.08	2.99±0.08	2.98±0.08	0.67	2.71 – 3.13	2.2 - 4.0
Lysine	2.48±0.05	2.48±0.06	2.47±0.06	0.94	2.34 – 2.64	1.55 - 2.86
Methionine	0.54±0.02	0.54±0.02	0.54±0.02	0.91	0.5 – 0.58	0.43 - 0.76
Phenylalanine	1.97±0.05	1.98±0.06	1.96±0.06	0.26	1.83 – 2.08	1.60 - 2.39
Proline	1.82±0.07	1.83±0.07	1.82±0.07	0.75	1.71 – 1.94	1.69 - 2.33
Serine	1.97±0.07	1.98±0.06	1.99±0.06	NA[1]	1.77 – 2.13	1.11 - 2.48
Threonine	1.55±0.04	1.54±0.04	1.53±0.04	0.25	1.44 – 1.62	1.14 - 1.89
Tryptophan	0.45±0.03	0.44±0.03	0.44±0.03	0.12	0.39 – 0.54	0.36 - 0.67
Tyrosine	1.40±0.04	1.40±0.04	1.40±0.04	0.58	1.32 – 1.48	0.1 - 1.62
Valine	1.89±0.06	1.88±0.05	1.87±0.06	0.61	1.66 – 2.03	1.50 - 2.44

[1]NA = not applicable because of TREAT x SITE interaction

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 FG72 and the control.

The level of glycitein was below the limit of quantitation. The levels of the remaining isoflavones are given in Table 14 and show the following:

- The mean percentage dry weights of daidzein and daidzin in soybean FG72 were not significantly different from the levels in 'Jack'.
- For genistein, there was a significant TREAT x SITE interaction but the majority of the individual by-site analyses did not show significant differences between treatments (A vs B and A vs C).
- The level of genistin in soybean FG72 seeds (sprayed and unsprayed treatments) was significantly lower than the level in control seeds. The means for both the GM line and the control were all within the tolerance range and the combined literature range.
- The level of glycitin in soybean FG72 seeds (sprayed and unsprayed treatments) was significantly higher than the level in control seeds. The means for both the GM line and the control all exceeded both the tolerance range and the combined literature range.

Table 14 : Mean weight (µg/g dry weight) of isoflavones in soybean FG72 and 'Jack' seed

Analyte	A 'Jack' (µg/g dw)	B FG72-unsprayed (µg/g dw)	C FG72-sprayed (µg/g dw)	Overall treat effect (P-value)	Tolerance range (µg/g dw)	Combined literature range (µg/g dw)
Daidzein (free)	11.0±2.0	10.6±1.4	10.3±1.0	0.15	<LOQ-14.6	trace – 65[2]
Genistein (free)	11.5±2.1	11.2±1.8	10.5±0.8	NA[1]	<LOQ – 20.6	0.5 – 62[2]
Daidzin conjugate	1035±350	1034±356	994±357	0.56	568 - 2530	60 – 2454[3]
Genistin conjugate	1817±482	1682±465	1640±446	<0.001	1130 - 3290	144 – 2837[3]
Glycitin conjugate	365±39	414±43	400±56	<0.001	142 - 315	15.3 – 310[3]

[1]NA = not applicable because of TREAT x SITE interaction

[2]daidzein and genistein literature values compiled from Wang & Murphy (1994), Kim *et al*. (2005a), Kim *et al*. (2005b), Kim *et al*. (2005c).

[3]these figures are taken from ILSI (2010) and represent the total isoflavone (sum of free and conjugated) isomers expressed as aglycon equivalents. Since the free levels are negligible in FG72, the total isoflavone level is essentially equivalent to the conjugate level.

6.3.5 Anti-nutrients

Levels of key anti-nutrients are given in Table 15. No significant differences between means were obtained for lectin, phytic acid and stachyose. For raffinose the mean levels in seeds from FG72 (sprayed and unsprayed) were higher than the mean level in seeds from 'Jack'. For trypsin inhibitor, while the overall P value indicated a significant difference, neither of the P-values for the A vs B and A vs C analyses indicated significance.

Table 15 : Mean levels of anti-nutrients in soybean FG72 and 'Jack' seed.

Analyte	A 'Jack'	B FG72-unsprayed	C FG72-sprayed	Overall treat effect (P-value)	Tolerance range	Combined literature range
Lectin (hemagglutinat. units/mg)	1.74±0.6	1.4±0.5	1.54±0.42	0.05	0.46 – 8.63	0.11 - 129
Phytic acid (%dw)	1.4±0.16	1.37±0.23	1.35±0.23	0.12	0.96 – 1.5	0.634 - 2.74
Raffinose (%dw)	0.36±0.04	0.38±0.05	0.38±0.06	0.03	0.29 – 0.5	0.11 – 1.28
Stachyose (%dw)	2.49±0.24	2.42±0.18	2.50±0.19	0.27	2.23 – 2.96	1.21 – 6.3
Trypsin inhibitor (trypsin inhibitor units/mg)	33.0±6.6	30.1±6.1	33.9±5.7	0.04	23.5 – 60.1	19.6 - 119

6.3.6 Minerals

Levels of 6 minerals were measured and the results are given in Table 16. There were no significant treatment effects on iron and phosphorus levels. For calcium and sodium, mean levels in FG72 seeds (from both sprayed and unsprayed plants) were significantly higher than in seeds of the control but were within the tolerance and literature ranges. For

magnesium, mean levels in FG72 seeds (from both sprayed and unsprayed plants) were significantly lower than in seeds of the control but, again, were within the tolerance and literature ranges. For potassium, there was a significant TREAT x SITE interaction. The majority of the individual by-site analyses for A vs B did not show significant differences between treatments, while for the A vs C analysis there were equal numbers of significant and non-significant results across the sites.

Table 16: Mean values for mineral levels in seed from 'Jack' and FG72.

Analyte	A 'Jack'	B FG72-unsprayed	C FG72-sprayed	Overall treat effect (P-value)	Tolerance range	Combined literature range
Calcium (% dw)	0.282±0.023	0.258±0.024	0.259±0.026	<0.001	0.212 – 0.347	0.12 – 0.34
Iron (mg/kg)	93.3±41.8	82.6±13.3	84.1±18.9	0.12	58.8 - 175	55.4 – 172
Magnesium (% dw)	0.241±0.01	0.226±0.012	0.226±0.01	<0.001	0.197 – 0.263	0.21 – 0.32
Phosphorus (% dw)	0.626±0.053	0.618±0.062	0.620±0.065	0.49	0.499 – 0.651	0.49 – 0.94
Potassium (% dw)	1.93±0.08	1.85±0.08	1.85±0.09	NA[1]	1.84 – 2.11	1.4 – 2.3
Sodium (%dw)	0.012±0.003	0.015±0.007	0.016±0.008	0.01	<LOQ – 0.026	0.002 – 0.02

[1]NA = not applicable because of TREAT x SITE interaction

6.3.7 Vitamins

Levels of eight vitamins were measured (refer to Table 17). The level of total tocopherols was determined by summation. The results showed the following:

- There were no significant treatment effects on Vitamin B2 and folic acid levels.
- For vitamin B1, further analysis of the significant P-value obtained in the ANOVA showed that this was due only to a significant difference between A vs C .
- For γ-tocopherol, further analysis of the significant P-value obtained in the ANOVA showed that this was due only to a significant difference between A vs B.
- For total tocopherols, mean levels in FG72 seed (from both sprayed and unsprayed plants) were significantly higher than levels in the control.
- There was a significant TREAT x SITE interaction for each of vitamin A, vitamin K, α-tocopherol and δ-tocopherol. For vitamin A, the majority of the individual by-site analyses showed that there were significant differences between treatments (A vs B and A vs C) although it should be noted that at five of the trial sites, readings for Treatments A, B and C were all below the LOQ.

  For vitamin K, α-tocopherol and δ-tocopherol, the majority of the individual by-site analyses did not show significant differences between treatments (A vs B and A vs C).

Treatment means for all analytes fell within both the tolerance and combined literature ranges, except for Vitamin A in which the means for all treatments were above the tolerance range (which was below the LOQ) but were within the range reported in the scientific literature.

Table 17: Mean weight (µg/g dry weight) of vitamins in seed from 'Jack' and FG72

Analyte	A 'Jack' (µg/g dw)	B FG72-unsprayed (µg/g dw)	C FG72-sprayed (µg/g dw)	Overall treat effect (P-value)	Tolerance range (µg/g dw)	Combined literature range (µg/g dw)
Vit B1 (thiamin)	3.59±0.76	3.44±0.95	3.16±0.91	0.009	1.6 – 4.7	1.01 – 16.02
Vit B2 (riboflavin)	4.42±0.88	4.52±0.89	4.80±0.84	0.253	3.36 – 6.38	1.90 – 14.5
Folic acid	2.98±0.35	3.07±0.3	3.12±0.34	0.117	2.19 – 4.33	2.39 – 4.71
Vitamin A	0.217±0.047	0.261±0.112	0.284±0.117	NA[1]	<LOQ	0.26 – 4.37
Vitamin K	0.191±0.069	0.203±0.078	0.215±0.087	NA[1]	<LOQ – 0.263	0.38 – 0.51
α-tocopherol	17.4±3.9	19.0±5.1	20.7±5.8	NA[1]	12.2 – 24.9	2 - 70
γ-tocopherol	195±16	200±14	198±11	0.038	153 - 237	18 - 461
δ-tocopherol	74.1±7.4	75.2±8.3	74.0±11.1	NA[1]	41.5 – 99.2	31 - 186
Total tocopherols	286±16	294±14	293±13	0.017	225 - 346	120 - 674

[1]NA = not applicable because of TREAT x SITE interaction

6.3.8 Summary of analysis of key components

Statistically significant differences in the analyte levels found in seed of soybean FG72 and 'Jack' are summarised in Table 18. For all analytes the soybean FG72 (both sprayed and unsprayed) means fall within the combined literature range. It is noted that for palmitic acid the levels in seeds from FG72 plants (both sprayed and unsprayed) are lower than the lowest level in the tolerance range.and that for oleic acid the levels in seeds from FG72 plants (both sprayed and unsprayed) are higher than the highest level in the tolerance range. Levels of glycitin in both 'Jack' and FG72 were higher than the highest level in the tolerance range.

These statistically significant differences do not raise safety concerns, given that there are no trends in the results, that for the majority of analytes there were no significant differences, and that where significant differences did occur, the means were within the normal biological range for soybean.

Table 18: Summary of analyte levels found in seed of soybean FG72 that are significantly (P < 0.05) different from those found in seed of 'Jack'.

Analyte	Unit of measure.	A Jack	B FG72-unsprayed	C FG72-sprayed	Soybean FG72 within tolerance range	Soybean FG72 within literature range
Ash	%dw	5.24	5.07	5.06	yes	yes
Palmitic acid (C16:0)	% total fat	10.06	9.34	9.38	no	yes
Stearic acid (C18:0)	% total fat	4.28	4.52	4.51	yes	yes
Oleic acid (C18:1)	% total fat	21.97	24.65	24.12	no	yes
Linoleic acid (C18:2)	% total fat	54.56	52.65	53.08	yes	yes
Linolenic acid (C18:3)	% total fat	8.27	7.94	8.01	yes	yes
Arachidic acid (C20:0)	% total fat	0.312	0.324	0.324	yes	yes
Eicosenoic acid (C20:1)	% total fat	0.161	0.165	0.166	yes	yes

Analyte	Unit of measure.	A Jack	B FG72-unsprayed	C FG72-sprayed	Soybean FG72 within tolerance range	Soybean FG72 within literature range
Behenic acid (C22:0)	% total fat	0.319	0.339	0.327	yes	yes
Genistin	µg/g dw	1817	1682	1640	yes	yes
Glycitin	µg/g dw	365	414	400	no ('Jack' also outside range)	No ('Jack also outside range)
Raffinose	%dw	0.36	0.38	0.38	yes	yes
Calcium	%dw	0.282	0.258	0.259	yes	yes
Magnesium	%dw	0.241	0.226	0.226	yes	yes
Sodium	%dw	0.012	0.015	0.016	yes	yes
Vitamin B1	µg/g dw	3.59	3.44	3.16	yes	yes
Vitamin A	µg/g dw	0.217	0.261	0.284	no	yes ('Jack outside range)
γ-tocopherol	µg/g dw	195	200	198	yes	yes
Total tocopherols	µg/g dw	286	294	293	yes	yes

6.4 Compositional data for processed commodities

Studies submitted:

Kowite, W.J. (2009). Production of processed commodities from transgenic event FG72 soybenas and the non-transgenic counterpart (2009). Study ID# M-357014-01-1, Bayer CropScience (unpublished).

Oberdörfer, R. (2009). Nutritional impact assessment report for the double-herbicide-tolerant soybean (transformation event FG72). Study ID# M357280-01-1, Bayer CropScience (unpublished).

The Applicant obtained comparative compositional data for a range of soybean processed commodities (hulls, meal, toasted meal, protein isolate, crude oil, refined, bleached and deodorised (RBD) oil and crude lecithin) derived from line FG72 (both sprayed with isoxaflutole and glyphosate and unsprayed) and 'Jack' grown in a single trial in Adel (Iowa) in 2008. For crude lecithin, the analytes measured were L-α-phosphatidic acid, L-α-phosphatidyl choline, L-α-phosphatidylethanolamine and L-α-phosphatidylinositol. No statistical analyses of the compositional results were done because of the low sample number. The means obtained for each analyte were compared to a combined literature range.

FSANZ has considered the results from these analyses as presented in the Oberdörfer (2009) study but the data are not presented in this Assessment. Of relevance to human food consumption are the results for protein isolate, RBD oil and lecithin:

- For protein isolate, the means across the three treatments, for each analyte measured, showed little variation. In three cases (alanine, proline and trypsin inhibitor) the means for the three treatments were all below the literature range. The level of lectin in protein isolate derived from 'Jack' was over twice the level of that in FG72 (from both sprayed and unsprayed treatments) and was also approximately twice the level reported in the literature.
- The mean level of oleic acid in oil from FG72 plants (both sprayed and unsprayed) appeared to be approximately 15% higher than oil from 'Jack' while the level of linoleic acid was approximately 4% lower in oil from FG72. Otherwise, the fatty acid profile in RBD oil obtained from the three treatments was very similar. Vitamin K was approximately 30% higher in the FG72 oil than in oil from 'Jack' and the means for the two FG72 treatments both exceeded the literature range. While the level of α-tocopherol in oil from unsprayed FG72 was approximately 15% higher than that from

both ‘Jack’ and sprayed FG72, the total tocopherol levels in oil from the three treatments was similar.

- The phospholipid profile of the three crude lecithin samples was similar. Phosphatidylcholine appeared to be approximately 15% higher in lecithin from FG72 (both sprayed and unsprayed). The mean level of phosphatidylinositol in lecithin from sprayed FG72 plants was nearly twice the level in lecithin from ‘Jack’ and unsprayed FG72.

Overall, the levels of analytes in protein isolate, RBD oil and lecithin were similar across the three treatments.

6.5 Assessment of endogenous allergenic potential

Soybean naturally contains allergenic proteins and is one of a group of known allergenic foods including milk, eggs, fish, shellfish, wheat, peanuts, tree nuts and sesame. This group of foods accounts for approximately 90% of all food allergies (Metcalfe *et al*., 1996). The presence of allergenic proteins in the diet of hypersensitive individuals can cause severe adverse reactions. The allergenic effect of soybeans is attributed to the globulin fraction of soybean proteins that comprise about 85% of total protein (OECD, 2001). Soybean-allergic individuals will also be allergic to soybean FG72.

Since soybean is associated with allergic effects in susceptible individuals, a study was done to assess whether seed from soybean line FG72 may have an endogenous allergen content that was different from the non-GM parent line.

Study submitted:

Rouquie, D. (2009). FG72 soybean 2-dimensional gel electrophoresis analysis of soybean endogenous food allergens. ID# M-357490-01-1, Bayer CropScience (unpublished).

Protein extracts were prepared from ground seeds of soybean line FG72, the non-GM parent ‘Jack’, and three commercial non-GM lines (Stine® 2686-6, Stine® 2788 and Stine® 3000-0). The main comparison was between FG72 and ‘Jack’ and the three commercial lines were used to give an indication of biological variability. Two dimensional polyacrylamide gel electrophoresis (2D-PAGE) was used to separate the proteins. The protein profiles were compared, after Coomassie blue staining, using PDQuest image analysis. A total of 37 allergen proteins representing the five known soybean food allergen families (glycinin, Kunitz trypsin inhibitor, β-conglycinin, Bd 28K and Bd 30K) as identified in the AllergenOnline database (www.allergenonline.com) were analysed on the 2D-PAGE maps. A quantitative analysis (spot intensity) for each protein, generated means and standard deviations and allowed statistical comparison between FG72 and ‘Jack’ using a t-test.

All 37 allergens were detected in all the gels generated. For only three of these allergens were statistical differences found between FG72 and ‘Jack’ (refer to Table 19). The mean quantities of these three allergens were smaller in the FG72 seeds than in the ‘Jack’ seeds. The means for the two Kunitz trypsin inhibitors in FG72 fell within the range of biological variability observed in the three non-GM commercial lines. The mean for GLCA 17 was slightly below the range observed in the commercial lines.

Table 19: Summary of differences in allergens found in FG72 and 'Jack'

Allergen ID	Allergen family	No. of allergens tested in the family	Amount by which FG72 differed from 'Jack'
KTI3_3	Kunitz trypsin inhibitor	7	-43%
KTI (1,2,3)_5	Kunitz trypsin inhibitor	7	-29%
GLCA_17	β-conglycinin	4	-21%

Overall, these results suggest that protein from seed of soybean FG72 and 'Jack' is similar in protein profile and content. Thus, soybean line FG72 appears to be equivalent to the non-transgenic counterpart in terms of its endogenous allergen content.

6.6 Conclusion

Detailed compositional analyses were done to establish the nutritional adequacy of seed-derived products from soybean line FG72 under both herbicide sprayed and unsprayed conditions. Analyses were done of 77 analytes encompassing proximates (crude fat/protein, carbohydrate and ash), ADF, NDF, fatty acids, amino acids, isoflavones, anti-nutrients, minerals, and vitamins. The levels were compared to levels in the seeds of the non-GM parent 'Jack'.

These analyses indicated that the seeds of soybean line FG72 are compositionally equivalent to those of the parental line. Out of all the analytes tested, there were significant differences between the non-GM control and soybean FG72 in only 19 analytes. In all of these, the mean levels observed in seeds of soybean FG72 were within the range of natural variation reported in the literature. There were no consistent trends in the effect that herbicide spraying of soybean FG72 had on mean analyte levels.

Mean levels of a range of analytes were also obtained for processed products derived from soybean. There were no meaningful differences between the control and the GM line for any analyte measured in processed products used for human consumption.

In addition, no difference between seeds of soybean line FG72 and 'Jack' were found, in terms of presence and mean level of endogenous allergens.

The compositional data are consistent with the conclusion that there are no biologically significant differences in the levels of key components in seed and processed products derived from soybean line FG72 when compared with the non-GM control or with the range of levels found in non-GM commercial soybean cultivars.

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, the evidence to date indicates that feeding studies using target livestock species

will add little to the safety assessment and generally are not warranted (OECD, 2003; EFSA, 2008).

In this case soybean line FG72 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 FG72, indicate it is equivalent in composition to conventional soybean cultivars. The introduction of food from soybean line FG72 into the food supply is therefore expected to have little nutritional impact.

However, the Applicant submitted the results of a feeding study with toasted soybean meal in broiler chickens. This has been evaluated by FSANZ as additional supporting information.

7.1 Feeding study

Study submitted:

Stafford, J. (2009). Broiler chicken feeding study with FG72 soybeans. ID# M-358025-01-1, Bayer CropScience (unpublished).

This 42-day study compared growth performance and carcass yield of Ross #308 broiler chickens (*Gallus gallus domesticus*) fed diets containing 20% toasted soybean seedmeal from seeds of the T_8 generation of soybean line FG72 (Group C) with those fed diets obtained from seedmeal from two non-GM soybean lines ('Jack' (Group A) plus one commercial cultivar, Stine 3000-0 (Group D)).

Broilers were housed 10 broilers per pen in 14 replicate pens (7 male pens and 7 female pens) to give 140 broilers in each of the three soybean seedmeal treatments.

Diets using seedmeal from the three sources were formulated to meet nutrient requirements of a typical commercial broiler diet and were isoenergetic, isoproteic and as similar as possible relative to limiting amino acids. Both the seedmeal and the formulated diets were sampled for compositional and DNA analyses. The diets were fed in three phases according to standard commercial poultry farming practice - starter (days 0-7), grower (days 8-21) and finisher (days 22-42). Feed and drinking water were available *ad libitum* throughout the study.

Birds were observed daily for overall health, behaviour and/or evidence of toxicity. Body weights were taken on days 7, 21, 35 and 42 and feed consumption was determined every seven days. At study termination, all surviving birds were processed to collect carcass and carcass part yield data. Market dressed carcass, muscle (thigh, breast, wing, leg) and abdominal fat pad weights were collected from 126 individual birds (21 birds/gender/diet).

A total of 77 birds showed clinical symptoms during the feeding study (23 in Group A; 22 in Group C and 32 in Group D). None of these was considered to be treatment related. A total of 45 birds died during the study but again, there was no relation to treatment.

While transient significant differences were noted for some traits, no statistically significant differences attributable to diet were observed in growth performance (body weight gain, mortality, feed efficiency) or carcass yield (breast, thigh, leg , wing, and abdominal fat). Since no adverse effects of feeding FG72 seedmeal in the diet were noted, it is concluded that seeds from soybean line FG72 are nutritionally equivalent to those from both the non-GM control soybean with a comparable genetic background, and a non-GM commercial cultivar.

References[11]

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.

Awata, H., Endo, F. and Matsuda, I. (1994) Structure of the human 4-hydroxyphenylpyruvic acid dioxygenase gene (HPD). *Genomics* 23(3):534-539.

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.

Beltrán, E., Fenet, H., Cooper, F. and Coste, C.-M. (2003) Fate of isoxaflutole in soil under controlled conditions. *Journal of Agricultural and Food Chemistry* 51:146-151.

Boudec, P., Rodgers, M., Dumas, F., Sailland, A. and Bourdon, H. (2001) Mutated hydroxyphenylpyruvate dioxygenase, DNA sequence and isolation of plants which contain such a gene and which are tolerant to herbicides. (US6245968 B1): http://www.patentlens.net/patentlens/patents.html?patnums=US_6245968&language=en&query=(US6245968B1%20in%20publication_number)&stemming=true&returnTo=structured.html%3Fquery%3DUS6245968B1.

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.

Carrington, J.C. and Freed, D.D. (1990) Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. *Journal of Virology* 64(4):1590-1597.

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.

Chang, T.M., Chuang, Y.C., Su, J.H. and Chang, M.C. (1997) Cloning and sequence analysis of a novel hemolysin gene (vllY) from *Vibrio vulnificus*. *Applied and Environmental Microbiology* 63(10):3851-3857.

Codex (2001) *Codex Standard for Named Vegetable Oils.* Report No. CX-STAN 210 – 1999, Codex Alimentarius. http://www.codexalimentarius.net/web/standard_list.do?lang=en.

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

[11] All website references were current as at 28 October 2011

of protein safety in the context of agricultural biotechnology. *Food and Chemical Toxicology* 46:S71-S97.

Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M. (1982) Nopaline synthase: transcript mapping and DNA sequence. *Journal of Molecular and Applied Genetics* 1(6):561-573.

Douglas, J.S. (1996) Recommended compositional and nutritional parameters to test in soybean. Technical Assessment Services, Washington.

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.

Galvez-Valdivieso, G., Pineda, M. and Aguilar, M. (2010) Functional characterization and expression analysis of p-hydroxyphenylpyruvate dioxygenase from the green alga *Chlamydomonas reinhardtii* (Chlorophyta). *Journal of Phycology* 46(2):297-308.

Gish, W. and States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nature Genetics* 3:266-272, doi:10.1038/ng0393-266.

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.

Hamer, L., Adachi, K., Montenegro-Chamorro, M.V., Tanzer, M.M., Mahanty, S.K., Lo, C., Tarpey, R.W., Skalchunes, A.R., Heiniger, R.W., Frank, S.A., Darveaux, B.A., Lampe, D.J., Slater, T.M., Ramamurthy, L., DeZwaan, T.M., Nelson, G.H., Shuster, J.R., Woessner, J. and Hamer, J.E. (2001) Gene discovery and gene function assignment in filamentous fungi. *Proceedings of the National Academy of Sciences* 98(9):5110-5115.

Harrigan, G.G., Lundry, D., Drury, S., Berman, K., Riordan, S.G., Nemeth, M.A., Ridley, W.P. and Glenn, K.C. (2010) Natural variation in crop composition and the impact of transgenesis. *Nature Biotechnology* 28(5):402-404.

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.

Heoedus, Z.L. and Nayak, U. (1994) Homogentisic acid and structurally related compounds as intermediates in plasma soluble melanin formation and in tissue toxicites. *Archives of Physiology and Biochemistry* 102(3):175-181.

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.

ILSI (2010) *International Life Sciences Institute Crop Composition Database Version 4.0*. http://www.cropcomposition.org/query/index.html.

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.

Jay, J.M., Loessner, M.J. and Golden, D.A. (2005) *Modern food microbiology*. 7th ed, Springer, New York. Food Science Text.

Joshi, C.P., Zhou, H., Huang, X. and Chiang, V.L. (1997) Context sequences of translation initiation codon in plants. *Plant Molecular Biology* 35:993-1001.

Kakade, M.L., Simons, N.R., Liener, I.E. and Lambert, J.W. (1972) Biochemical and nutritional assessment of different varieties of soybeans. *Journal of Agricultural and Food Chemistry* 20(1):87-90.

Kim, J.J., Kim, S.H., Hahn, S.J. and Chung, I.M. (2005a) Changing soybean isoflavone composition and concentrations under two different storage conditions over three years. *Food Research International* 38:435-444.

Kim, S.H., Jung, W.S., Ahn, J.K. and Chung, I.M. (2005b) Analysis of isoflavone concentration and composition in soybean [*Glycine max* (L.)] seeds between the cropping year and storage for 3 years. *European Food Research & Technology* 220:207-214.

Kim, S.-H., Jung, W.-S., Ahn, J.-K., Kim, J.-A. and Chung, I.-M. (2005c) Quantitative analysis of the isoflavone content and biological growth of soybean (*Glycine max* L.) at elevated temperature, CO_2 level and N application. *Journal of the Science of Food and Agriculture* 85:2557-2566.

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.

Klein, T.M., Wolf, E.D., Wu, R. and Sanford, J.C. (1987) High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature* 327:70-73.

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.

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.

Liener, I.E. (1994) Implications of antinutritional components in soybean foods. *Critical Reviews in Food Science and Nutrition* 34(1):31-67.

Matringe, M., Sailland, A., Pelisser, B., Rolland, A. and Zink, O. (2005) *p*-Hydroxyphenylpyruvate dioxygenase inhibitor-resistant plants. *Pest Management Science* 61:269-276.

Messina, M. (2005) Isoflavones. In: Coates, P., Blackman, M.R., Cragg, G., Levine, M., Moss, J., and White, J. eds. *Encyclopedia of Dietary Supplements*. Informa Healthcare.

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

Miletich, J.P. and Broze Jr., G.J. (1990) β Protein C is not glycosylated at asparagine 329. *The Journal of Biological Chemistry* 265:11397-11404.

Munger, P., Bleiholder, H., Hack, H., Hess, M., Stauss, R., van den Bloom, T. and Weber, E. (1997) Phenological growth stages of the soybean plant (*Glycine max* (L.) Merr.). Codification and description according to the general BBCH scale - with figues. *Journal of Agronomy and Crop Science* 179:209-217.

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.

Nickell, C.D., Noel, G.R., Thomas, D.J. and Waller, R. (1990) Registration of 'Jack' soybean. *Crop Science* 30:1365.

Novak, W.K. and Haslberger, A.G. (2000) Substantial equivalence of antinutrients and inherent plant toxins in genetically modified novel foods. *Food and Chemical Toxicology* 38:473-483.

OECD (1997) *Consensus document on information used in the assessment of environmental applications involving pseudomonas*. Series on Harmonization of Regulatory Oversight in Biotechnology No. 6. Report No. OCDE/GD(97)22, Organisation for Economic Co-operation and Development, Paris. http://www.oecd.org/officialdocuments/displaydocumentpdf?cote=ocde/gd(97)22&doclanguage=en.

OECD (2001a) *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 (2001b) *Test No. 420: Acute Oral Toxicity - Fixed Dose Procedure*. OECD Guidelines for the Testing of Chemicals / Section 4: Health Effects. Organisation for Economic Co-operation and Development.

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.

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.

Pallett, K.E., Cramp, S.M., Little, J.P., Veerasekaran, P., Crudace, A.J. and Slater, A.E. (2001) Isoxaflutole: the background to its discovery and the basis of its herbicidal properties. *Pest Management Science* 57:133-142.

Pallett, K.E., Little, J.P., Sheekey, M. and Veerasekaran, P. (1998) The mode of action of isoxaflutole. I. Physiological effects, metabolism and selectivity. *Pesticide Biochemistry and Physiology* 62(2):113-124.

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.

Pattison, R.J. and Amtmann, A. (2009) N-glycan production in the endoplasmic reticulum of plants. *Trends in Plant Science* 14(2):92-99.

Paulsen, I.T., Press, C.M., Ravel, J., Kobayashi, D.Y., Myers, G.S., Mavrodi, D.V., Deboy, R.T., Seshadri, R., Ren, Q., Madupu, R., Dodson, R.J., Durkin, A.S., Brinkac, L.M., Daugherty, S.C., Sullivan, S.A., Rosovitz, M.J., Gwinn, M.L., Zhou, L., Schneider, D.J., Cartinhour, S.W., Nelson, W.C., Weidman, J., Watkins, K., Tran, K., Khouri, H., Pierson, E.A., Pierson III, L.S., Thomashow, L.S. and Loper, J.E. (2005) Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nat.Biotechnol.* 23(7):873-878.

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. (2000) Nα-terminal Acetylation of Eukaryotic Proteins. *Journal of Biological Chemistry* 275(47):36479-36482.

Roche, P.A., Moorehead, T.J. and Hamilton, G.A. (1982) Purification and properties of hog liver 4-hydroxyphenylpyruvate dioxygenase; primary structure and chromosomal localization of the gene. *Archivesof Biochemistry and Biophysics* 216(1):62-73.

Rüetschi, U., Dellsén, A., Sahlin, P., Stenman, G., Rymo, L. and Lindstedt, S. (1993) Human 4-hydroxyphenylpyruvate dioxygenase. Primary structure and chromosomal localization of the gene. *European Journal of Biochemistry* 213(3):1081-1089.

Rupprecht, J.K., Liu, A., Kelly, I. and Allen, R. (2004) Degradation of RPA 202248 [U-14C-phenyl]alpha(-(cyclopropylcarbonyl)-2-(methylsulfonyl)-beta-oxo-4-(trifluromethyl)benzenepropanenitrile), the primary degradation product of isoxaflutole, in an outdoor aquatic microcosm system. *Journal of Environmental Science and Health B* 39(5-6):725-736.

Ryle, M.J. and Hausinger, R.P. (2002) Non-heme iron oxygenases. *Current Opinion in Chemical Biology* 6(2):193-201.

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.

Sillankorva, S., Neubauer, P. and Azeredo, J. (2008) *Pseudomonas fluorescens* biofilms subjected to phage philBB-PF7A. *BMC Biotechnology* 8:epub-doi:10.1186/1472-6750-8-79.

Steinert, M., Flügel, M., Schuppler, M., Helbig, J.H., Supriyono, A., Proksch, P. and Lück, P.C. (2001) The Lly protein is essential for *p*-hydroxyphenylpyruvate dioxygenase activity in *Legionella pneumophila*. *FEMS Microbiology Letters* 203:41-47.

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.

U.S.Pharmacopeia. (1990) *United States Pharmacopeia, Volume 23*. United States Pharmacopeia Convention, Inc, Rockville, MD, p1788-1789.

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.

Vaidehi, M.P. and Kadam, S.S. (1989) Soybeans. In: Salunkhe, D.D. and Kadam, S.S. eds. *Handbook of world food legumes, Volume III*. Chapter 1. CRC Press Inc., Boca Raton, pp. 1-21.

Viviani, F., Little, J.P. and Pallett, K.E. (1998) The mode of action of isoxaflutole II. Characterization of the inhibition of carrot 4-hydroxyphenylpyruvate dioxygenase by the diketonitrile derivative of isoxaflutole. *Pesticide Biochemistry and Physiology* 62(2):125-134.

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.

Wada, G.H., Fellman, J.H., Fujita, T.S. and Roth, E.S. (1975) Purification and properties of avian liver p-hydroxyphenylpyruvate hydroxylase. *Journal of Biological Chemistry* 250:6720-6726.

Wang, H. and Murphy, P.A. (1994) Isoflavone composition of American and Japanese soybeans in Iowa: effect of variety, crop, year and location. *Journal of Agricultural and Food Chemistry* 42:1674-1677.

Xiong, J. (2006) *Essential bioinformatics*. Cambridge University Press, p352 pp.