Bayer CropScience

Application to FSANZ to Vary Food Standard 1.5.2 to Include the Double-Herbicide-Tolerant Soybean Event FG72

Prepared by
Felicity Keiper
Sandra Mackie
Dirk Fredrichsdorf
Nina McCormick

June 2010

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Appendix 17 Capt A. (2009d) 2mEPSPS protein overall amino acid sequence homology search with known toxins and allergens. Unpublished Bayer CropScience Report, November 27, 2008. M-273792-02-1.

Appendix 18 Rouquie D. (2007) 2mEPSPS protein heat stability study. Unpublished Bayer CropScience Report, November 27, 2007. M-293053-02-1.

Appendix 19 Rascle J.B. (2009a) HPPD W366 protein heat stability study. Unpublished Bayer CropScience Report, August 26, 2009. M-354574-01-1.

Appendix 20 Rouquie D. (2006a) 2mEPSPS protein *in vitro* digestibility study in simulated gastric fluid. Unpublished Bayer CropScience Report, August 29, 2006. M-276964-01-1.

Appendix 21 Rouquie D (2006b) 2mEPSPS protein *in vitro* digestibility study in simulated intestinal fluid. Unpublished Bayer CropScience Report, July 28, 2006. M-275371-01-1.

Appendix 22 Rascle J.B. (2009b) HPPD W336 *in vitro* digestibility study in human simulated gastric fluid. Unpublished Bayer CropScience Report, September 15, 2009. M-356196-01-1.

Appendix 23 Rascle J.B. (2009c) HPPD W336 protein *in vitro* digestibility study in human simulated intestinal fluid. Unpublished Bayer CropScience Report, September 15, 2009. M-356198-01-1.

Appendix 24 Rouquie D. (2006c) 2mEPSPS protein acute toxicity by oral galvage in mice. Unpublished Bayer CropScience Report, August 29, 2006. M-276952-01-1.

Appendix 25 Rouquie D. (2008) 2mEPSPS protein acute toxicity by intravenous injection in mice. Unpublished Bayer CropScience Report, February 7, 2008. M-297233-01-1.

Appendix 26 Rascle J.B. (2009d) HPPD W336 acute toxicity by oral gavage in female mice. Unpublished Bayer CropScience Report, November 10, 2009. SA 09201.

Appendix 27 Rascle J.B. (2009e) HPPD W336 acute toxicity by intravenous injection in mice. Unpublished Bayer CropScience Report, October 13, 2009. M-357601-01-1.

Appendix 28 Rattemeyer-Matschurat V. (2009) Analysis of substantial equivalence of double-herbicide-tolerant soybean, by means of T-test differences. Unpublished Bayer CropScience Report. M-356445-03-1.

Appendix 29 Mackie S.J.W. (2009) Composition of seed from FG72 soybean and its non-transgenic counterpart. Unpublished Bayer CropScience Report, July 1, 2009. M-355723-01-1.

Appendix 30 Rouquie D. (2009) FG72 soybean 2-dimensional gel electrophoresis analysis of soybean endogenous food allergens. Unpublished Bayer CropScience Report, October 15, 2009. M-357490-01-1.

Appendix 31 Stafford J. (2009) Broiler chicken feeding study with FG72 soybeans. Unpublished Bayer CropScience Report, October 22, 2009. M-358025-01-1.

Executive Summary

Bayer CropScience Pty Ltd seeks to vary FSANZ Standard 1.5.2 to allow the use of genetically modified soybean (*Glycine max* L. Merr.) derived from transformation event FG72 in the Australian and New Zealand food industries. Five food products are derived from soybean: whole soybeans, oil, meal, hulls and protein. Soybean oil is the primary food product consumed by humans in Australia, with the other products used either as food products or as components of animal feed.

Soybean varieties containing event FG72 comprise the stably integrated gene *2mepsps* and *hppdPf W336.* The incorporation and expression of the FG72 transgenic locus in the *G. max* genome has been characterised according to international standards for the safety assessment of biotechnology products. This information is included with this application to support the food safety of the 2mEPSPS and HPPD W336 proteins. Soybean varieties containing the FG72 event will be grown commercially in the major soybean producing countries of the world.

The *2mepsps* gene was generated by introducing mutations into the wild-type *epsps* (wt *epsps*) gene from *Zea mays*, leading to a modified EPSPS protein with two amino acid substitutions (2mEPSPS). This modification confers decreased binding affinity for glyphosate, thus allowing sufficient enzyme activity in the presence of the herbicide. The EPSPS protein has been widely used to confer glyphosate tolerant properties to crops, and FSANZ has previously assessed the EPSPS protein in transgenic soybean (A338, A592), cotton (A355, A553, A614), corn (A362, A416, A548), canola (A363), wheat (A524), sugarbeet (A378, A525) and lucerne (A575) expressed by other transgenic crops.

The wild-type (wt) *hppd* gene was isolated and cloned from *Pseudomonas fluorescens*. A single amino acid substitution introduced in the wt *hppd* gene resulted in the modified *hppdPf W336* gene. The expressed protein, HPPD W336 possesses greater than 99.5% homology to the native HPPD protein from *P. fluorescens* and confers tolerance to isoxazole herbicides.

The *2mepsps* and *hppdPf W336* genes were introduced into the *G. max* genome in a single gene construct via direct-gene transfer. The regulatory sequences used in this construct are derived from common plants or plant pathogens that are routinely used in plant biotechnology and have a history of safe use.

In the molecular characterisation of the FG72 transgenic locus, bioinformatics analysis of the full DNA sequence revealed no evidence supporting cryptic gene expression or unintended effects resulting from the genetic modification. The transgenic locus also shows structural stability over different generations and growing environments, and in different genetic backgrounds.

Food safety evaluation of the 2mEPSPS and HPPD W336 proteins was undertaken utilising guidance provided by Codex (2003). No health-related adverse effects have been associated with the proteins. The source organism for the 2mEPSPS protein, maize, is a safe crop plant widely used for food and feed. EPSPS proteins are ubiquitous in nature, being widely expressed in food and feed crops. The 2mEPSPS protein has no amino acid sequence homology to know allergens and is rapidly degraded in simulated gastric fluid and simulated intestinal fluid assays. The 2mEPSPS protein has no amino acid sequence similarity to know toxins and exhibited no effects in an acute oral mouse toxicity test.

The HPPD protein source organism, *Pseudomonas fluorescens*, is a non-pathogenic bacterium which is ubiquitous in nature and has a good history of safe use. The HPPD W336 protein showed no amino acid sequence similarities to know allergens or toxins. The protein was rapidly degraded in simulated gastric and simulated intestinal fluid assays and showed no toxicity in an acute oral mouse test.

Analysis of novel protein content revealed detectable quantities of the 2mEPSPS and HPPD W336 proteins in all processed fractions derived from FG72 soybeans, but no protein content in refined oil, the primary soybean food product consumed by humans in Australia. Levels of 2mEPSPS and HPPD W336 proteins were measured in event FG72 soybean tissues and exposure levels to humans and animals of both proteins are expotentially lower than the doses tested in the acute oral mouse study. In the nutritional and compositional analyses of commodities derived from soybean containing the FG72 event, no differences of biological relevance were identified between transgenic and non-

transgenic soybean. Therefore, it is concluded that the FG72 event has negligible impact on soybean nutritional value and furthermore, that no significant adverse effects to animal or human health will result based on the food and feed safety assessment.

Part 1 General Information on the Application

1.1 Applicant Details

(a) Applicant's name

Bayer CropScience Pty Ltd
Contact person: Dr Nina McCormick

(b) Company/organisation name

Bayer CropScience Pty Ltd

(c) Address (street and postal)

Bayer CropScience Pty Ltd
391-393 Tooronga Road
Hawthorn East
Victoria 3123
Australia

(d) Telephone and facsimile numbers

Tel: +61 3 9248 6832
Fax: +61 3 9248 6605

(e) Email address

nina.mccormick@bayercropscience.com

(f) Nature of applicant's business

Plant biotechnology, plant breeding, seed and trait research and development.

(g) Details of other individuals, companies or organisations associated with the application.

On June 3, 2002, Bayer CropScience was formed by the acquisition of Aventis CropScience by Bayer AG. From that date, Bayer CropScience is the agricultural business unit of Bayer that is engaged in the research, development, and marketing of crop protection, seed technology, turf and ornamentals, professional pest and vector control, and home and garden products.

On December 15, 1999, Aventis S.A. was formed by the completion of the merger between Hoechst AG and Rhône-Poulenc S.A. Aventis CropScience was formed as part of a worldwide merger between Rhone-Poulenc S.A. and Hoechst AG. A portion of that merger created Aventis CropScience Holding S.A. that included interests from Hoechst AG and Schering AG. Hoechst AG and Schering AG were the parent companies of AgrEvo USA Company which were all merged into the Aventis companies.

Some of the activities described in the technical information supporting this application were undertaken before the merger and acquisition. Consequently, the names Aventis CropScience, AgrEvo USA Company, AgrEvo, and Hoechst Schering AgrEvo GmbH may appear in reports submitted to support this application.

M.S. Technologies, LLC, is an Iowa limited liability company, with offices at 103 Avenue D, West Point, Iowa 52656, U.S.A. The FG72 transformation event is owned by M.S. Technologies, LLC.

In November of 2007, M.S. Technologies, LLC and Bayer CropScience AG entered into an agreement for the joint development of herbicide tolerant soybeans, including the FG72 transformation event.

Some of the activities described in this application were undertaken in the context of the agreement between Bayer CropScience AG and MS Technologies, LLC, for example some of the field activities were conducted by MS Technologies that are contained in reports submitted to support this application.

Please note that all inquiries regarding this report and the data contained herein should be addressed to Dr Nina McCormick, as described in the applicant details in (a)-(e) above.

1.2 Purpose of the Application

This application, on behalf of Bayer CropScience Pty Ltd and MS Technologies LLC, seeks to vary FSANZ Standard 1.5.2 to allow the use of genetically modified soybeans (*Glycine max* L.) derived from event FG72 in the Australian and New Zealand food industries.

Soybeans are cultivated for the production of seed which have a multitude of food, feed and industrial uses. Soybeans are one the major sources of vegetable oil in the human diet. Soybeans are also a source of high protein meal for livestock.

Soybeans are cultivated primarily in the United States, Argentina, Brazil, China and India. Soybean varieties containing the FG72 event will be commercially cultivated in some of these countries. It is therefore anticipated that food products derived from soybean varieties containing the FG72 event will enter the Australian and New Zealand food supply via imports from countries of production.

1.3 Justification for the Application

The FG72 transformation event introduced two genes to the *G. max* genome. These genes confer two novel traits: tolerance to broad spectrum herbicides with glyphosate and isoxazole as the active ingredients. Soybean varieties containing the FG72 event will be produced commercially in the major soybean producing countries of the world.

<u>Advantages of FG72 soybeans</u>
The novel traits expressed by soybean varieties containing the FG72 event provide several agronomic benefits over conventional soybean varieties and other transgenic soybeans. These include:

- Glyphosate is a broad spectrum, post-emergence weed control system that provides an alternative to pre-emergent and residually active compounds, and encourages herbicide use on an as-needed basis.
- Isoxazole is a broad spectrum, post-emergence weed control system that provides an alternative to glyphosate and to pre-emergent and residually active compounds, and encourages herbicide use on an as-needed basis
- Broad spectrum weed control reduces cultivation needs.

<u>Food safety</u>
Information is provided in this application to support the safety of the 2mEPSPS and HPPD W336 proteins expressed by the FG72 event.

FSANZ has previsouly evaluated the 2mEPSPS protein, as expressed by Glytol cotton (A614).

Costs and benefits, and impacts on trade

Varying FSANZ Standard 1.5.2 to include commercial soybean varieties containing event FG72 is unlikely to have a detrimental impact on the Australian soybean industry. Despite being a small soybean producer, Australian soybean is sourced for food and feed products on the domestic market and also, culinary quality soybeans produced out of season are exported to the main northern hemisphere producers. Soybean food and feed ingredients are also obtained from imported soybean products, with the US a major source of imports. Once soybean varieties containing the FG72 event are launched for commercial production in the US as well as other parts of the world, food and feed products derived from soybean containing this event are likely to enter the domestic food and feed supply.

If the soybean event FG72 is not incorporated into the FSANZ Standards, this could have wide ranging impacts on the price of food and feed products containing the ingredients derived from soybean. These would arise from the need to source varieties that do not contain the FG72 event. These products may attract a premium price that must be met by the manufacturer, with those costs eventually passed on to the consumer. This would be compounded by the costs of segregating FG72 soybean products, where trading partners are willing to comply with this requirement. Other factors to consider include disruptions to the food supply, and the significant costs of recalling food products if the FG72 event were to be distributed in the local food and feed supply.

Varying the FSANZ Standards to include FG72 will contribute to maintaining stable food prices, consumer choice in the marketplace, and decreased production costs for transgenic soybean varieties in the longer term. The potential trade implications of not including soybean event FG72 in the FSANZ Standards are significant. Segregating FG72 soybean products from other soybean products has compliance and identification requirements that are difficult and costly to meet. The US is the major trading partner of Australia, and approved transgenic crops are considered to be substantially equivalent to conventional crops. Therefore, in the US, there are no intentions of segregating or labelling transgenic crops or their products. Products containing event FG72 imported into Australia from the US, or other tranding partners with similar treatments of transgenic crops, may need to be removed from sale. This could expose Australia to disputes with trading partners at the World Trade Organisation.

1.4 Assessment Procedure

We consider that the appropriate assessment for this application is the General Procedure. Event FG72 expresses two novel proteins as a result of genetic modification. The 2mEPSPS, and other similar EPSPS proteins, have been assessed for their safety by FSANZ previously. The HPPD W336 protein expressed by the FG72 event has not been previously assessed.

1.5 Confidential Commercial Information

Parts of the Bayer CropScience reports provided in Appendices 2 to 6 contain confidential commercial information. A formal request for this information to be treated as such has been submitted to FSANZ.

1.6 Exclusive Capturable Commercial Benefit (ECCB)

The application is expected to confer an ECCB upon Bayer CropScience since it will contribute to facilitating commercial activities with soybean varieties containing the FG72 event in Australia.

1.7 International and Other Standards

The Bayer CropScience reports and studies included in the information supporting this application have been conducted according to international standards. In the safety assessment of biotechnology products, Bayer CropScience refers primarily to the *Codex Alimentarius* Commission weight-of-evidence approach (CAC, 2003), and the relevant Codex Standard is:

Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants. CAC/GL 45-2003. Adopted in 2003, Annexes II and III adopted in 2008.

Other guidelines and recommendations are also considered including those of the World Health Organisation (WHO), the United Nations Food and Agriculture Organisation (FAO), the Organisation for Economic Cooperation and Development (OECD), the United States Food and Drug Administration (US-FDA), the United States Environment Protection Agency (US-EPA), and the European Food Safety Agency (EFSA) (see CAC, 2003; EFSA, 2006; FAO/WHO, 2001; OECD, 2001a, 2001b; US-EPA, 2002; US-FDA, 2006).

1.8 Statutory Declaration

Included in the application cover letter to FSANZ.

1.9 Checklist for Standards Related to New Foods

APPLICATION REQUIREMENT CHECKLIST	SECTION IN THIS APPLICATION	PAGE NUMBER
General Requirements **(Application Handbook section 3.1)**		
Form of application		
Applicant details	1.1	13
Purpose of the application	1.2	14
Justification of the application	1.3	14
Information to support the application	Parts 2, 3 and 4	18 - 92
Assessment procedure	1.4	15
Confidential Commercial Information	1.5	15
Exclusive Capturable Commercial Benefits	1.6	15
International standards	1.7	16
Statutory Declaration	1.8	See application cover letter
Foods Produced Using Gene Technology **(Application Handbook section 3.5.1)**		
Nature and identity of the GM food	2.1	18
History of use of host and donor organisms	2.2	19
Nature of genetic modification	2.3	32
Labelling information on GM food	2.4	59
Antibiotic resistance marker genes (of used)	3.1	62
Characterisation of novel protein(s)/substances	3.2	62
Toxicity of novel protein(s)/substances	3.3	70
Potential allergenicity of novel protein(s)	3.4	74
Compositional analysis of GM food	3.5	76
Nutritional impact of GM food	4.1	91
Animal feeding studies (if available)	4.2	91

Part 2 Technical Information on the Genetically Modified Food

2.1 Nature and Identity of the Genetically Modified Food

(a) A description of the GM organism from which the new GM food is derived. The description must include the nature and purpose of the genetic modification.

The GM organism is soybean (*Glycine max*) transformed with event FG72. Transformation of the soybean variety "Jack" (Nickell *et al*., 1990) was achieved using direct gene transfer transformation methodology (see Section 2.3(a)).

Event FG72 introduced two novel genes to the *G. max* genome:

(i) The double mutant 5-enol pyruvylshikimate-3-phosphate synthase (*2mepsps*) gene that encodes for the 2mEPSPS protein. The *2mepsps* coding sequence was developed by introducing two point mutations to the wild-type *epsps* gene cloned from maize (*Zea mays*). Expression of the 2mEPSPS protein confers tolerance to glyphosate herbicides. The *2mepsps* gene has been used to confer glyphosate tolerant properties to crops including maize, cotton, canola and soybean (Herouet-Guicheney *et al*., 2009). FSANZ has previously assessed the 2mEPSPS protein, as expressed by the *2mepsps* gene, in the Bayer CropScience application for GlyTol cotton (A614).

(ii) The *hppdPf W336* gene encodes for the HPPD W336 protein. The wild type p-hydroxyphenylpyruvate dioxygenase (HPPD) protein is found throughout nature: it is produced by bacteria, fungi, plants and animals including mammals. The HPPD W336 protein was developed by introducing a single amino acid change in the wild type HPPD protein produced by *Pseudomonas fluorescens*. Over production of the HPPD W336 protein in plants confers tolerance to isoxazole herbicides (such as isoxaflutole; IFT).

(b) The name, number or other identifier of each of the new lines or strains of GM organism from which the food is derived.

The double-herbicide-tolerant soybean event is named "FG72", and soybeans transformed with this event are known as FG72 lines before they are introgressed into elite germplasm. The OECD identification number for event FG72 is: MST-FGØ72-3.

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

This is unknown as this application is related to a commodity crop rather than a specific food or additive.

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

Soybean is cultivated primarily for the production of seed that has many food, feed and industrial uses. In the human diet, soybeans are one of the major sources of edible vegetable oil that is used as a purified oil, or utilized in margarines, shortenings and cooking and salad oils. Soybeans may also be consumed directly without any processing as soybean seeds, and many non-fermented and fermented oriental soybean foods such as soy sprouts, milk, tofu, tempeh, miso, natto, and soy sauce. Edible soy protein products including grits and flours, concentrates and isolates are used as food ingredients. The different soy protein products are added in bakery products, snack foods, noodle products and comminuted meat products (Hui, 1992; CRC, 1983). Meal derived from soybeans is used as a high protein supplement in feed rations for livestock. Industrial uses of soybeans range from the production of yeasts and antibodies to the manufacture of soaps and disinfectants.

2.2 History and Use of the Host and Donor Organisms

(a) A description of all the donor organism(s) from which the genetic elements are derived, including:

(i) Common and scientific names and taxonomic classification;

The taxonomic classifications of the organisms from which the genetic elements of event FG72 are derived are presented below in Table 1.

Glycine max
The host species, *G. max* belongs to the subgenus *Soja*, which also contains *G. soja* and *G. gracilis*. *Glycine soja* is the wild species of soybean and is endemic in many Asian countries. Cytological, morphological and molecular evidence suggest that *G. soja* is the ancestor of *G. max*. *Glycine gracilis* is considered to be a weedy or semi-wild form of *G. max*, with some phenotypic characteristics intermediate to those of *G. max* and *G. soja*. *Glycine gracilis* may be an intermediate in the speciation of *G. max* from *G. soja* or a hybrid between *G. soja* and *G. max* (OECD, 2000).

2mepsps gene
The coding sequence of the EPSPS protein was originally isolated from *Zea mays*, specifically "Black Mexican Sweet", an old commercial sweet maize variety. Maize is one of the few major crops indigenous to the Western Hemisphere and is grown in nearly all areas of the world over a wide range of climatic conditions (Hallauer *et al*., 1988). Black Mexican Sweet maize is a cultivar of New England sweet maize first introduced in 1864, most likely in the US state of New York. Maize is categorized as a vegetable and used mainly for human consumption directly, with no processing.

hppdPf W336 gene
The coding sequence of the of the 4-hydroxyphenylpyruvate dioxygenase (HPPD) protein was isolated from the *Pseudomonas fluorescens* strain A32 (Genebank A69533; McKellar, 1982) via PCR amplification. The PCR approach was based on the amino acid sequence of the HPPD protein present in *Pseudomonas fluorescens* strain P.J. 874. The resulting DNA sequence was modified to produce the HPPD W336 protein with enhanced tolerance against HPPD inhibitors.

Pseudomonas fluorescens Migula 1895 (type strain ATCC 13525; taxonomy ID: 136843; Skerman, 1980)*, Pseudomonas putida* and *Pseudomonas chlororaphis* are closely related to each other and are seen as forming a complex within the fluorescent subgroup of the *Pseudomonas* genus. In addition, *P. fluorescens* is a heterogeneous species comprising several biovars, each of which may deserve species rank, but which are so interconnected that adequate methods have not been devised to clearly separate them (OECD, 1997). *Pseudomonas fluorescens* are Gram-negative, rod-shaped, motile, asporogenous, aerobic bacteria that produce fluorescent pigments and are catalase and oxidase-positive. *Pseudomonas fluorescens* strains are generally not able to grow above 42°C, but grow at 5°C (OECD, 1997; Palleroni, 1981). This organism is a nonpathogenic saprophyte which inhabits soil, water and plant surface environments. It is able to produce a soluble, greenish fluorescent pigment, which relates to its name.

Regulatory Sequences
In the *2mepsps* gene expression cassettes, the *2mepsps* gene coding sequence is under the control of the H4 promoter of *Arapidopsis thaliana* (Ph4A748; Chabouté *et al*., 1987), followed by the first intron of gene II of the histone H3.III variant of *Arabidopsis thaliana* (Chaubet *et al*., 1992) and by the optimized transit peptide as described by Lebrun *et al*. (1996), and terminated by the 3′ untranslated region of the histone H4 gene of *Arabidopsis thaliana* (Chabouté *et al.*, 1987).

The *hppdPf W336* gene coding sequence is under the control of the duplicated H4 promoter of *Arabidopsis thaliana* (Ph4A748; Chabouté *et al*., 1987), followed by the enhancer sequence of the Tobacco etch virus (Carrington and Freed, 1990), and by the optimized transit peptide as described by Lebrun *et al.* (1996), and terminated by the 3′ untranslated region of the nopaline synthase from *Agrobacterium tumefaciens*.

The plant species *Arabidopsis thaliana,* common name mouse-ear cress, is member of the family Brassicaceae. *Zea mays*, more commonly know as maize or corn, is a member of the Poaceae family.

Helianthus annuus, or sunflower, is a member of the family Asteraceae. These plant species are not considered to cause disease in humans, plants or animals.

Tobacco etch virus is a plant pathogenic virus of the family Potyviridae. The virus comprises single-stranded positive sense RNA and has a host range primarily of tomato, pepper, tobacco and various Solanaceous weeds. It is not know to cause disease in animals or humans.

The nopaline synthase gene (3′ nos) is derived from the T-DNA of the tumour-inducing plasmid, pTiT37, of *Agrobacterium tumefaciens* (Depicker *et al*., 1982). *Agrobacterium tumefaciens* is a soil born, gram-negative bacterium that has been extensively studied since it was identified as the causative agent of crown gall disease in plants. *Agrobacterium tumefaciens* and *A. rhizogenes* are two well know prokaryotic organisms capable of transferring DNA to the eukaryotic cell (De Groot *et al*., 1998). This gene transfer ability may have evolved from bacterial conjugal transfer systems which mobilise plasmids for transfer between bacterial cells (Stachel and Zambryski, 1986) and is exploited in biotechnology. Consequently, *A. tumefaciens* is a widely used transformation system in plant biotechnology

Table 1 Taxonomy of the donor organisms from which the genetic elements of the FG72 event are derived

GENETIC ELEMENT	DONOR ORGANISM TAXONOMY							
	Kingdom	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Scientific Name**	**Common Name**
Plant Genome								
Genomic DNA	Plantae	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Glycine	*Glycine max* (L.) Merr.	soy bean
Gene Construct								
3'nos								
hppdPf W336	Bacteria	Proteobacteria	*Gammaproteobacteria*	Pseudomonadales	Pseudomonadaceae	Pseudomonas	*Pseudomonas fluorescens*	Strain ATCC 13525; taxonomy ID 136843[A]
TPotp Y	Plantae	Anthophyta	Liliopsida	Poales	Poaceae	Zea	*Zea mays*	corn
	Plantae	Magnoliophyta	Magnoliopsida	Asterales	Asteraceae	Helianthus	*Helianthus annuus*	sunflower
5'tev	Viruses	-	-	-	Potyviridae	Potyvirus	Tobacco etch virus	
Ph4a748 ABBC	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress
Ph4a748	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress
intronl h3At	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress
TPotp C	Plantae	Anthophyta	Liliopsida	Poales	Poaceae	Zea	*Zea mays*	corn
	Plantae	Tracheophyta	Magnoliopsida	Asterales	Asteraceae	Helianthus	*Helianthus annuus*	sunflower
2mepsps	Plantae	Anthophyta	Liliopsida	Poales	Poaceae	Zea	*Zea mays*	corn
3'histonAt	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress

[A] Skerman, 1980. Approved lists of bacterial names. *Int. J. Syst. Bacteriol.* 30:225-420. Available at http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=bacname (accessed on March 03, 2009).

(ii) Information about any known pathogenicity, toxicity or allergenicity of relevance to the food; and

Soybean

Anti-nutrients in Soybean
There are a several compounds in legumes, and therefore also in soybeans, which are not favourable for human or animal nutrition. These anti-nutritional factors include phytic acid, raffinose and stachyose, protease inhibitors, and hemagglutinins (lectins).

Phytic acid
In most plant tissues, large portions of phosphorus are present in form of phytic acid (1,2,3,4,5,6-hexakis (dihydrogen phosphate) myo-inositol). Phytic acid is regarded as the primary storage form of phosphorus and inositol in almost all seeds. During seed germination, phytin, the calcium-magnesium salt of phytic acid, is hydrolysed by the enzyme phytase and serves as a source of inorganic phosphorus and cations for the emerging seedling. The term phytate is used for the mono to dodeca anion of phytic acid (Ravindran *et al.* 1994; Maga, 1982).

Two-thirds of the phosphorus in soybeans is bound as phytate and unless freed is mostly unavailable to animals (Liener, 1994). Ruminants are able to utilise considerably more phosphorus, since rumen microbes produce phytase that breaks down phytate and releases phosphorus. Phytic acid also chelates mineral nutrients including calcium, magnesium, potassium, iron, and zinc, rendering them unavailable to monogastric animals consuming the beans. In fact, phytic acid chelation of zinc present in corn-soybean meal diets used for growing swine requires supplements of zinc to avoid a parakeratosis (OECD, 2001c). It is becoming common for feed formulators to add a phytic acid degrading enzyme, phytase, to swine and poultry diets to release phytin-bound phosphorus, so that the amount of this mineral added to the diet can be decreased, potentially reducing excess phosphorus in the environment. Phytic acid also impacts on protein bioavailability and enzyme activity since it is a strong anion and it can interfere with the polar side groups of proteins leading to complexion of nutritional proteins or changes in the molecular conformation of enzymes (Fretzdorff and Brümmer, 1992).

Phytic acid contents reported for soybean seeds are 1.0 - 1.5% (Liener, 1994). However, higher values, up to 2.74% have also been reported (Douglas, 1996).

Raffinose and stachyose
The low molecular weight carbohydrates, stachyose and raffinose, are present in defatted toasted soybean meal, as well as in raw soybeans. Raffinose is a trisaccharide containing galactose, glucose and fructose. Stachyose is a tetrasaccharide built of two galactose, one glucose and one fructose molecule. They are considered anti-nutrients, because they remain unhydrolysed in the small intestine of monogastric animals and humans due to a lack of galactosidase and hence are not absorbed. They then pass into the large intestine where microbial fermentation converts them to CO_2, the main components of flatus (Vaidehi and Kadam, 1989). The raffinose content of soybean seeds ranges from 0.1- 0.9 g per 100 g on a fresh weight basis, while stachyose content is from 1.4 - 4.1 g per 100 g. Further processing of soybean meal into concentrate or isolate, reduces or removes, these oligosaccharides.

Protease inhibitors
Protease inhibitors are typical anti-nutritional compounds present in soybeans, cereals and potatoes. Two types of protease inhibitors are present in soybeans: the Kunitz inhibitor and the Bowman-Birk inhibitor. Trypsin inhibitors are proteins with molecular weights between 6 - 46 kDalton, which form inactive complexes with the proteinase trypsin. The Kunitz inhibitor and the Bowman-Birk inhibitors are active against trypsin, while the latter is also active against chymotrypsin (Liener, 1994). These protease inhibitors interfere with the digestion of proteins resulting in decreased animal growth. The activity of these inhibitors is destroyed when the bean or meal is toasted or heated during processing.

Lectins
Lectins are proteins that bind to carbohydrate-containing molecules. Lectins in raw soybeans can inhibit growth and cause death in animals and it is expected that similar effects would occur

in humans (Liener, 1994). The ability of lectins to act as hemagglutinins that cause blood clotting is the basis for most quantitative analytical methods. Soybean lectin is sometimes referred to as soybean hemagglutinin. Lectins are rapidly degraded upon heating but are quite resistant to dry heat.

Isoflavones

Soybeans naturally contain a number of isoflavone compounds reported to possess biochemical activity, including estrogenic, anti-estrogenic, and hypocholesterolemic effects, in mammalian species. These compounds have been implicated in adversely affecting reproduction in animals fed diets containing large amounts of soybean meal (Shutt, 1976). However, it is not universally accepted that isoflavones are antinutrients as they have also been reported to have beneficial anti-carcinogenic effects (Messina and Messina, 1991).

The isoflavones in soybeans and soy products have three basic types: daidzein, genistein, and glycitein. Each of these three isomers, known as aglucones or free forms, can also exist in three conjugate forms: glucoside, acetylglucoside, or malonylglucoside. Therefore, in total there are twelve isomers of isoflavones in soybeans. The isoflavone content of soybeans is greatly influenced by many factors, including variety, growing locations, planting year, planting date and harvesting date. In literature reports on isoflavone contents of soybeans, the specific substances investigated, the analytical methods and the reporting conventions have differed widely from report to report (Douglas, 1996). Isoflavones are heat stable and not destroyed by toasting of soybean meal (Liener, 1994).

Allergies to Soybeans

Several soybean food allergies have been recorded in most countries of the world (Ballmer-Weber and Vieths 2008). Clinical reactions are similar to those observed with other major food allergens (Besler *et al*., 2000). In the absence of epidemiological data, the estimated prevalence of soybean allergies could be 0.5% in the general population (Sicherer and Sampson, 2006; Ballmer-Weber and Vieths 2008). Due to its widespread use in the food and beverage industry, soybean allergens are often hidden ingredients. Therefore, labelling regulations (e.g. from Codex, US Food and Drug Administration, European Union) incorporate soybean as part of the major allergenic food lists that should be labelled (Codex 1999; EU, 2000, 2003; US-FDA, 2004).

Saline extracts of soybeans have been reported to contain several antigenic proteins that stimulate the rabbit systemic immune system after injection and/or orally sensitise guinea pigs, calves, pigs, and humans. The presence of these allergenic proteins in the diet of hypersensitive individuals can cause severe adverse reactions in the gastrointestinal tract. The allergenic effect is attributed to the globulin fraction of soybean proteins that comprise about 85% of total protein. When compared to soybean seeds, sprouts exhibit a similar ability to bind IgE from soy-allergic individuals. A number of immunological or immunochemical tests have been developed to examine allergenic proteins usually based on sera from sensitive subjects (OECD, 2001c).

Many soybean allergenic proteins have been identified, characterized and recorded in multiple allergen databases. AllergenOnline (www.allergenonline.org; update in January 2009), from the Food Allergy Research and Resource Program (FARPP) program, contains the highest number of allergens (see Table 2) and provides a robust resource for searching potential similarities with other proteins. These allergens belong to five major protein families: beta-conglycinin, glycinin, Kunitz trypsin inhibitor, Bd 28K, and Bd 30K. These families have conserved structural features in relation with their biological activity, which explains the wide immunochemical cross-recognition observed among members of the legume family (Ballmer-Weber and Vieths 2008).

Table 2 Food Soybean Allergens

FAMILIES OF PROTEINS	PROTEIN NAMES	GI NUMBERS
Beta-conglycinin (7S-cupin, 7S-globulin, vicilin, Gly m 5)	Beta-conglycinin, alpha chain [Precursor] Beta-conglycinin-alpha subunit Beta-conglycinin storage protein CG4 beta-conglycinin	18536 169927 169929 256427
Bd 30K (Cysteine thiol-protease C1)	34 kDa maturing seed vacuolar thiol protease precursor Gly m Bd 30K P34 probable thiol protease [Precursor]	3097321 1199563 129353
Bd 28K (7S-cupin)	Gly m Bd 28K	12697782
Glycinin G1 (11S-globulin, legumin, Gly m 6)	Glycinin G1 [Precursor] Glycinin G1	18635 18615
Glycinin G2 (11S-globulin, legumin, Gly m 6)	Glycinin G2 [Precursor] Glycinin G2	18637 18609
Glycinin G3 (11S-globulin, legumin, Gly m 6)	Glycinin G3 [Precursor]	18639
Glycinin G4 (11S-globulin, legumin, Gly m 6)	Glycinin G4 [Precursor] Glycinin G4 Glycinin G4 A5A4B3 subunit	732706 18641 806556
Glycinin G5 (11S-globulin, legumin, Gly m 6)	Glycinin G5 [Precursor] Glycinin G5 Gy5 protein	169971 169969 736002
Kunitz trypsin inhibitor (Kunitz-legume)	Kunitz trypsin inhibitor Kti Kunitz trypsin inhibitor KTi1 Kunitz trypsin inhibitor Kti2 Trypsin inhibitor subtype A Trypsin inhibitor subtype B ..... Kunitz trypsin inhibitor	256429 256635 256636 18770 18772 510515

Source: www.allergenonline.org; accessed 23rd of October 2009)

hppdPf W336

Pathogenicity to humans
The gene *hppdPf W336* in event FG72 is derived from *P. fluorescens. P. fluorescens* can be an opportunistic pathogen in immunocompromised patients (McKellar, 1982). Some cases of septicemia have been reported due to *P. fluorescens* contamination of transfused blood and blood products, given its ability to grow at 5°C (Gibb *et al*., 1995, Puckett *et al*., 1992). Some *P. fluorescens* strains were also reported to create biofilms on compounded sterile products like catheters and have led to rare infections in immunocompromised populations (Gershman *et al.,* 2008). However, the general virulence of *P. fluorescens* is low due to its inability to multiply rapidly at body temperature and having to compete with defense mechanisms of the host (Liu, 1964).

Pathogenicity to animals
P. fluorescens can infect a wide range of animals including horses, chickens, marine turtles, and many fish and invertebrate species. However, since it is unable to grow at elevated temperatures, it is probably only an opportunistic pathogen for warm-blooded animals (OECD, 1997).

Pathogenicity to plants
Generally *P. fluorescens* is considered saprophytic but it may be an opportunistic pathogen causing soft rot in plants (OECD, 1997).

Allergenicity
In general fluorescent pseudomonads have not been described as allergens. However, they do possess an endotoxin (lipopolysaccharide) which may induce an allergic response in some individuals (OECD, 1997).

2mepsps
Assessments of the maize (*Zea mays* L.) source organism, the *2mepsps* gene, and the 2mEPSPS protein indicate that they are not pathogenic, allergenic, or toxic to mammals:

The maize source organism is a safe crop plant widely used for food and feed with little pathogenic, toxic, or allergenic effects on humans and animals. The 2*mepsps* gene is composed of the same

essential nucleic acids found in any food or feed DNA, which is commonly consumed as part of human or animal diets. Decades of research have indicated that dietary DNA poses no direct toxicity to human health. The EPSPS proteins are ubiquitous in nature, widely expressed in food and feed crops (e.g. soybean, tomato, maize). No health-related adverse effects have been associated with these proteins. Since the 2mEPSPS protein is derived from maize and has only two amino acid modifications, the safety profile of the novel protein is expected to remain unchanged relative to its wild-type counterpart. The 2mEPSPS protein is highly homologous to, and shares similar molecular weight and functionalities with other shikimate synthase proteins which have been demonstrated to be non-toxic and non-allergenic over the years through consumption. Its identity with the wt EPSPS enzyme is greater than 99.5%.

Regulatory sequences

The promoter and terminator sequences used in FG72 are derived from common plants or plant pathogens. These genetic elements constitute a minute component of their respective genomes, no genes that may be implicated in human disease, allergies or toxic effects have been transferred. Many of these organisms from which these elements are derived are model species in plant science with a history of safe use. These elements are described in Table 5, Section 2.3(c)(i).

(iii) Information about the history of use of the organism in the food supply or history of human exposure to the organism through other than intended food use (e.g. as a normal contaminant).

Glycine max

Cultivated soybean, *Glycine max* (L.) Merr, is grown as a commercial crop in over 35 countries. Today the major producers of soybeans are the United States, China, Democratic People's Republic of Korea and Republic of Korea, Argentina and Brazil. Soybean is one of the oldest cultivated crops, native to North and Central China. The first recording of soybeans was in a series of books known as Pen Ts'ao Kong Mu written by the emperor Sheng Nung in the year 2838 B.C., in which the various plants of China are described. Historical and geographical evidence suggests that soybeans were first domesticated in the eastern half of China between the 17th and 11th century B.C. (OECD, 2000). Domestication occurred over many centuries and was highlighted during the Shang Dynasty about 1700-1100 B.C. During the period of strong emperors, soybean remained only in China. In later centuries, increased trading and emigration brought soybean germplasm to other areas of Southeast and South-central Asia, which became the secondary centre of soybean germplasm. These events occurred during the 1st through the 15 - 16th century A.D. (Hymowitz *et al*., 1981). Soybeans were first introduced into the United States, now a major producer, in 1765 (OECD, 2000), and became established as an oilseed crop by the late 1920s. By World War II soybeans attained major commercial importance, and in the present day soybeans belong to the four principal oilseed crops in the US (soybean, cottonseed, peanuts and sunflowers) (Hui, 1992).

Soybean is grown primarily for the production of seed, has a multitude of uses in the food and industrial sectors, and represents one of the major sources of edible vegetable oil and of proteins for livestock feed use. A major food use is purified oil, utilised in margarines, shortenings and cooking and salad oils. Other food products include tofu, soya sauce, simulated milk and meat products. Soybean meal is also used as a high protein supplement in feed rations for livestock. Industrial uses of soybeans range from the production of yeasts and antibodies to the manufacture of soaps and disinfectants (OECD, 2000).

2mepsps gene

The coding sequence for the 2mEPSPS protein was isolated from *Zea mays* (maize). Maize is one of the few major crops that are indigenous to the Western Hemisphere and it is grown in nearly all areas of the world (Hallauer *et al*., 1988). There are many food/feed and industrial products that contain ingredients derived from maize. It is an important crop in human and animal nutrition because of its high levels of starch, protein, oil and other nutritionally valuable components. Consequently, maize has a very long history of safe use.

FSANZ has previously evaluated other transgenic crops that express EPSPS proteins (Table 3) and determined the protein indicates no potential for allergenicity or toxicity in humans.

Table 3 Transgenic crops expressing EPSPS proteins previously evaluated by FSANZ.

TRANSGENIC CROP	SOURCE OF TRANSGENE GENE	FSANZ APPLICATION
Canola	CP4 EPSPS from *A. tumefaciens* + GOX from *Ochrobactrum anthropi*	A363
Cotton	CP4 EPSPS from *A. tumefaciens*	A355, A553, A614
Lucerne	CP4 EPSPS from *A. tumefaciens*	A575
Maize	CP4 EPSPS from *A. tumefaciens* or 2mEPSPS from *Zea mays*	A362, A416, A548
Soybean	CP4 EPSPS from *A. tumefaciens*	A338, A592
Sugarbeet	CP4 EPSPS from *A. tumefaciens*	A378, A525
Wheat	CP4 EPSPS from *A. tumefaciens*	A524

hppdPf W336 gene

The *hppdPf W336* gene was isolated from *Pseudomonas fluorescens* strain A32. *Pseudomonas fluorescens* are ubiquitous bacterium in the natural environment and are frequently present in water, soil and plant rhizosphere (Bossis *et al*., 2000). The bacterium can be isolated from water, animals, human clinical specimens, the hospital environment, and spoiled foodstuffs such as fish and meat. The survival of *P. fluorescens* is affected by number of biotic and abiotic factors such as soil density, temperature, pH and humidity (OECD, 1997).

The natural properties of *P. fluorescens* are exploited in agriculture for plant growth-promotion (Fliessbach *et al*., 2009; OECD, 1997) and pest control. As a growth control agent, the bacterium can enhance plant growth through production of siderophores, which efficiently complex environmental iron rendering it unavailable to other organisms of the soil microflora. As a biopesticide, *P. fluorescens* is able to prevent the growth of frost-forming bacteria on leaves and blossoms of crops and fruits (Compant *et al*., 2005; Raaijmakers *et al*., 2006; US-EPA. 2008a), and prevent damping off diseases caused by fungi (Haas and Defago, 2005; Thrane *et al*., 2001; Voisard *et al*., 1989) and nematodes (Hamid *et al*., 2003) when used as a seed treatment. Naturally occurring strains of *P. fluorescens* have been registered commercially for the control of frost injury and fire blight on pear (Wilson and Lindow, 1993). Since 1992, four products containing *P. fluorescens* strains as active ingredients were approved by US-EPA (US-EPA, 2008b). The US-EPA recognized that this bacterial active ingredient is not expected to cause any adverse health effects in humans, based on various studies that found no evidence that these bacteria are harmful to mammals (US-EPA. 2008a). The US-EPA also established a tolerance exemption for residues of *P. fluorescens* in or on the raw agricultural commodity mushrooms (US-EPA, 1994). The pesticidal activity of *P. fluorescens* is attributed to three mechanisms: competition for an ecological niche or a substrate, production of inhibitory chemicals and induction of systemic resistance in host plants to a broad spectrum of pathogens (Compant *et al*., 2005; Haas and Defago, 2005).

In other applications, strains of *P. fluorescens* have been genetically modified to encapsulate crystal δ-endotoxins (Cry proteins) from the bacterium *Bacillus thuringiensis* (Bt) (Downing *et al*., 2000; Peng *et al*., 2003). The Cry proteins encapsulated by *P. fluorescens* showed high insecticidal activity and retained this activity for two to three times longer than Bt formulations (Peng *et al*., 2003). In pharmaceutical uses, *P. fluorescens* produces the antibiotic pseudomonic acid (also called mupirocin), which is used to prevent *Staphylococcus aureus* infections (Hothersall *et al*., 2007; Tacconelli *et al*., 2003). Further, in addition to the metabolic diversity of *P. fluorescens*, it may be used in bioremediation applications. The bacterium is able to degrade a wide variety of compounds, including 3-chlorobenzoic acid, naphthalene, phenathrene, fluorene and fluoranthene, chlorinated aliphatic hydrocarbons, styrene, pure hydrocarbons and crude oil (OECD, 1997).

In summary, the source of the *hppdPf W336* gene is ubiquitous in the environment, including soil, water and food. It has many beneficial uses in agriculture, human health and bioremediation. Despite this widespread presence, it is not described as allergenic, toxic or pathogenic to healthy humans and animals.

The promoter and terminator sequences used in FG72 are derived from common plants or plant pathogens. These genetic elements constitute a minute component of their respective genomes, no genes that may be implicated in human disease, allergies or toxic effects have been transferred. Many of the organisms from which these elements are derived are model species in plant science with a history of safe use.

(b) A description of the host organism into which the genes were transferred and its history of safe use for food, including:

(i) Any relevant phenotypic information;

As detailed above in Section 2.2(a)(iii), *G. max* is an established agricultural field crop that has been grown for millennia as a source of food and feed, and has a long history of safe use. Cultivated soybean is an erect, bushy herbaceous annual that can reach a height of 1.5 metres. Amongst the cultivated soybean varieties there are three types of growth habit: determinate, semi-determinate and indeterminate (Bernard and Weiss, 1973). Determinate growth is characterised by the cessation of vegetative activity of the terminal bud when it becomes an inflorescence at both axillary and terminal racemes. Determinate genotypes are primarily grown in the southern US (Maturity Groups V to X). Indeterminate genotypes continue vegetative activity throughout the flowering period and are grown primarily in central and northern regions of North America (Maturity Groups 000 to IV). Semideterminate types have indeterminate stems that terminate vegetative growth abruptly after the flowering period. No cultivated soybean varieties are frost tolerant and they are unable to survive freezing winter conditions (OECD, 2001c).

Cultivated soybeans are characterised by primary leaves that are unifoliate, opposite and ovate; secondary leaves that are trifoliolate and alternate; and compound leaves with four or more leaflets. Soybean has a nodulated root system consisting of a taproot from which the lateral root system emerges. The plants of most cultivars are covered with fine trichomes, but glabrous types also exist. The papilionaceous flower consists of a tubular calyx of five sepals, a corolla of five petals (one banner, two wings and two keels), one pistil and nine fused stamens with a single separate posterior stamen. The pod is straight or slightly curved, varies in length from two to seven centimetres, and consists of two halves of a single carpel which are joined by a dorsal and ventral suture. The shape of the seed, usually oval, can vary amongst cultivars from almost spherical to elongate and flattened (OECD, 2001c).

Soybean is a quantitative short day plant. Consequently, photoperiodism and temperature response is important in determining areas of cultivar adaptation. Seed will germinate when the soil temperature reaches 10°C and will emerge in a 5 - 7 day period under favourable conditions. In new areas of soybean production an inoculation with *Bradyrhizobium japonicum* will be necessary for optimum efficiency of the nodulated root system. Soybeans do not yield well on acid soils and the addition of limestone may be required. Soybeans are often rotated with such crops as corn, winter wheat, spring cereals and dry beans (OECD, 2001c).

(ii) How the organism is typically propagated for food use;

Soybean is considered a self-pollinated species that is propagated commercially for food use by seed. Artificial hybridisation is used to breed commercial cultivars. The soybean flower stigma is receptive to pollen approximately 24 hours before anthesis and remains receptive 48 hours after anthesis. The anthers mature in the bud and directly pollinate the stigma of the same flower. As a result, soybeans exhibit a high percentage of self-fertilisation, and cross pollination is usually less than one percent (Caviness, 1966). A soybean plant can produce as many as 400 pods, with two to twenty pods at a single node. Each pod contains one to five seeds. Neither the seedpod, nor the seed, has morphological characteristics that would encourage animal transportation (OECD, 2000).

(iii) What part of the organism is typically used as food;

The two primary products of soybeans used in food and feed, oil and meal respectively, are derived from the bean or seed. The various food (and feed) uses of these products are detailed above in Sections 2.1(d) and 2.2(a)(iii).

(iv) Whether special processing is required to render food derived from the organism safe to eat; and

Three basic methods are used to process soybeans for use as food as feed: solvent extraction, hydraulic extraction and expeller extraction. Almost all soybean oil is extracted from the seed using

the solvent process. Prior to processing, seeds are cleaned, cracked to loosen the seed coat or hulls, dehulled and then conditioned to 10 - 11% moisture. The conditioned meats are then flaked and extracted with hexane to remove the oil. Hexane and oil in the miscella are separated by evaporation and the hexane is recovered. Residual hexane in the flakes is removed by steam treatment in a desolventiser-toaster. The heat treatment inactivates antinutritional factors, such as trypsin inhibitors and lectins, in the raw flakes and increases protein digestibility. A metric ton of soybeans yields about 180 kg oil and 790 kg meal. (Hui,1992). Figure 1 below shows the solvent extraction process.

Soybean oil

Soybean oil is the most valuable of the soybean products and is consumed almost entirely (more than 95%) as food. Food-grade soybean oil is used as salad and cooking oil, shortenings and margarines. For non-food uses, soybean oil is converted into alkyd resins for protective coatings, plasticisers, dimer acids, surfactants and a number of other products (Hui, 1992). To be suitable for human consumption, the extracted oil must undergo further processing, which is referred to as refining. Figure 2 below shows the oil refining procedure.

Soybean meal

Most soybean meal obtained via processing is used as a protein supplement in animal feeds. Only in the last 30 years have appreciable amounts been converted into products for human consumption, and these have been almost exclusively derived from defatted soybean flakes (Hui, 1992).

Soybean meal normally contains 41 – 50% protein, depending on the amount of hull removed. Because of their high protein content, protein meals are essential ingredients of poultry and livestock feeds. Soybean meal is often blended with corn meal in animal feeds because the two protein sources complement each other; soy supplies the lysine and corn the methionine necessary to provide a balanced ration at relatively low cost (Hui, 1992).

Soybean hulls

The hull is the tough protective covering of the seed which must be removed before the oil can be extracted. The primary use for soybean hulls is animal feed. Hulls are routinely removed during crushing of soybeans but are returned to the processing stream to be added to the meal fraction. Hulls are withheld from the meal only if their inclusion would cause the product to exceed the limit of allowable fibre. Excess hulls may be sold as feedstuffs or discarded as waste.

Soybean protein products

Three classes of protein products are derived from soybeans: defatted flours and grits, protein concentrates and protein isolates. Flours and grits (containing 40 – 50% protein) are made by grinding and sieving flakes. Concentrates (containing about 70% protein) are prepared by extracting and removing the soluble sugars from the defatted flakes by leaching with dilute acid at pH 4.5 or leaching with aqueous ethanol. Isolated soy proteins are obtained by extracting the soluble proteins with water at pH 8-9, precipitating at pH 4.5, centrifuging the resulting protein curd, washing, redispersing in water, and finally spray drying. Flours and concentrates are further processed into textured products that are used as meat extenders and substitutes. Protein isolate is used primarily as adhesives for clays used in coating of paper and paperboard to render surfaces suitable for printing (Hui, 1992).

Soy Lecithin - Phospholipids

Soybean has the highest phospholipid content of the common oilseeds. Crude lecithin is obtained by degumming the crude soybean oil. This process involves mixing the crude oil with about 2% water at a temperature of 60 – 80°C. The mixture is then centrifuged to separate the lecithin emulsion which is vacuum dried in a thin film evaporator to a water content of 0.2 - 0.8%. Crude lecithin consists of 45 – 60% phosphatides and 30 – 35% triglycerides, the remaining 5 – 10% are free fatty acids, carbohydrates, glycolipids, sterols, and tocopherols (Pardun, 1989). Soy lecithin is used as ingredient in margarine, chocolate, icecream and baked goods. Its non-food applications are in cosmetics, pharmaceuticals and as additives in technical products.

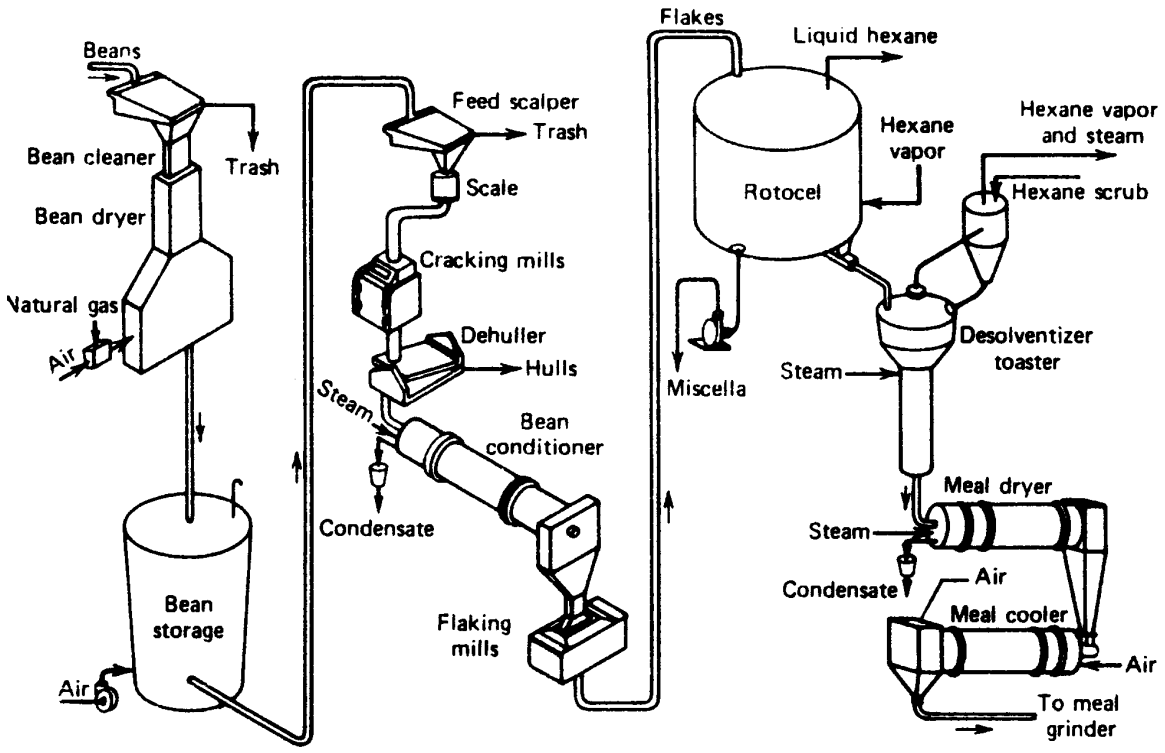
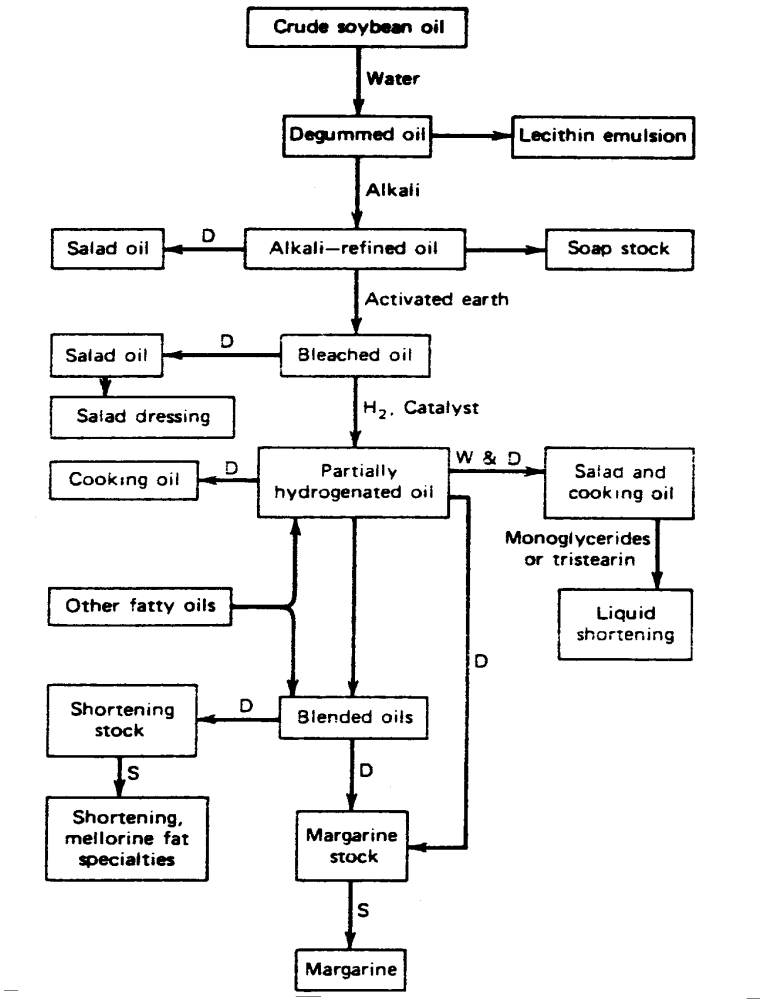


Figure 1 Processing of soybean into oil and meal by solvent extraction, courtesy of Dravo Corp. (Hui, 1992).



D = Deodorization, W = Winterisation, S = Solidification

Figure 2 Soybean oil refinement and edible soybean oil products, courtesy of the American Soybean Association and the American Oil Chemists' Society (Hui, 1992).

(v) *The significance to the diet in Australia and New Zealand of food derived from the host organism.*

Table 4 below details the import and export statistics for soybeans and process commodities for Australia and New Zealand.

Table 4 Soybean import and export statistics for Australia and New Zealand

Import & Export Quantity (tonnes)																
	2000		2001		2002		2003		2004		2005		2006		2007	
Commodity	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**
Australia																
Cake of Soybeans	26,001	2*	166,020*	2	320,081	6	349,587	31	226,490	1,333	241,333	199	294,654	116	686,189	76
Soya Sauce	6,992	81	6,515	141	6,644	173	7,820	121	8,536	84	8,313	152	8,923	191	9,327	106
Soybean oil	8,160	3,346	8,978	1,229	12,153	2,376	12,826	1,736	13,397	1,730*	10,511	578*	32,179	806	22,134	1,701
Soybeans	1,137	8,625	245*	6,978	594*	7,096	74,264	3,189	9,412	7,540	629	5,536	750	3,074	10,465	2,522
New Zealand																
Cake of Soybeans	43228	3	58096	0	59834	0	68738	2	64838	0	100722	47	65916	6	102661	220
Soya Sauce	1926	55	1243	44	1655	22	1526	30	1733	26	1585	28	1646	41	1789	27
Soybean oil	16383	40	13373	117	19281	46	17226	76	22479	71	26412	118	18951	668	19724*	204
Soybeans	470	21	503	26	749	3	730	3	807	7	860	6	853	4	1062	6

* Unofficial figure

Import & Export Value (1000 $)																
	2000		2001		2002		2003		2004		2005		2006		2007	
Commodity	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**	**Import**	**Export**
Australia																
Cake of Soybeans	6,133	5	39,633	4	73,706	4	78,721	30	57,817	485	68,834	99	64,195	79	208,077	85
Soya Sauce	9,656	161	9,549	212	10,336	310	11,774	238	13,806	95	14,269	185	14,462	275	16,200	300
Soybean oil	4,174	2,274	4,286	689	6,540	1,531	7,933	1,164	9,220	1,320	7,441	498	20,162	632	19,739	1,687
Soybeans	244	2,868	140	2,295	367	2,701	18,793	1,386	3,136	3,931	616	2,725	490	1,586	4,691	1,728
New Zealand																
Cake of Soybeans	10424	2	14522	0	15649	0**	19587	1	22258	0	27824	19	18552	3	35732	117
Soya Sauce	1841	67	1842	110	2120	64	1918	117	2220	105	2212	92	2248	128	2793	72
Soybean oil	8479	32	6135	83	10295	53	11088	105	16089	98	16996	184	13106	416	18627	458
Soybeans	213	9	231	10	349	2	385	2	490	7	483	5	607	3	816	6

** FAO estimate

FAOSTAT accessed 26th October 2009

2.3 The Nature of the Genetic Modification

(a) A description of the method used to transform the host organism.

Glycine max plants of variety Jack were genetically modified by means of direct gene transfer of a purified *Sal*I fragment from plasmid pSF10 into an embryogenic *G. max* cell line. Transformed cells were selected using isoxaflutole, and after a round of multiplication cycles in the presence of the selection agent, were regenerated into embryos and shoots in the absence of the selective agent. The regenerated plantlets were then transferred to the greenhouse and glyphosate was used as a selection agent and for herbicide tolerance evaluation. Surviving plantlets were allowed to flower and set seeds.

(b) Information about the intermediate host organisms (e.g. bacteria) used for all laboratory manipulations prior to transformation of the host organism.

The only intermediate organism used prior to transformation was *Escherichia coli*. Strains of *E. coli* are natural residents of the normal intestinal microbial flora of humans and animals. Standard *E. coli* strains used in laboratory techniques are non-pathogenic (Mühldorfer and Hacker, 1994).

(c) A description of the gene construct and the transformation vectors used, including:

(i) The size, source and function of all the genetic components including marker genes, regulatory and other elements; and

The genetic components comprising soybean event FG72 are detailed in Table 5 below, and in Criel (2009; Appendix 1). These components are show in Figure 3 in Section 2.3(c)(ii) below.

Table 5 Genetic elements comprising the pSF10 vector used in soybean event FG72

GENETIC ELEMENT	NT POSITION	SIZE (KB)	ORIENTATION	DESCRIPTION AND FUNCTION	REFERENCE
3'nos	3262 - 355		Counter Clockwise	Sequence including the 3' untranslated region of the nopaline synthase gene from the T-DNA of pTiT37	Depicker *et al.*, 1982
hppdPf W336	3554 - 4630	1077	Counter Clockwise	The coding sequence of the 4-hydroxyphenylpyruvate dioxygenase of *Pseudomonas fluorescens* strain A32 modified by the replacement of the amino acid Glycine 336 with a Tryptophane	Boudec *et al.*, 2001
TPotp Y	4631 - 5002		Counter Clockwise	Coding sequence of an optimized transit peptide derivative (position 55 changed into Tyrosine), containing sequence of the RuBisCO small subunit genes of Zea mays (corn) and *Helianthus annuus* (sunflower)	Lebrun *et al*., 1996
5'tev	5003 - 5143		Counter Clockwise	Sequence including the leader sequence of the tobacco etch virus	Carrington and Freed, 1990
Ph4a748 ABBC	5144 - 6433		Counter Clockwise	Sequence including the promoter region of the histone H4 gene of *Arabidopsis thaliana,* containing an internal duplication	Chaboute *et al.*, 1987
Ph4a748	6434 - 7448		Clockwise	Sequence including the promoter region of the histone H4 gene of *Arabidopsis thaliana*	Chaboute *et al.*, 1987
intronl h3At	7449 - 7929		Clockwise	First intron of gene II of the histone H3.III variant of *Arabidopsis thaliana*	Chaubete *et al.*, 1992
TPotp C	7930 - 8301		Clockwise	The coding sequence of the optimized transit peptide, containing sequence of the RuBisCO small subunit genes of Zea mays (corn) and *Helianthus annuus* (sunflower)	Lebrun *et al*., 1996
2mepsps	8302 – 9639		Clockwise	The coding sequence of the double-mutant 5-enolpyruvylshikimate-3-phosphate synthase gene of Zea mays (corn)	Lebrun *et al.*, 1997
3'histonAt	9640- 10326		Clockwise	The sequence including the 3' untranslated region of the histone H4 gene *of Arabidopsis thaliana*	Chaboute *et al.*, 1987
	10327-10398 1 -232			Sequence of the pMCS5 vector	Hoheisel, 1994
	233 - 457			Sequences of the pUC19 vector	Yanisch-Perron *et al.*, 1985
ORI ColE1	458 - 1244			Fragment including the origin of replication from the plasmid pBR322 for replication in *Escherichia coli*	Bolivar *et al.*, 1977
	1245 - 1403			Sequences of the pUC19 vector	Yanisch-Perron *et al.*, 1985
bla	1404 - 2264		Counter clockwise	Fragment including the beta-lactamase gene of plasmid pBR322 of *Escherichia coli*	Bolivar *et al.*, 1977
	2265 - 2394			Sequences of the pUC19 vector	Yanisch-Perron *et al.*, 1985
ORI f1	2395 - 2840			Fragment including the origin of replication of the filamentous phage f1	Dotto *et al.*, 1982
	2841 - 3261			Sequences of the pUC19 vector	Yanisch-Perron *et al.*, 1985

(ii) A detailed map of the location and orientation of all the genetic components contained within the construct and vector, including the location of relevant restriction sites.

The plasmid vector pSF10 that was used in soybean event FG72 is shown in Figure 3 below and described by Criel (2009; Appendix 1).

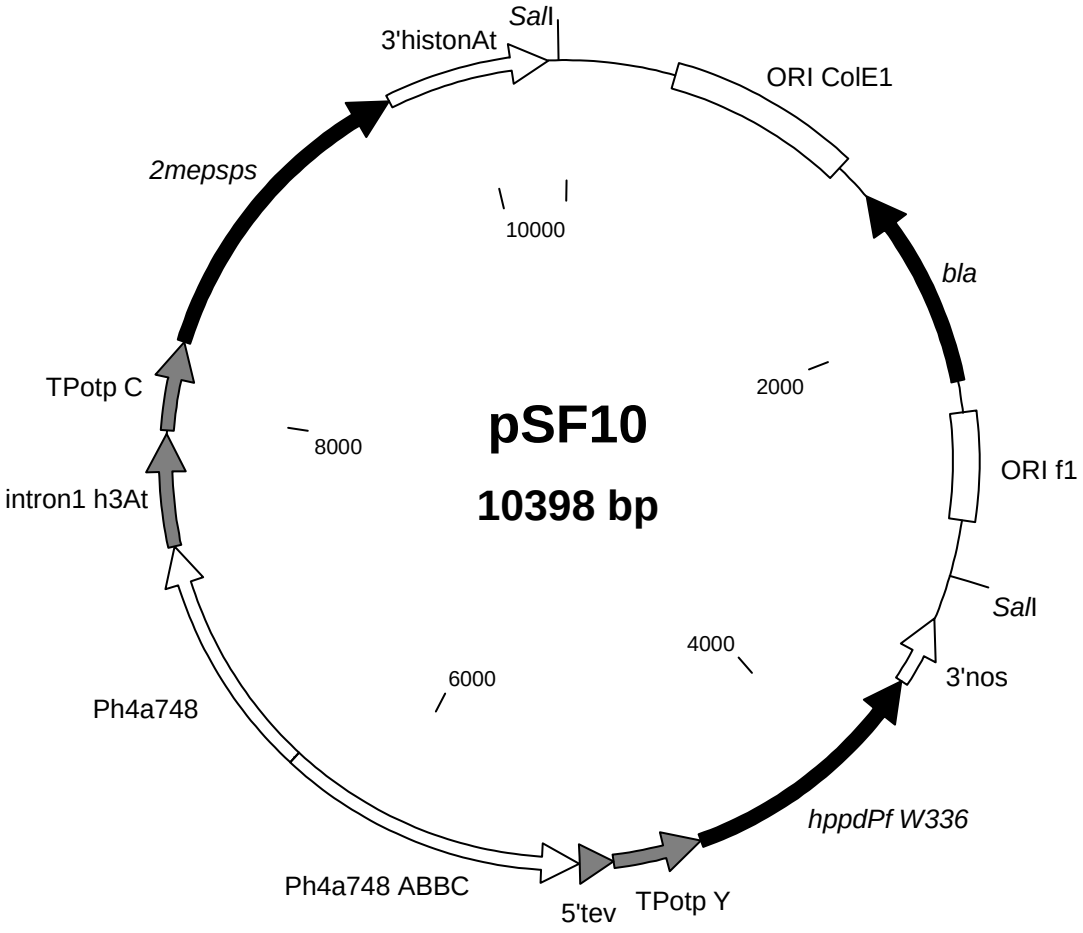


Figure 3 Map of plasmid vector pSF10 used in event FG72

(d) A full molecular characterisation of the genetic modification in the new organism, including:

(i) Identification of all transferred genetic material and whether it has undergone any rearrangement;

Transgenic Locus
For molecular characterization of the transgenic locus, genomic DNA was isolated from plants of *G. max* FG72 plants, and from non-transgenic *G. max* plants of variety Jack. The organization of the inserted genetic elements was assessed by means of Southern blot analysis and full DNA sequence determination of the transgenic locus.

Genomic DNA was isolated from FG72 plants and digested with ten different restriction enzymes: *Hinc*II, *Sac*I, *Hind*III, *Bsp*HI, *Apa*I, *Stu*I, *Nco*I, *Sca*I, *Eco*RI and *Bsu36*I. Wild type genomic DNA (non-transgenic variety Jack) digested with *Hind*III was used as negative control; wild type genomic DNA digested with *Hind*III and supplemented with an equimolar amount of pSF10 plasmid DNA digested with *Hind*III was used as positive control. The resulting DNA fragments were separated by agarose gel-electrophoresis, transferred to a membrane and hybridized with different radioactive labelled probes:

eight probes targeted single genetic elements present in the pSF10 vector used for the transformation (PT015, PT016, PT024, PT059, PT060, PT061, PT062 and PT063), and a ninth probe targeted the complete transfer DNA (PT058). Table 6 below provides details of the probes used in the Southern blot analysis, and the analysis is described in detail in Verhaeghe (2009a; Appendix 2).

The Southern blot analysis demonstrated that the transgenic sequence in event FG72 consists of two partial 3′histonAt sequences in a head to head orientation, followed by two complete transfer DNA copies (the full FG72 insert) arranged in a head to tail orientation. Upon integration of the FG72 insert into the soybean genome, a non-transgenic region translocated to a new position, which is joined at the 3′ junction by 158 bases of Ph4a748 promoter sequences (Verhaeghe, 2009a; Appendix 2 and Verhaeghe, 2009b; Appendix 3). Figure 4 below shows a schematic presentation of the genetic elements inserted into the *G. max* genome in event FG72.

The organisation of the FG72 transgenic locus and the Southern blot strategy is shown in Figure 5, with indication of the restriction enzymes and probes used and the expected hybridization fragments. The expected and obtained hybridization fragments for each of the ten different restriction enzymes are provided in Table 7. The hybridization results support the model of the FG72 insert organization as described above and depicted in Figures 4 and 5 (Verhaeghe, 2009a; Appendix 2).

The full DNA sequence of the FG72 transgenic locus was determined via PCR amplification of six overlapping DNA fragments: FG72-TR1, FG72-TR2, FG72-TR3, FG72-TR4, FG72-TR6, FG72-TR6 (see Figure 4 below). The sequences of the newly created junctions resulting from the translocation of non-transgenic sequences were also determined via the amplification of two fragments spanning these junctions: FG72-TL1, FG72-TL2 (see Figure 4 below). Details of the PCR strategy used to amplify the FG72 transgenic locus and translocated sequences are provided below in Table 8. Following 4-fold sequencing of all PCR fragments, the FG72-TR consensus sequence was determined by combining the sequence results of fragments FG72-TR1 through FG72-TR6 (see Table 8 below). Alignment of the FG72-TR consensus sequence, and the FG72-TL1 and FG72-TL2 sequences with the pSF10 plasmid sequence and wild type (non-transgenic, variety Jack) DNA sequences was performed using Clone Manager software. The results of this analysis are presented in Figure 4 and Table 8 below. Full details of the DNA sequence analysis are presented in Verhaeghe (2009b; Appendix 3).

Table 6 Details of the probes used in the Southern blot analysis for characterisation of the FG72 transgenic locus

PROBE TEMPLATE ID	GENETIC ELEMENT	SIZE PROBE TEMPLATE (BP)	POSITION IN PSF10 (BP)
PT015-1	*2mepsps*	1351	8309 → 9659
PT016-1	3′histonAt	753	9614 → 10366
PT024-2	3′nos	214	3265 → 3478
PT058-4	T-DNA probe	7204	3142 → 10345
PT059-1	Ph4a748	959	6491 → 7449
PT060-1	Intron1 h3At	507	7446 → 7952
PT061-1	5′tev + TPotp Y	460	4650 → 5109
PT062-1	Ph4a748B	430	6866 → 7295
PT063-1	*hppdPf W336*	1055	3558 → 4612

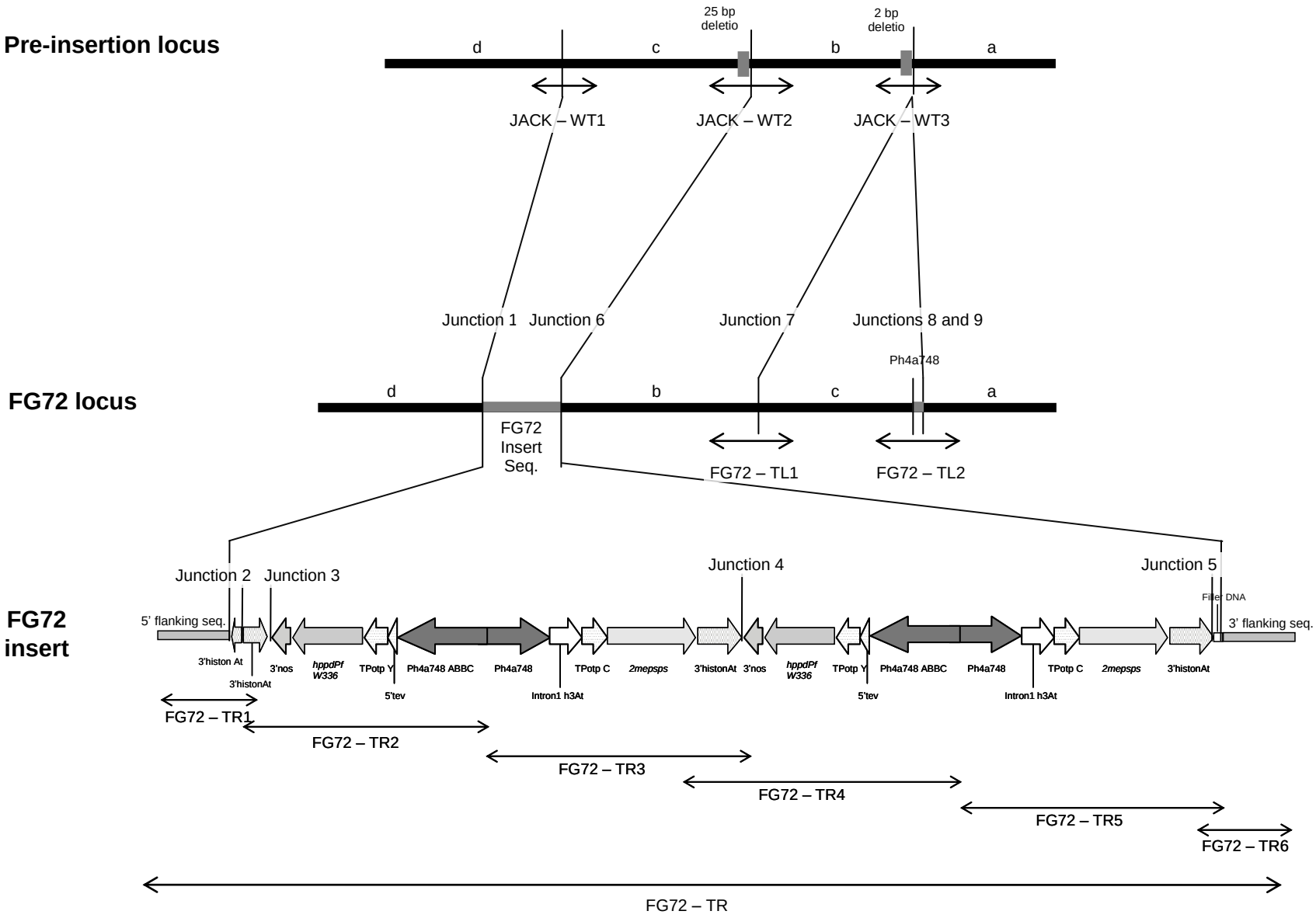


Figure 4 Overview of the FG72 transgenic locus

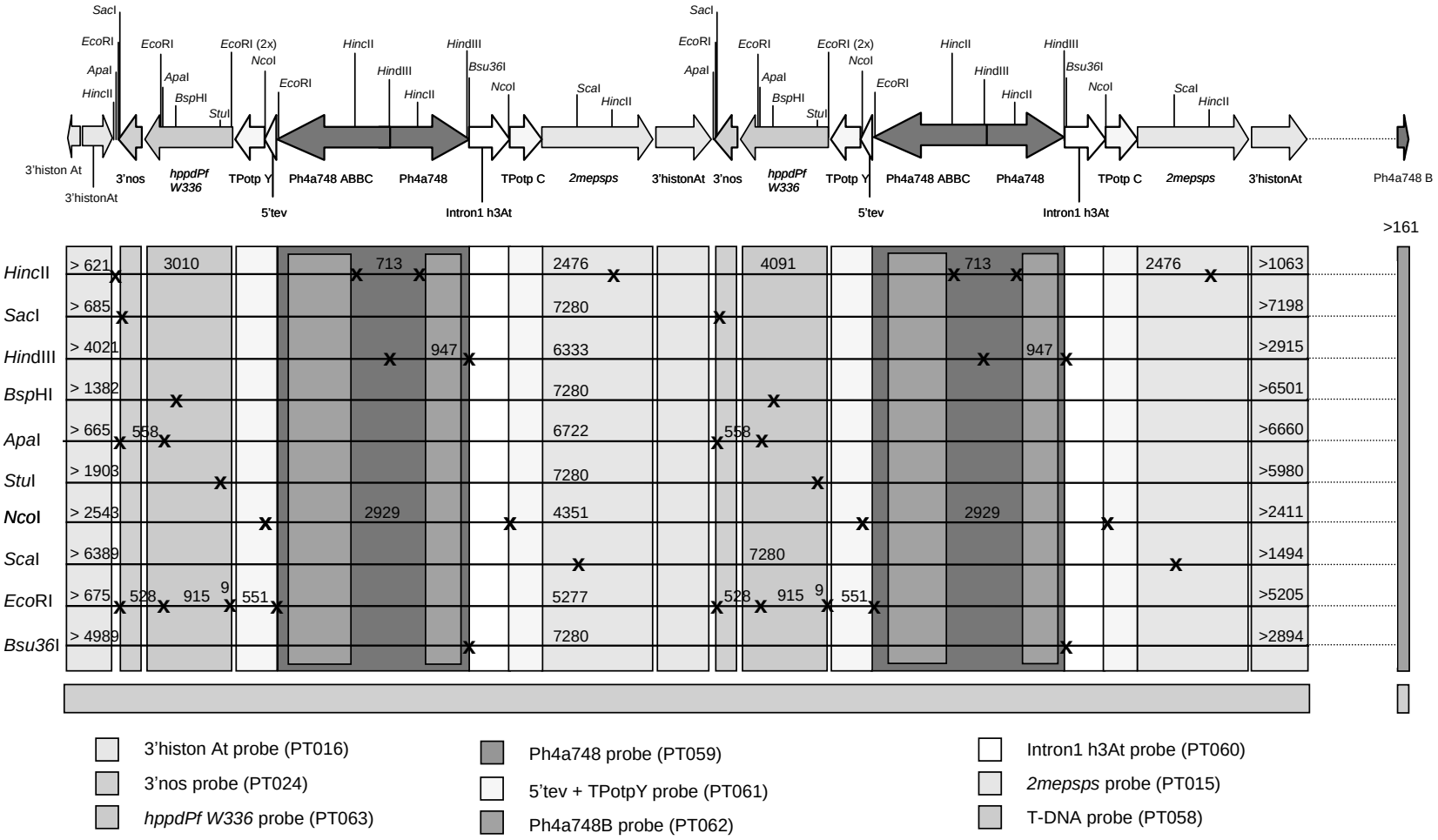


Figure 5 Schematic drawing of the event FG72 insert indicating restriction enzymes, probes and expected fragment lengths

Table 7 Expected (Exp) and obtained (Obt) hybridization fragments in the Southern blot analysis of the FG72 transgenic locus

Digest	Description	Expected fragment sizes (bp)	Obtained fragment sizes (bp)	PT016-1: 3'histonAt		PT024-2: 3'nos		PT063-1: hppdPf W336		PT061-1: 5'tev+TPotp Y		PT059-1: Ph4a748		PT062-1: Ph4a748B		PT060-1: Intron1 h3At		PT015-1: 2mepsps		PT058-4: T-DNA probe	
				Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
HincII	5' integration fr.	> 621	5250	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	No[a]
	internal fr.	3010	3010	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes
	internal fr.	4091	4091	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
	internal fr.	713	713	No	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No	No	Yes	No[a]
	internal fr.	2476	2476	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' integration fr.	> 1063	1130	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	1300	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No[a]
SacI	5' integration fr.	> 685	6060	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes, weak
	internal fr.	7280	7280	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' integration fr.	> 7198	>14 kb	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	>14 kb	No	No	No	No	No	No	No	No	Yes	No[a]	Yes	Yes	No	No	No	No	Yes	No[a]
HindIII	5' integration fr.	> 4021	9550	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes
	internal fr.	6333	6333	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	internal fr.	947	947	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes
	3' integration fr.	> 2915	5500	Yes	Yes	No	No	No	No	Yes	Yes	Yes	No[a]	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	1480	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No[a]
BspHI	5' integration fr.	> 1382	3200	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	internal fr.	7280	7280	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' integration fr.	> 6501	7480	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	4260	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No[a]
ApaI	5' integration fr.	> 665	10760	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes, weak
	internal fr.	6722	6722	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	internal fr.	558	558	No	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	3' integration fr.	> 6660	7900	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Digest	Description	Expected fragment sizes (bp)	Obtained fragment sizes (bp)	PT016-1: 3'histonAt		PT024-2: 3'nos		PT063-1: hppdPf W336		PT061-1: 5'tev+TPotp Y		PT059-1: Ph4a748		PT062-1: Ph4a748B		PT060-1: Intron1 h3At		PT015-1: 2mepsps		PT058-4: T-DNA probe	
				Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
	Partial fr.	/	8570	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
	3' junction translocation	> 158	>14 kb	No	No	No	No	No	No	No	No	Yes	No[a]	Yes	Yes, weak	No	No	No	No	Yes	No[a]
StuI	5' integration fr.	> 1903	4710	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	internal fr.	7280	7280	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' integration fr.	> 5980	6210	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	>14 kb	No	No	No	No	No	No	No	No	Yes	No[a]	Yes	Yes, weak	No	No	No	No	Yes	No[a]
NcoI	5' integration fr.	> 2543	7660	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	Yes	Yes
	internal fr.	2929	2929	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
	internal fr.	4351	4351	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No[a]	Yes	Yes	Yes	Yes
	3' integration fr.	> 2411	10270	Yes	Yes	No	No	No	No	Yes	Yes	No	No	No	No	Yes	No[a]	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	>14 kb	No	No	No	No	No	No	No	No	Yes	Yes, weak	Yes	Yes	No	No	No	No	Yes	No[a]
ScaI	5' integration fr.	> 6389	9900	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	internal fr.	7280	7280	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' integration fr.	> 1494	4730	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	11430	No	No	No	No	No	No	No	No	Yes	No[a]	Yes	Yes	No	No	No	No	Yes	No[a]
EcoRI	5' integration fr.	> 675	5110[d]	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	internal fr.	5277	5277[d]	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	internal fr.	528[c]	528[c]	No	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	internal fr.	915	915	No	No	No	No	Yes	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	internal fr.	9[b]	/	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
	internal fr.	551[c]	551[c]	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	Yes	Yes
	3' integration fr.	> 5205	9610	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' junction translocation	> 158	4670	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No[a]
Bsu36I	5' integration fr.	> 4989	7650[e]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes
	internal fr.	7280	7280[e]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3' integration fr.	> 2894	10350	Yes	Yes	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Digest	Description	Expected fragment sizes (bp)	Obtained fragment sizes (bp)	PT016-1: 3'histonAt		PT024-2: 3'nos		PT063-1: hppdPf W336		PT061-1: 5'tev+TPotp Y		PT059-1: Ph4a748		PT062-1: Ph4a748B		PT060-1: Intron1 h3At		PT015-1: 2mepsps		PT058-4: T-DNA probe	
				Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
	3' junction translocation	> 158	7670[e]	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes
	Partial fr.	/	11590	No	No	No	No	No	No	No	No	No	Yes, weak	No	Yes, weak	No	No	No	No	No	No
	Partial fr.	/	>14 kb	No	No	No	No	No	No	No	No	No	Yes, weak	No	Yes, weak	No	No	No	No	No	No
WT genomic DNA-HindIII				/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
WT genomic DNA - HindIII + pSF10 - HindIII	Positive control	3420	3420	No	No	Yes	Yes	Yes	Yes	Yes	Yes, weak	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes
	Positive control	947	947	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No[a]
	Positive control	2961	2961	Yes	Yes	No	No	No	No	Yes	Yes, weak	Yes	No[a]	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	Positive control	3070	/	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Hybridization ID				H1/09-005/01-F2		H1/09-005/20-F3		H1/09-005/08-F1		H1/09-005/19-F7		H1/09-005/06-F2		H1/09-005/07-F2		H1/09-005/18-F1		H1/09-005/05-F2		H2/09-005/19-F1	

[a] to a small overlap between the fragments and the T-DNA probe and/or to the ratio fragment length/probe length, these expected fragments could not be visualized.
[b] This fragment is too small to be visualized.
[c,d, e] These fragments can appear as a single fragment.

Table 8 Details of the PCR amplification strategy for sequencing the FG72 transgenic locus.

Fragment ID	Template DNA	Primer pair	Primer position in pSF10	Length of amplicon (bp)	Overlap between fragments (bp)
FG72 - TR1	*Hin*dIII digested FG72 gen. DNA	STV162	/	ca. 1927	
		MAE082	10190 → 10171		ca. 263
FG72 - TR2	*Hin*dIII digested FG72 gen. DNA	STV026	/	ca. 3819	
		STV065	6454 → 6436		ca. 30
FG72 - TR3	FG72 gen. DNA	STV063	6425 → 6443	ca. 4141	
		DPA288	3285 → 3266		ca. 1049
FG72 - TR4	FG72 gen. DNA	MLD108	9517 → 9536	ca. 4218	
		STV065	6454 → 6436		ca. 30
FG72 - TR5	FG72 gen. DNA	STV063	6425 → 6443	ca. 4229	
		JDB003 *	/		ca. 556
FG72 - TR6	FG72 gen. DNA	SHA096	10098 →10123	ca. 1454	
		SMP185	/		N.A.
FG72 - TL1	FG72 gen. DNA	TVS032	/	ca. 2257	
		TVS030	/		N.A.
FG72 - TL2	FG72 gen. DNA	TVS031	/	ca. 2481	
		STV068	/		

* Comparison of the JDB003 primer sequence with the obtained consensus sequence of fragment FG72-TR revealed that this primer is not 100% homologous to the consensus sequence

Vector Backbone Sequences

Southern blot analysis was performed to verify the absence of pSF10 vector backbone sequences in soybean event FG72. Genomic DNA was isolated from FG72 plants and digested with two restriction enzymes: *Hin*dIII and *Hinc*II. Wild type *G. max* genomic DNA (non-transformed) digested with *Hin*dIII was used as negative control, and a positive control was prepared by supplementing wild type genomic DNA digested with *Hin*dIII with an equimolar or 0.1x equimolar amount of pSF10 plasmid DNA digested with *Hin*dIII. Restricted DNA fragments were separated by means of agarose gel electrophoresis, transferred to a membrane and hybridized with two overlapping vector backbone probes covering the complete vector backbone sequences of the pSF10 transformation vector, followed by re-hybridization with a T-DNA probe that targeted the transfer DNA. Details of the probes used in this analysis are presented in Table 9 and Figure 6 below.

Table 9 Details of probes used in the Southern blot analysis to verify the absence of vector backbone sequences in event FG72.

Probe template ID	Description	Primer pair/ Restriction digest	Position in pSF10	Size probe template (bp)	Overlap between probe templates (bp)
PT056	Vector backbone probe	KM033	bp 10356 → bp10373	1730	573
		DPA010	bp 1687 → 1666		
PT057	Vector backbone probe	VH055	bp 1115 → bp 1134	1982	
		STV039	bp 3096 → bp 3077		N.A.
PT058	T-DNA probe	*Sac*I/*Sma*I	bp 3142 → bp 10345	7204	

Only the expected hybridisation fragments were obtained in the Southern blot analysis for the detection of vector backbone sequences. The expected Southern blot profile was obtained in the FG72 samples after hybridization with the T-DNA probe (Table 10). Both vector backbone probes contain regions that are also present in the T-DNA sequence, therefore several fragments originating from inserted transgenic DNA also hybridized with the vector backbone probes. This analysis confirmed the absence of vector backbone sequences in the genome of the *G. max* transformation event FG72. This analysis is detailed in Verhaeghe (2009c; Appendix 4).

Preinsertion Locus

To determine any wild type *G. max* sequences that might have been interrupted as a result of the transformation process, genomic DNA was isolated from the non-transgenic *G. max* variety Jack, and four fragments were amplified spanning the 5′ integration site of the FG72 insert (JACK-WT1), the 3′ integration site of the FG72 insert (JACK-WT2a and JACK-WT2b) and the reintegration site of the translocating genomic sequences (JACK-WT3). Alignment of these three wild type sequences with the transgenic locus consensus sequences FG72-TR1, FG72-TL1 and FG72-TL2, revealed:

(i) For the 2303 bp JACK-WT1 fragment, bases 1-1166 are completely identical to bases 286-1451 of FG72-TR, representing the 5′ flanking sequence of the FG72 locus, and bases 1167-2303 are completely identical to bases 1081-2217 of FG72-TL1, representing the 5′ sequence of the translocated region.

(ii) For the 2991 bp JACK-WT2 fragment, bases 648-1798 are identical to bases 1-1151 of FG72-TL2, representing the 3′ sequence of the translocated region, and bases 1824-2991 are identical to bases 16639-17805 of FG72-TR, representing the 3′ flanking sequenc of the FG72 locus. Bases 1799-1823 of fragment JACK-WT2 did not show homology with FG72 sequences and are annotated as bases deleted upon transformation.

(iii) For the 2212 bp JACK-WT3 fragment, bases 1-1082 are identical to bases 1-1082 of FG72-TL1, representing the 5′ flanking sequence of the translocated region, and bases 1083-2212 are identical to bases 1310-2439 of FG72-TL2, representing the 3′ flanking sequence of the translocated region. Bases 1801-1802 of fragment JACK-WT3 are bases deleted upon transformation.

Shorter transgenic sequences were evident compared to wild type fragments due to primer design. An overview of this analysis is presented in Figure 4 above, and detailed in Verhaeghe (2009b; Appendix 3).

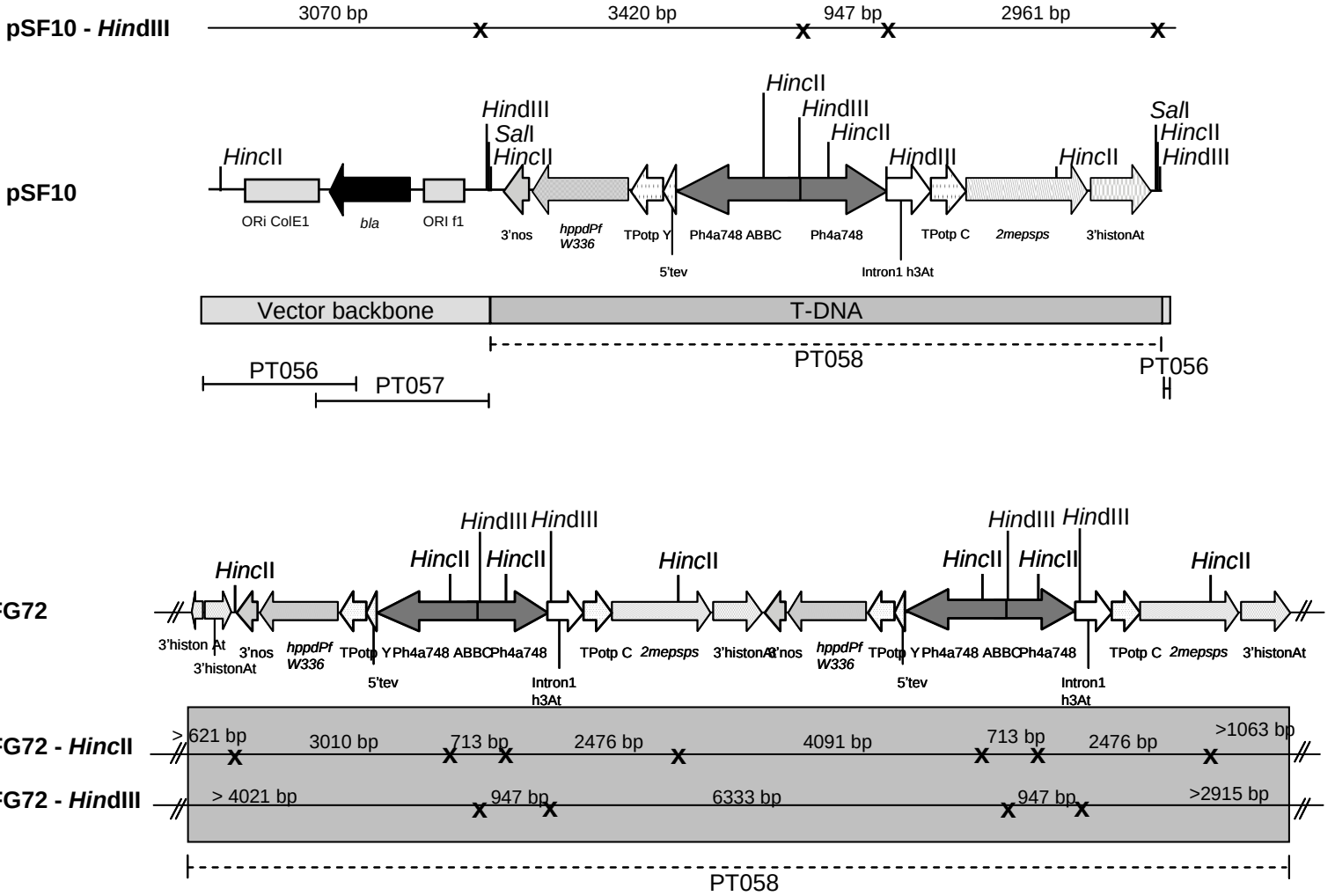


Figure 6 Schematic drawing of pSF10 with indication of the relevant restriction sites and the position of the used probe templates.

Table 10 Expected (exp) and observed (obs) hybridisation fragments in the Southern blot analysis for the detection of vector backbone sequences in the FG72 transgenic locus

Sample	Expected T-DNA or plasmid fragment sizes	Fragment description	M/09-003/06				M/09-003/07			
			PT056-1 Vector backbone probe		PT058-6 T-DNA probe		PT057-1 Vector backbone probe		PT058-6 T-DNA probe	
			Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
FG72 - *Hin*dIII	9550 bp [c]	5' integration fr.	Yes [d]	Yes	Yes	Yes	Yes [d]	Yes	Yes	Yes
	947 bp	internal fragment	Yes [d]	Yes	Yes	Yes	Yes [d]	No	Yes	Yes
	6333 bp	internal fragment	Yes [d]	Yes	Yes	Yes	Yes [d]	Yes	Yes	Yes
	5500 bp [c]	3' integration fr.	No	No	Yes	Yes	No	No	Yes	Yes
	1480 bp [c]	3' junction translocation	No	No	Yes [b]	No	No	No	Yes [b]	No
FG72 – *Hinc*II	5250 bp [c]	5' integration fr.	Yes [d]	No	Yes [b]	No	No	No	Yes [b]	No
	3010 bp	internal fragment	Yes [d]	Yes	Yes	Yes	Yes [d]	Yes	Yes	Yes
	713 bp	internal fragment	Yes [d]	Yes	Yes	Yes, weak	Yes [d]	No	Yes	Yes, weak
	2476 bp	internal fragment	No	No	Yes	Yes	No	No	Yes	Yes
	4091 bp	internal fragment	Yes [d]	Yes	Yes	Yes	Yes [d]	Yes	Yes	Yes
	1130 bp [c]	3' integration fr.	No	No	Yes	Yes	No	No	Yes	Yes
	1300 bp [c]	3' junction translocation	No	No	Yes [b]	No	No	No	Yes [b]	No
WT - *Hin*dIII	/	/	/	/	/	/	/	/	/	/
WT - *Hin*dIII + 0.1 equimolar amount pSF10 - *Hin*dIII	3420 bp	positive control	Yes [d]	No	Yes	No [a]	Yes [b]	No	Yes	No [a]
	947 bp	positive control	Yes [d]	No	Yes	No [a]	Yes [d]	No	Yes	No [a]
	2961 bp	positive control	Yes [b]	No	Yes	No [a]	No	No	Yes	No [a]
	3070 bp	positive control	Yes	Yes	Yes [b]	No	Yes	Yes	Yes [b]	No
WT - *Hin*dIII + 1 equimolar amount pSF10 - *Hin*dIII	3420 bp	positive control	Yes [d]	No	Yes	Yes	Yes [b]	No	Yes	Yes
	947 bp	positive control	Yes [d]	No	Yes	No [a]	Yes [d]	No	Yes	No [a]
	2961 bp	positive control	Yes [b]	No	Yes	Yes	No	No	Yes	Yes
	3070 bp	positive control	Yes	Yes	Yes [b]	No	Yes	Yes	Yes [b]	No
		Hybridization ID	H1/09-003/06-F2		H3/09-003/06-F1		H1/09-003/07-F1		H3/09-003/07-F1	

[a] These fragments of the positive control could not or very weakly be visualized after hybridization with the T-DNA probe. This has no impact on the interpretation of the results. - [b] The overlap between the probe and the fragment can be too small to visualize this fragment. - [c] Expected fragment sizes as determined in the detailed insert characterization study (Appendix 2). - [d] Since part of the sequence of both vector backbone probes is as well present in the T-DNA sequences, fragments originating from inserted transgenic sequences are hybridizing with vector backbone probes.

(ii) A determination of the number of insertion sites, and the number of copies at each insertion site;

As detailed above in Section 2.3 (d)(i), Southern blot analysis and DNA sequencing of the FG72 transgenic locus revealed that the inserted genetic material consists of two partial 3′histonAt sequences in a head to head orientation, followed by two complete transfer DNA copies (the full FG72 insert) arranged in a head to tail orientation. Upon integration of the FG72 insert into the soybean genome, a non-transgenic region translocated to a new position, which is joined at the 3′ junction by 158 bases of Ph4a748 promoter sequences from the FG72 transfer DNA (Verhaeghe, 2009a; Appendix 2 and Verhaeghe, 2009b; Appendix 3). Figure 6 above demonstrates the number of insertion sites and the number of copies of the genetic elements introduced by the FG72 transformation event.

(iii) Full DNA sequence data of each insertion event, including junction regions with the host DNA, sufficient to identify any substances expressed as a consequence of the inserted material, or where more appropriate, other information such as analysis of transcripts or expression products to identify any new substances that may be present in the final food;

Full DNA sequence of the FG72 transgenic locus, the insertion site flanking regions and the pre-insertion locus are provided in Verhaeghe (2009b; Appendix 3).

(iv) A map depicting the organisation of the inserted genetic material at each insertion site; and

The organisation of the FG72 transgenic locus is demonstrated in Figure 6 above.

(v) The identification and characterisation of any unexpected open reading frames within the inserted DNA or created by insertion with contiguous genomic DNA, including those that could result in fusion proteins or unexpected protein expression products.

As described in Sections 2.3(d)(i-iii) above, Southern blot and full sequence analyses of the FG72 transgenic locus revealed that the insert sequence of 15188 bp consists of two partial 3′histonAt sequences in a head to head orientation, followed by two complete transfer DNA copies (the full FG72 insert) arranged in a head to tail orientation. Also, upon integration of the insert into the soybean genome, a non-transgenic region translocated to a new position, which is joined at the 3′ junction by 158 bases of Ph4a748 promoter sequences from the FG72 transfer DNA (Verhaeghe, 2009a; Appendix 2 and Verhaeghe, 2009b; Appendix 3).

Bioinformatics analyses were performed on the FG72 transgenic locus sequences and pre insertion locus sequences using current databases and bioinformatics tools to predict the presence of potential newly created open reading frames (ORFs) leading to the unintended expression of new proteins (Verhaeghe, 2009d; Appendix 5). An overview of the transgenic and pre-insertion query sequences and junctions analysed in the bioinformatics analyses are described below and indicated in Figure 4 in Section 2.3 (d)(i) above.

Pre-insertion Locus
The sequences of three regions of *G. max* wild type genomic DNA that were interrupted upon insertion of the FG72 transgenic locus were used as query sequences in the bioinformatic analysis:

(1) Fragment JACK-WT1 (2303 bp) comprising the 5′ integration site of the FG72 insert. This fragment contains 1 junction:
    (i) Junction 10 between the 5′ pre insertion locus sequences and the 5′ end of the translocating region.
(2) Fragment JACK-WT2 (2991 bp) comprising the 3′ integration site of the FG72 insert. This fragment contains 2 junctions:
    (i) Junction 11 between the 3′ end of the translocating region and bases deleted upon transformation.
    (ii) Junction 12 between bases deleted upon transformation and the 3′ pre insertion locus

sequences.

(3) Fragment JACK-WT3 (2212 bp) comprising the reintegration site of the translocating sequences. This fragment contains 2 junctions:
   (i) Junction 13 between the 5′ flanking sequences of the reintegration site and bases deleted upon transformation
   (ii) Junction 14 between bases deleted upon transformation and the 3′ flanking sequences of the reintegration site.

Transgenic Locus

Three transgenic fragments were used as query sequences:

(1) Fragment FG72-TR (17806 bp) containing the inserted DNA and 5′ and 3′ flanking regions (see Section (i) above). This sequence contains 6 junctions:
   (i) Junction 1 between the 5′ flanking sequences and the first partial 3′histonAt sequences.
   (ii) Junction 2 between 2 partial 3′histonAt sequences.
   (iii) Junction 3 between the second partial 3′histonAt sequence and the first complete T-DNA copy.
   (iv) Junction 4 between the 2 complete T-DNA copies.
   (v) Junction 5 between the second complete T-DNA copy and the filler DNA.
   (vi) Junction 6 between the filler DNA and the 3′ flanking sequences.

(2) Fragment FG72-TL1 (2217 bp) containing the 5′ end of the translocated region and the sequences flanking this region. This fragment contains 1 junction:
   (i) Junction 7 between the 5′ end of the translocated region and its flanking sequences.

(3) Fragment FG72-TL2 (2439 bp) containing the 3’ end of the translocated region, 158 bp of Ph4a748 promoter sequences and the sequences flanking these promoter sequences. This fragment contains 2 junctions:
   (i) Junction 8 between the 3′ end of the translocated region and Ph4a748 promoter sequences.
   (ii) Junction 9 between Ph4a748 promoter sequences and its flanking sequences.

Homology with known functional genes or proteins

The Basic Local Alignment Search Tool (BLAST; Altschul *et al*., 1997) finds regions of local similarity between sequences. BLASTx compares the six-frame theoretical translation products of the nucleotide query sequence (both strands) against a protein sequence database. This analysis was performed on non transgenic *G. max* sequences in order predict gene sequences in the pre-insertion locus. This search identified sequence homology at the 5′ end of the translocated region only with part of a putative cysteine protease (Figure 7). This protein is not interrupted upon transformation with the FG72 event. No relevant known functional genes interrupted upon transformation were identified in this analysis.

Gene prediction and open reading frames

In an analysis to detect putative open reading frames (ORFs), current ORF and gene database resources for predicting the presence of potential coding sequences were used. ORFs were defined as the regions between standard start (ATG) and stop (TAA, TAG, TGA) translation codons, with a minimum coding region size of three amino acids. 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. In a second GetORF analysis, an ORF was defined as a region between a start codon (ATG) and a stop codon (TAA, TAG, TGA) with a minimum size coding for three amino acids. The ORF search program, GetORF, within EMBOSS (European Molecular Biology Open Software Suite) was used and only putative ORFs spanning the newly created junction regions, with the potential for interruption during transformation, were taken into consideration.

For the pre-insertion locus sequences, GetORF identified a total of 24 putative ORFs; 19 interrupted ORFs were predicted between two stop codons (ORF-47 to ORF-65) and 5 interrupted ORFs were predicted between a start and a stop codon (ORF-74 to ORF-78). These predicted ORFs are shown in Figures 7 to 9 below.

For the transgenic locus sequences, GetORF identified 46 newly created ORFs defined between two stop codons (ORF-1 to ORF-46) and 8 newly created ORFs between a start and a stop codon (ORF-66 to ORF-73). These predicted ORFs are shown in Figures 10 to 14 below.

Potentially expressed genes
The software FGENESH (Softberry Inc., version 2.4) was used to predict gene structure and the potential for gene expression. This software is able to recognise multiple genes within a query sequence, predict exon and intron structures based on statistical analyses, and predict transcription start sites and poly-adenylation signals through homology with consensus sequences for plant nuclear genomes. Only genes crossing a junction, i.e. newly created genes and genes interrupted upon transformation, are reported.

For the pre-insertion locus sequences, no genes crossing the insertion points were predicted by FGENESH. One gene is predicted (Gene-2, Figure 7) in the 5′ end sequence of the translocated region that corresponds to the putative cysteine protease identified previously by the BLASTx analysis.

FGENESH predicted 5 genes in the transgenic locus sequences:
(i) Two genes corresponding to the *hppdPf W336* genes, both preceded by TPotp Y.
(ii) Two genes corresponding to the *2mepsps* genes, both preceded by TPotp C. The second copy of the *2mepsps* gene was predicted with a poly-adenylation signal in the 3′ flanking sequence, which would lead to a prolonged transcript spanning junctions 5 and 6 (Gene–1, Figure 12).
(iii) One gene (Gene-2, Figure 7) corresponding to a putative cysteine protease. This gene is positioned in the 5′ end sequence of the translocated region and was identified in the analysis of the pre-insertion locus.

Prediction of regulatory elements
TSSP is a pattern-finding tool used to search for core promoter (TATA-box) and enhancer sequences listed in the RegSite Database (version 4, Softberry Inc.). Promoters relevant for predicted newly created ORFs containing a start codon, newly created promoters and interrupted promoters are reported.

For the pre-insertion locus sequences, three promoter regions spanning an insertion point that are interrupted upon transformation are predicted (Table 11; Figures 7 and 9).

Table 11 Overview of predicted promoter regions in non transgenic *Glycine max* sequences

Fragment	Promoter region	Spanning insertion point	Position in fragment (bp)
JACK-WT1	Promoter-12	Insertion point 1	955 → 1251
JACK-WT1	Promoter-13	Insertion point 1	975 ← 1275
JACK-WT3	Promoter-14	Insertion points 4-5	965 ← 1209

For the transgenic locus sequences, TSSP predicted 11 relevant promoter regions (Table 12; Figures 10 to 14). Promoter regions 2, 3, 4, 6, 9 and 10 spanned a junction and are newly created. Promoter regions 1, 2, 4, 5, 7, 8 and 11 are in front of one or more predicted newly created ORFs containing a start codon.

Table 12 Overview of predicted promoter regions in FG72 transgenic locus sequences

Fragment	Promoter region	Position in fragment (bp)	Spanning junction	Promoter in front of newly created ORF(s)
FG72-TR	Promoter-1	301 → 600	/	ORF-66
	Promoter-2	1166 → 1464	Junction 1	ORF-66
	Promoter-3	1260 ← 1558	Junction 1	/
	Promoter-4	1197 → 1488	Junction 1	ORF-66
	Promoter-5	4158 ← 4434	/	ORF-69, ORF-70, ORF-71
	Promoter-6	9348 → 9647	Junction 4	/
	Promoter-7	13315 → 13615	/	ORF-67
	Promoter-8	17192 ← 17459	/	ORF-68
FG72-TL1	Promoter-9	965 ← 1248	Junction 7	/
	Promoter-10	970 → 1264	Junction 7	/
FG72-TL2	Promoter-11	679 → 975	/	ORF-72

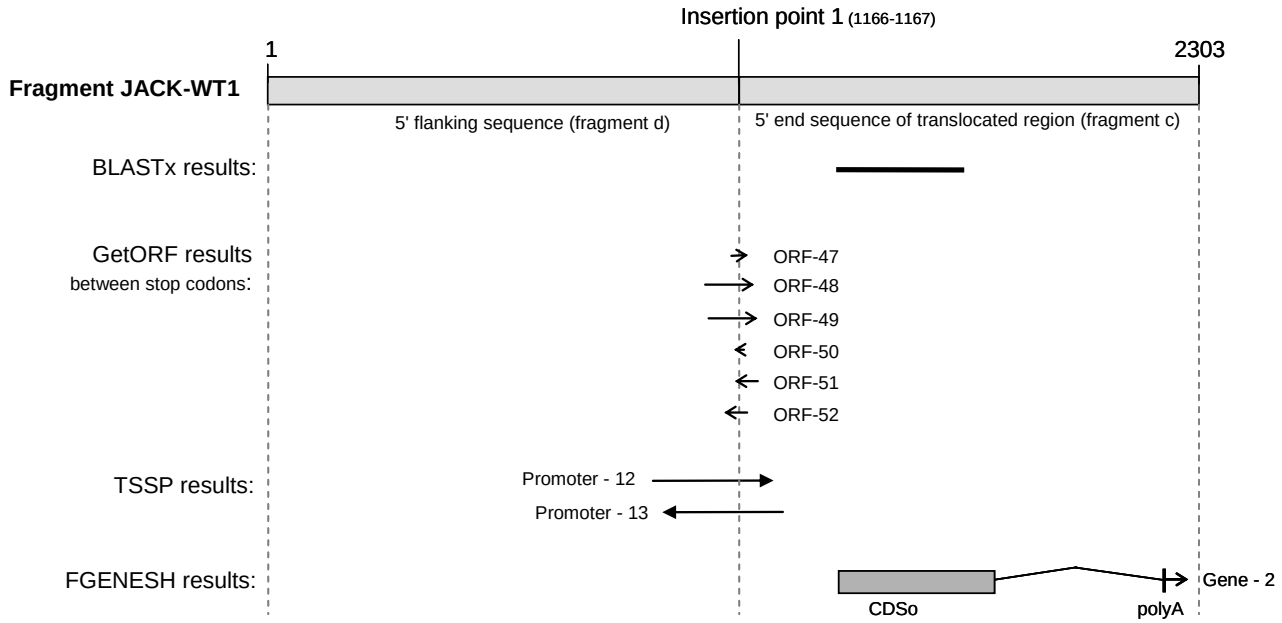


Figure 7 Putative open reading frames and promoter and enhancer sequences predicted in insertion point 1 of the FG72 pre insertion locus (fragment JACK-WT1)

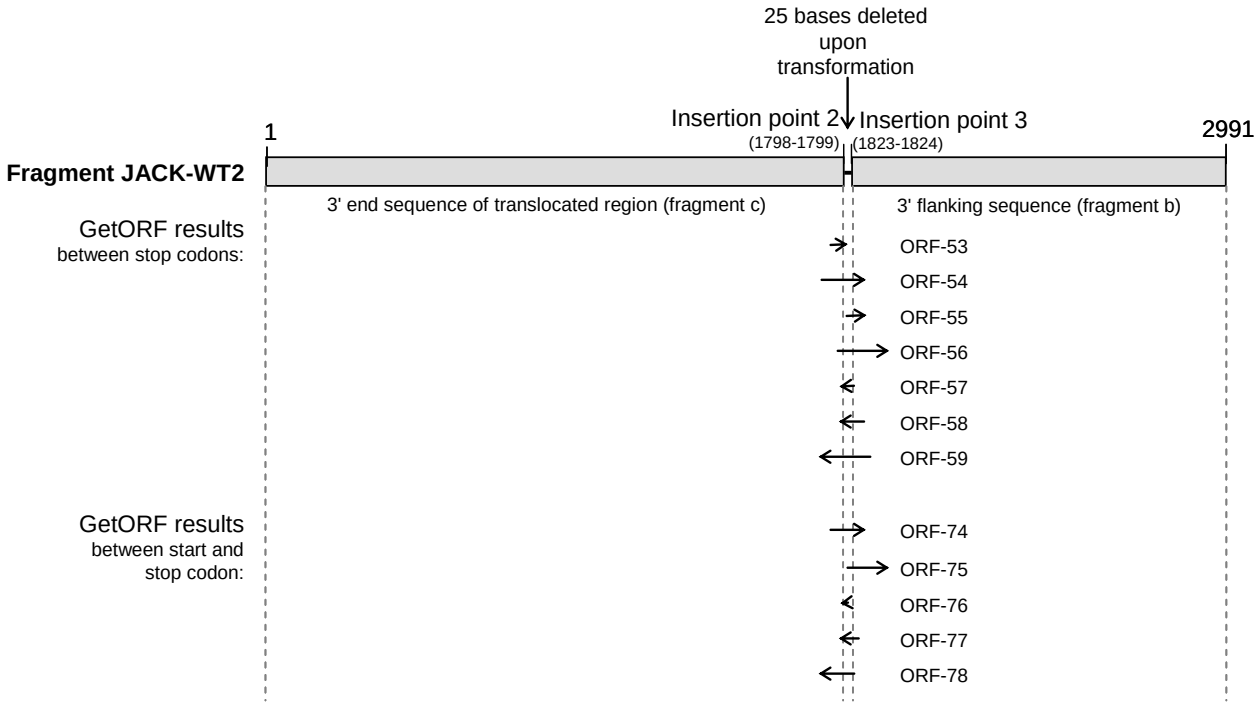


Figure 8 Putative open reading frames and promoter and enhancer sequences predicted in insertion points 2 and 3 of the FG72 pre insertion locus (fragment JACK-WT2).

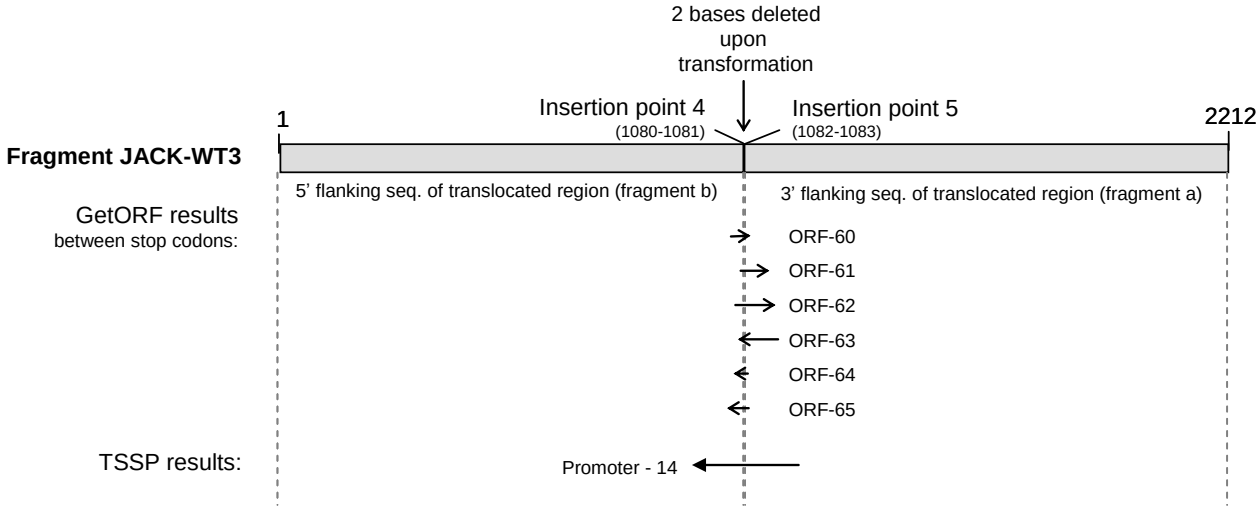


Figure 9 Putative open reading frames and promoter and enhancer sequences predicted in insertion points 4 and 5 of the FG72 pre insertion locus (fragment JACK-WT3).

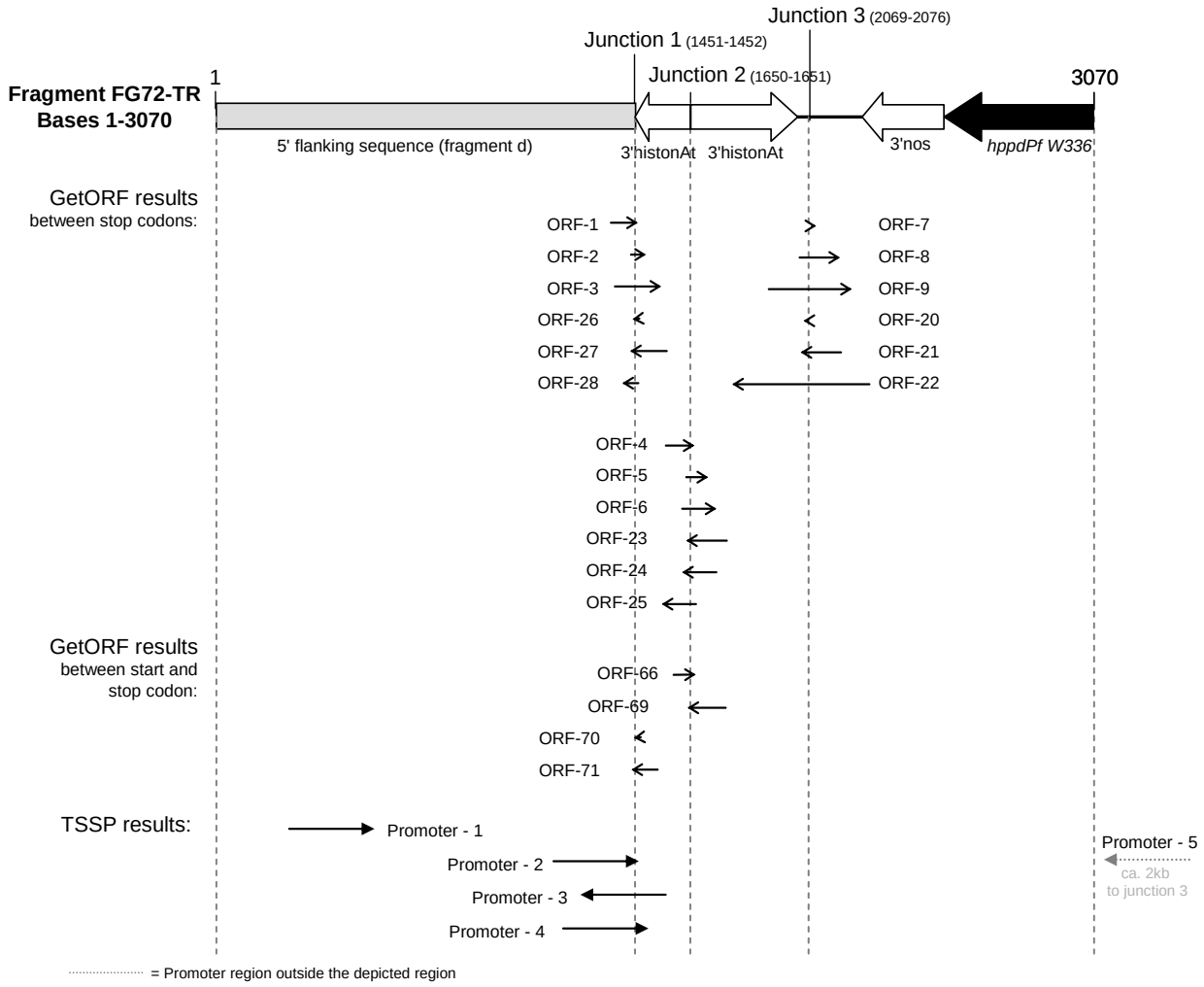


Figure 10 Putative open reading frames and promoter and enhancer sequences predicted junctions 1, 2 and 3 of the FG72 transgenic locus (fragment FG72-TR).

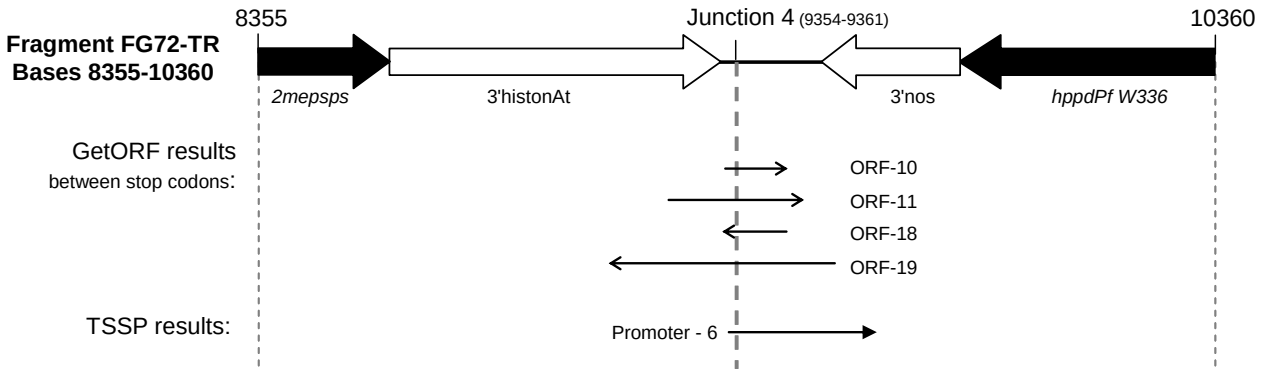


Figure 11 Putative open reading frames and promoter and enhancer sequences predicted junction 4 of the FG72 transgenic locus (fragment FG72-TR).

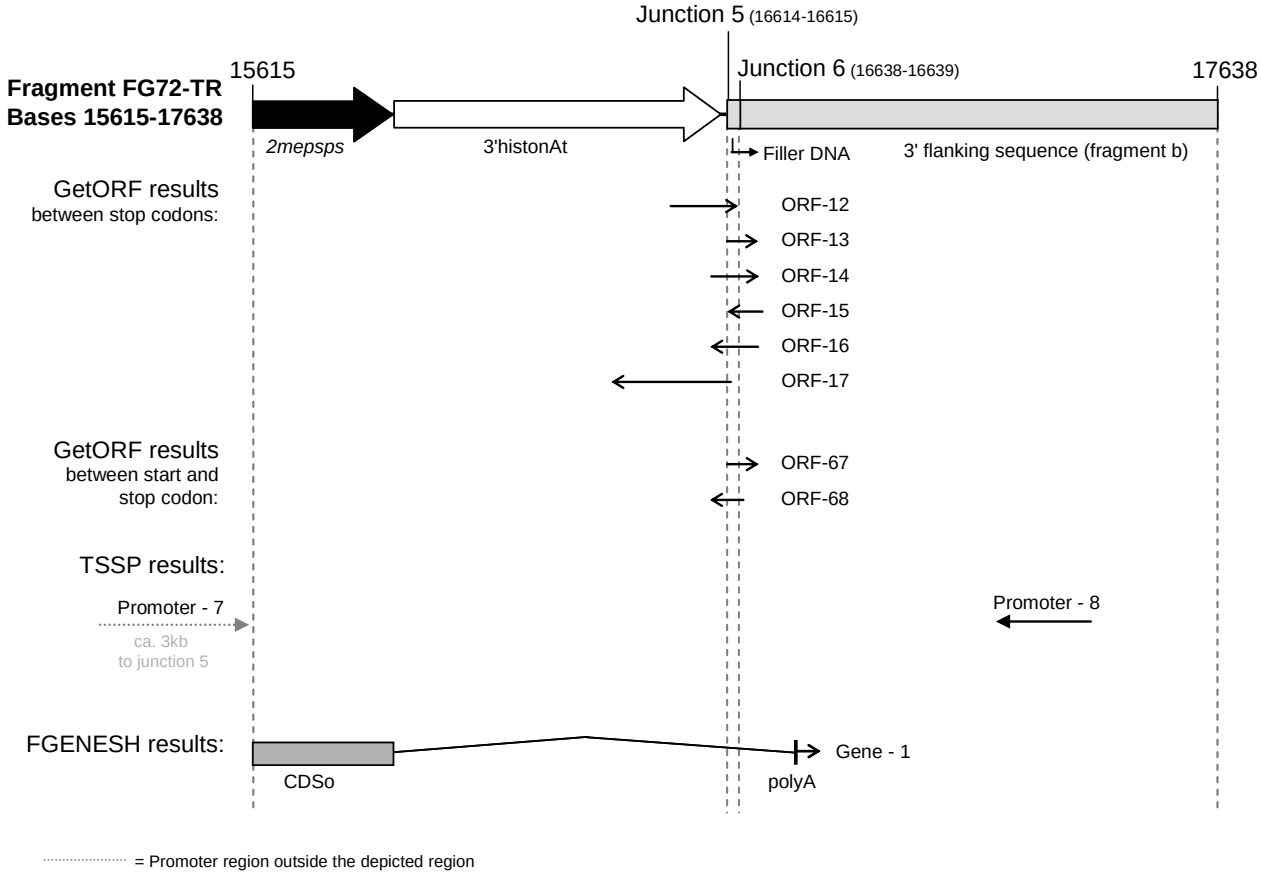


Figure 12 Putative open reading frames and promoter and enhancer sequences predicted junctions 5 and 6 of the FG72 transgenic locus (fragment FG72-TR).

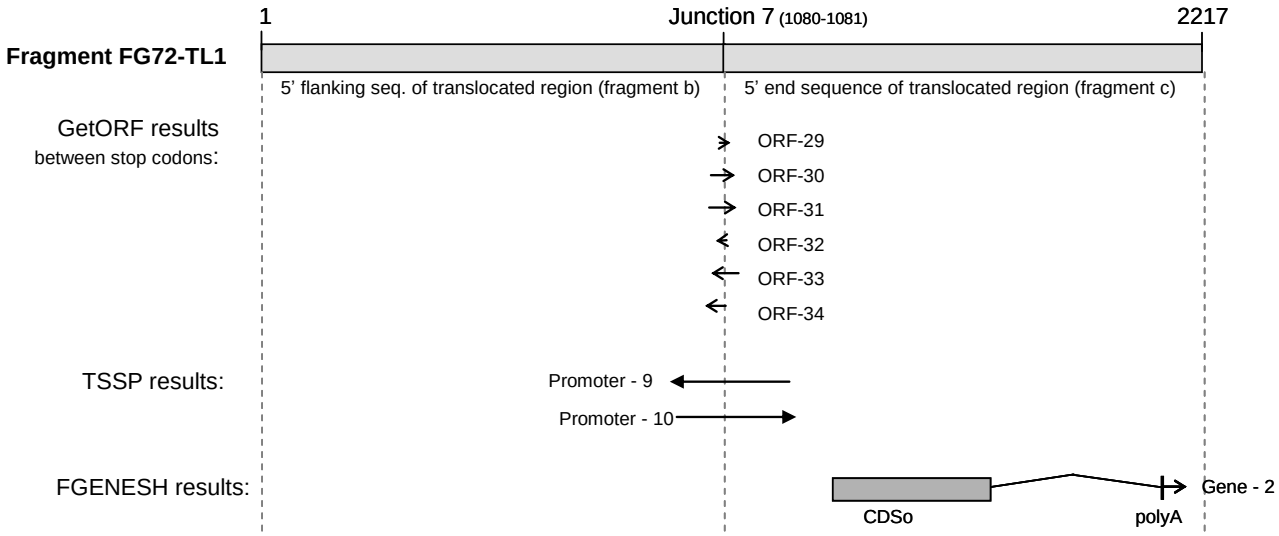


Figure 13 Putative open reading frames and promoter and enhancer sequences predicted junction 7 of the FG72 transgenic locus (fragment FG72- TL1).

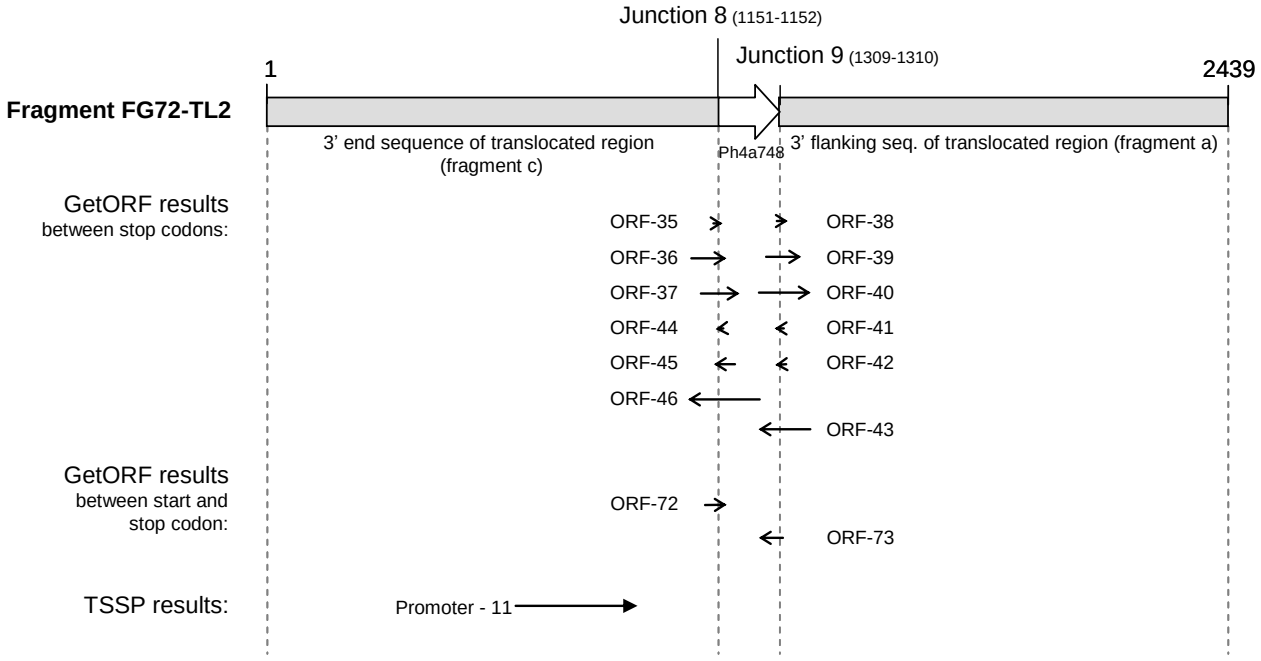


Figure 14 Putative open reading frames and promoter and enhancer sequences predicted junctions 8 and 9 of the FG72 transgenic locus (fragment FG72-TL2).

Prediction of ribosome binding sites

A consensus sequence has been determined for the ribosome binding site (RBS) based on a bioinformatics analysis of nucleotide frequencies at positions flanking the translation start codon of dicotyledon and monocotyledon plant genes (Joshi *et al.,* 1997). This sequence (aaaaaaaA(A/C)aATGGCtacta(c/t)ta) has been shown to be important for the initiation and efficiency of translation (Gallie *et al.,* 1987). The -3 and +4 positions (where the A of ATG is +1) are considered as the most important in determining a favorable context of initiator ATG.

For the transgenic locus sequences, comparison of the ATG context sequences of the predicted ORFs (ORFs 66 to 73) with the RBS consensus sequence was undertaken (Verhaeghe, 2009d; Appendix 5). Most of the essential nucleotides for the RBS are absent for all of these predicted ORFs, indicating that their translation is very unlikely.

To summarise the results of the bioinformatics analysis presented above, the evidence indicates that the predicted newly created ORFs are unlikely to lead to the expression of unintended proteins. No newly created genes or genes interrupted upon transformation are predicted.

(e) A description of how the line or strain from which food is derived was obtained from the original transformant (i.e. provide a family tree or describe the breeding process).

Following transformation of the *G. max* genome with event FG72, T_0 plants were treated with herbicides containing glyphosate and isoxazole to select for the expression of the *2mepsps* and *hppdPf W336* genes. The surviving plants were then self-pollinated to generate T_1 seed. This process of line purification through herbicide treatment and selfing to generate seed was repeated up to generation T_9 (see Figure 15 below). In the development of soybean varieties containing the FG72 event, T_6 plants were back-crossed into a conventional breeding line.

The breeding program for the development of event FG72 and its introgression into commercial soybean germplasm is demonstrated in Figure 15 below. Table 13 describes the FG72 generations used for analysis and the associated reports describing these studies.

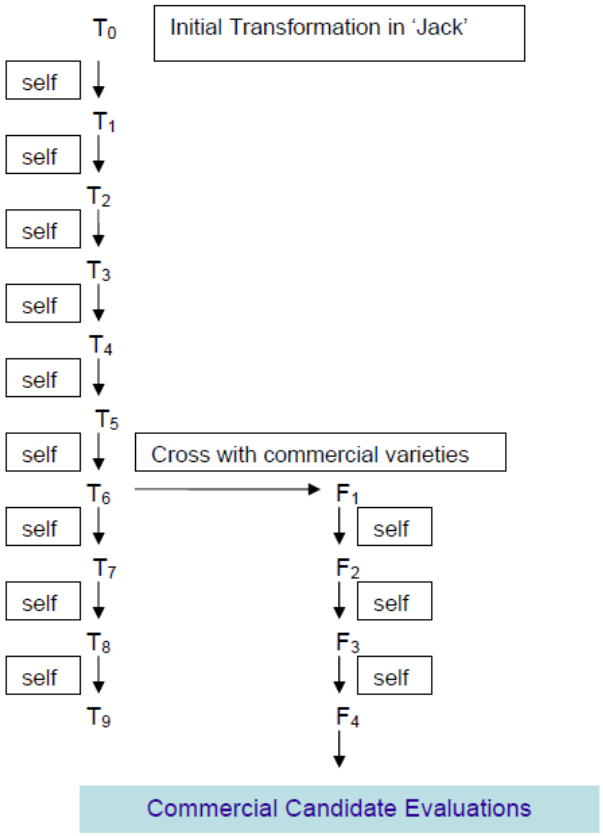


Figure 15 FG72 breeding program

Table 13 Generations used for analysis of event FG72

	Generation	Relevent Section and Reference
Mendelian inheritance analysis	T2, F2	Section 2.3(f)(i); Appendix 6
Structural stability	T2, T7, T9, F4	Section 2.3(f)(i); Appendix 6
Vector backbone	T7	Section 2.3(d)(i); Appendix 4
Insert characterization	T7	Section 2.3(d)(i),(ii)(iii),(iv); Appendix 2, 3
Protein expression *in planta*	T7	Section 3.2(c); Appendix 10
Grain nutrient composition	T8	Section 3.5(a), 4.1; Appendix 8
Protein expression in grain	T8	Section 3.2(c); Appendix 12
Chicken feeding study	T8	Section 4.2; Appendix 31
Full DNA Sequence	F2	Section 2.3(d)(iii); Appendix 3

(f) Evidence of the stability of the genetic changes, including:

(i) The pattern of inheritance of the transferred gene(s) and the number of generations over which this has been monitored; and

To demonstrate Mendelian inheritance of the FG72 transgenic locus, T_1 seed was harvested from self-pollinated plants of the primary transformation event (T_0) and planted in progeny rows in field trials conducted in 2001 in Puerto Rico. Three separate blocks were planted and sprayed with 0, 2, or 4 kg/ha glyphosate herbicide. Seed (T_2) was harvested from plants demonstrating the desired level of tolerance to the herbicide. The T_2 seed was then planted in six progeny rows in trials conducted in the US in 2002. Following glyphosate herbicide application, surviving plants were evaluated. Of the 172 individual plants, 124 were herbicide tolerant and 48 were sensitive to the herbicide.

In the progeny rows, the expected ratio for the inheritance of a single locus with herbicide tolerance is 1 fully tolerant row to 2 partially tolerant rows. For the individual plants, the expected ratio is 3 tolerant plants for each sensitive plant. Chi square analysis of the row segregation data (fully or partially tolerant) and of individual plants within rows (tolerant or sensitive) demonstrated Mendelian inheritance consistent with the inheritance of a single locus. This analysis is detailed in Table 14 below.

Table 14 Segregation data for progeny rows and individuals of self-pollinated soybean event FG72 treated with glyphosate herbicide

Parents and zygosity for the FG72 locus	Progeny	Fully Tolerant[a] Rows/	Partially Tolerant Rows/	Expected Ratio	χ^2 calculated[b]
		Tolerant Plants	Sensitive Plants		
Tolerant T_1 progenies of the self-pollinated T_0 transformants (1/4 FG72/FG72 ; 2/4 FG72/-)	T_2 Rows	1	5	1 to 2	0.485
Hemizygous T_1 plants (FG72/-) resulting in the partially resistant rows	T_2 Individual Plants	124	48	3 to 1	0.194

Table notes:
[a] Based upon survival to herbicide (glyphosate) application.
[b] Assumes a one locus model. There was no significant difference (p=0.05) for the χ^2 goodness-of-fit test for the hypothesis of one locus. To reject the null hypothesis, the χ^2 value must be greater than 3.84, with one degree of freedom.

Continued selection of herbicide tolerant plants was performed until a homozygous FG72 line was obtained. Plants from the sixth generation were crossed with conventional soybean breeding lines in the introgression program (see Figure 15 in Section 2.3(e) above). The resulting F_1 hybrid plants were grown to maturity and the F_2 seed was planted. Leaf samples were collected from 901 F_2 plants and analyzed using PCR probes designed to identify the zygosity of the FG72 insert. The expected ratio of 1:2:1 for a single insertion segregating by the rules of Mendel was observed. The results of this analysis are presented in Table 15 below.

Table 15 Segregation data for individual F_2 progeny of the FG72 x conventional line cross using zygosity PCR

ZYGOSITY OF THE FG72 LOCUS	PCR RESULT[A]	TOTAL PLANTS	RATIO	EXPECTED RATIO
Homozygous (nul/nul)	NN	212	0.24	0.25
Heterozygous (FG72/nul)	PN	471	0.52	0.5
Homozygous (FG72/FG72)	PP	218	0.25	0.25
		901		χ^2 value[B] = 0.172

[A] N = negative, P = positive

[B] Assumes a one locus model. There was no significant difference (p=0.05) for the χ^2 goodness-of-fit test for the hypothesis of one locus. To reject the null hypothesis, the χ^2 value must be greater than 3.84, with one degree of freedom.

(ii) The pattern of inheritance and expression of the phenotype over several generations and, where appropriate, across different environments.

The structural stability of soybean event FG72 was investigated in different genetic backgrounds, over different generations, and under different environmental conditions using Southern blot analysis (Verhaeghe, 2009e; Appendix 6).

As shown below in Table 16, the stability of the transgenic locus was tested using different *G. max* varieties and breeding lines representing three genetic backgrounds, three different generations and four different environments.

The isolated DNA was digested with the restriction enzyme *Hin*dIII, and hybridized with the Ph4a748B probe, followed by the T-DNA probe that targeted the transfer DNA. This analysis revealed the expected profile of hybridization fragments in all samples analysed, demonstrating the structural stability of the transgenic locus at the genomic level in the different generations, environments and genetic backgrounds tested. The hybridisation strategy used in the structural stability analysis is shown in Figure 16, and results of the Southern blot analysis are presented in Table 16 below.

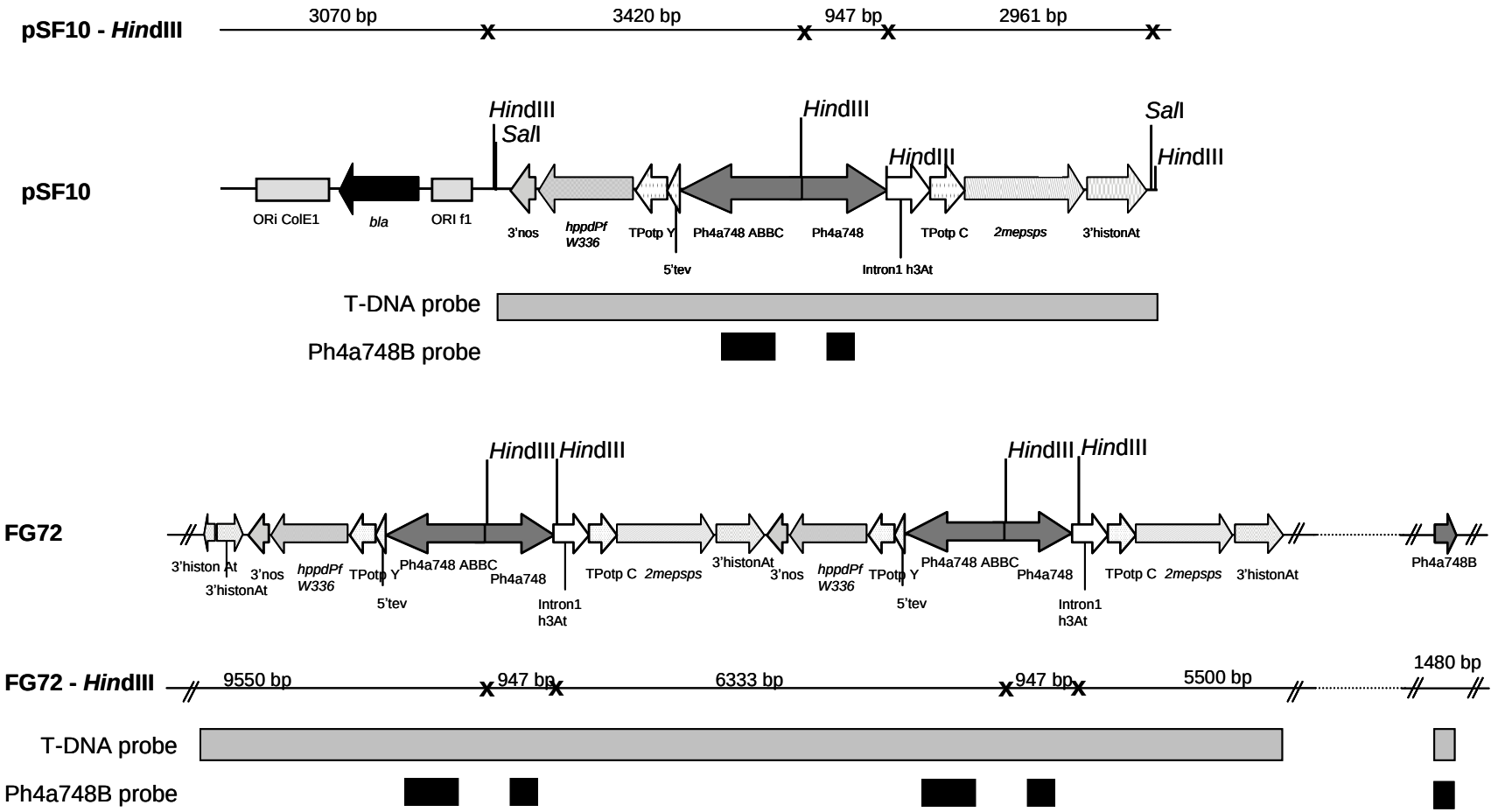


Figure 16 Southern blot analysis strategy for the structural stability analysis of the FG72 transgenic locus

Table 16 Expected and obtained hybridisation fragments in the genetic stability Southern blot analysis of soybean event FG72

Samples	Condition tested	Number of plants	Expected T-DNA or plasmid fragment sizes	Fragment description	Ph4a748B probe		T-DNA probe	
					Exp.	Obt.	Exp.	Obt
FG72 - *Hin*dIII	Location Adel, Iowa	22						
	Location Osborn, Missouri	22						
	Location Fithian, Illinois	21	9550 bp [d]	5' integration fr.	Yes	Yes	Yes	Yes
	Location Sharpsville, Indiana	16	947 bp	internal fragment	Yes	Yes	Yes	Yes
	Background 3068115-48 x Jack	21	6333 bp	internal fragment	Yes	Yes	Yes	Yes
	Background 3066617-48 x Jack	22	5500 bp [d]	3' integration fr.	No	No	Yes	Yes
	Generation T2	22	1480 bp [d]	3' junction translocation	Yes	Yes [a]	Yes [b]	No [c]
	Generation T7	13						
Non transgenic Jack - *Hin*dIII	Non-transgenic variety Jack		/	Negative control	/	/	/	/
Non transgenic Jack - *Hin*dIII + pSF10 - *Hin*dIII	Non-transgenic variety Jack + equimolar amount pSF10		3420 bp	positive control	Yes	Yes	Yes	Yes
			947 bp	positive control	Yes	Yes	Yes	No [c]
			2961 bp	positive control	No	No	Yes	Yes
			3070 bp	positive control	No	No	No	No

[a] In some hybridizations this fragment is very weak, but present for all samples
[b] The overlap between the probe and the fragment can be too small to visualize this fragment
[c] Not always visible after hybridization with T-DNA probe but presence is confirmed after hybridization with probe Ph4a748B.
[d] Expected fragment sizes as determined in the detailed insert characterization study (Appendix 2)

2.4 Information on the Labelling of the GM Food

(a) Information on whether novel DNA or protein is likely to be present in final food.

To enable an assessment of the potential exposure of humans and animals to the recombinant proteins expressed in soybean event FG72, seed samples and products derived from them were analysed for the content of the 2mEPSPS and HPPD protein.

The primary food product derived from soybean for human consumption is soybean oil. In the course of processing soybeans to produce refined vegetable oil of food grade quality, all protein components are destroyed by the high temperature and pressure of the screw pressing, or separated by extraction with a non-polar solvent and destroyed by the temperature of the solvent recovery. Remaining traces of protein in the crude oil are removed in the alkali treatment and deodorization steps of oil refining. The removal of proteins in soybeans derived from event FG72 was confirmed by the absence of any detectable 2mEPSPS and HPPD protein amounts in crude and food grade oil produced from them (Robinson, 2009; Appendix 7). Consequently, an intake of these recombinant proteins is not possible via soybean food grade oil or products containing this oil quality.

Using the "Pulses" information provided in the GEMS/Food Regional Diets publication of FAO/WHO, the potential intake of recombinant protein via the consumption of whole soybeans as pulses and, additionally, for all kind of oilseeds except groundnuts, was calculated. The highest recombinant protein contents were used for the calculation. These calculations were based on worst-case scenarios taking the highest recombinant protein amounts determined in the soybean commodity, assuming that all commercial soybean seeds taken to produce food or animal feed would be the double-herbicide-tolerant soybean event FG72, and that all kinds of oilseeds with the exception of groundnuts consumed by humans would be soybean seeds.

The predicted 2mEPSPS protein intake via whole soybeans as pulses was between 0.245 and 11.0 µg per person per day for the various regional diets. Based on the consumption of all kinds of oilseeds (except groundnuts), the 2mEPSPS protein intake was between 1.23 and 12.5 µg per person per day. The HPPD protein intake via whole soybeans as pulses for the various regional diets was between 0.126 and 5.67 µg per person per day. Based on the consumption of all kinds of oilseeds (except groundnuts), the HPPD protein intake was between 0.63 and 6.43 µg per person per day. Overall, the calculated predicted dietary intakes for the two recombinant proteins are very low, and the real per capita daily intake figures for both proteins are expected to be significantly lower than 5.0 µg per person per day for both proteins. Table 17 shows the predicted dietary intake of the 2mEPSPS and HPPD W336 proteins in different regional diets, with the full analysis detailed in Oberdoerfer (2009; Appendix 8).

Table 17 Predicted dietary intake of 2mEPSPS and HPPD W336 proteins

Parameter	Regional Diets				
	Europe [a]	Latin America	Middle East	Far East	Africa
Consumption of whole soybeans as pulses in gram per person per day	0.1 [b]	0.1 [b]	4.5	2	0.5
Consumption of all kind of oilseeds (except groundnuts) in gram per person per day	3.1	0.5	5.1	1.2	3.1
Highest 2mEPSPS protein content in FG72 seeds is 2.45 µg/g fw [c]					
Predicted daily intake of 2mEPSPS protein via whole soybeans (µg per person per day)	0.245	0.245	11	4.9	1.23
Predicted daily intake of 2mEPSPS protein via all kind of oilseeds (µg per person per day)	7.6	1.23	12.5	2.94	7.6
Highest HPPD protein content in FG72 seeds is 1.26 µg/g fw [c]					
Predicted daily intake of HPPD protein via whole soybeans (µg per person per day)	0.126	0.126	5.67	2.52	0.63
Predicted daily intake of HPPD protein via all kind of oilseeds (µg per person per day)	3.91	0.63	6.43	1.51	3.91

[a] The European diet includes countries with European-type diets, such as Australia, Canada and the USA
[b] Since whole soybean are not consumed in Europe or Latin America a default value of 0.1 g per person per day was assigned.
[c] Highest 2mEPSPS and HPPD contents in soybeans taken from Table 3.1.1, Appendix 8.

Soybean seeds and processed commodities are also used in animal feed. Therefore, the contribution of soybean event FG72 to animal feed was also calculated. The US-EPA residue chemistry test guidelines (EPA, 1996) list seven different plant fractions to be included in feedstuff including: seeds, aspirated grains fractions, forage, hay, silage, hulls and meal. The maximum theoretical 2mEPSPS and HPPD protein amounts were calculated for the case that FG72 seeds, hulls or meal was used to prepare the animal diets (Table 18).

Table 18 Contribution of soybean commodities to animal feed (US-EPA, 1996) and 2mEPSPS and HPPD protein amounts in the FG72 commodities

Agricultural commodity	Maximum contribution to animal diets (%)				Maximum 2mEPSPS protein in µg/g fw	Maximum HPPD protein in µg/g fw
	Beef cattle	Dairy cattle	Poultry	Swine		
Seed	15	15	20	25	2.45 [a]	1.26 [A]
Hulls	20	20	20	-[D]	0.50 [b]	0.96 [B]
Meal	15	15	40	25	< 0.020	ND [D]

[A] Results from 2mEPSPS and HPPD protein analyses of seeds taken from Table 4.1.1, Appendix 8.
[B] Results from 2mEPSPS and HPPD protein analyses of hulls and meal taken from Table 4.2.1, Appendix 8.
[C] Soybean hulls are not used to prepare diets for swine
[D] Not detected

Taking the highest 2mEPSPS and HPPD protein contents determined in seeds, hull and meal from glyphosate and IFT treated FG72 plants (Table 18), the maximum theoretical amounts of the recombinant proteins and the percentage of the proteins in the livestock diet are presented below in Table 19. The maximum theoretical concentrations of 2mEPSPS and HPPD protein were 0.61 µg/g fw or 6.1×10^{-5} % of the diet and 0.32 µg/g fw or 3.2×10^{-5} % of the diet, respectively, if soybean seeds were used to prepare animal feed for swine.

Table 19 Calculation of the maximum theoretical 2mEPSPS and HPPD protein amounts in animal diets produced with FG72 seeds or hulls

Agricultural commodity	2mEPSPS in µg/g animal diet				Percentage of 2mEPSPS in the animal diet			
	Beef cattle	Dairy cattle	Poultry	Swine	Beef cattle	Dairy cattle	Poultry	Swine
Seeds	0.37	0.37	0.49	0.61	3.7×10^{-5}	3.7×10^{-5}	4.9×10^{-5}	6.1×10^{-5}
Hulls	0.1	0.1	0.1	-	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	-
Agricultural commodity	**HPPD in µg/g animal diet**				**Percentage of HPPD in the animal diet**			
	Beef cattle	Dairy cattle	Poultry	Swine	Beef cattle	Dairy cattle	Poultry	Swine
Seeds	0.19	0.19	0.25	0.32	1.9×10^{-5}	1.9×10^{-5}	2.5×10^{-5}	3.2×10^{-5}
Hulls	0.19	0.19	0.19	-	1.9×10^{-5}	1.9×10^{-5}	1.9×10^{-5}	-

(b) Detection methodology for the GM food suitable for analytical purposes.

The two herbicide tolerance traits can be detected either on a molecular genetic level or on a protein biochemical level. These methods enable the detection and quantification of minute quantities ofDNA derived from the event FG72 and are specific enough to detect event-specific DNA within complex DNA pools.

The molecular genetic detection can be performed with a PCR based method to confirm the presence of the introduced material in the soybean plant material. The transgene PCR reaction will only amplify a product from the inserted DNA, making it possible to distinguish between nontransgenic and transgenic samples. There are specific protocols for each transgene within each type of plant. An individual protocol usually requires optimization to account for differences between labs, matrices, or reagents. This optimization is especially important when performing multiplex reactions. Some *loci* are more efficiently amplified than others due to base composition, length of product, and secondary structure. In multiplex reactions, the more efficiently amplified *loci* compete more effectively for available reaction components, and will negatively influence the yield of product from less efficient *loci*, making them less visible or undetectable. It is important to obtain reaction conditions that amplify equimolar quantities of both the enodogenous and transgenic sequences in a known transgenic DNA sample. An event-specific DNA polymerase chain reaction (PCR) assay can be provided to FSANZ on request.

Detection of the proteins expressed by the FG72 event can be achieved using standard immunoassay methodology such as the enzyme-linked immunosorbent assay (ELISA). The use of this methodology to detect the EPSPS and HPPD proteins in various plant tissues is described below in Section 3.2(c) and in Massengill (2009; Appendix 9). A pre-commercial ELISA kit is available for EPSPS-type proteins. This ELISA was validated in a soybean grain matrix for the 2mEPSPS protein. A pre-commercial ELISA kit is available for the HPPD W336 protein and was validated in a soybean grain matrix. These methods are based on the specific interaction between antibody and antigen (double antibody sandwich format) and have been validated for raw agricultural commodities of event FG72 (Massengill, 2009; Appendix 9).

Another protein detection method is the lateral flow strip device (LFS) which allows qualitative detection of the introduced protein, and can be performed under field and/or lab conditions. The LFS contains specific antibodies to the protein of interest and employs the sandwich format to detect the protein of interest. If the protein of interest is present in the sample, two bands appear on the strip (the control band and the positive band). The presence of the control band alone indicates a negative sample. LFS devices are available commercially for the 2mEPSPS protein and a LFS test for the HPPD W336 protein is being developed by MSTech.

Part 3 Information Related to the Safety of the Genetically-Modified Food

3.1 Information on Antibiotic Resistance Marker Genes (if used)

(a) Information on the clinical and veterinary importance, if any, in Australia and New Zealand of the antibiotic to which any transferred antibiotic resistance genes confer resistance.

Not applicable. The expression of the *2mepsps* and *hppdPf W336* genes were used as selective markers in the development of lines containing event FG72.

(b) Information on whether the presence in food of the enzyme or protein encoded by the antibiotic resistance marker gene would compromise the therapeutic efficacy of the orally administered antibiotic.

Not applicable.

(c) Information on the safety of the gene product.

Not applicable.

(d) If the new GM organism is a micro-organism, information on whether it will remain viable in the final food.

Not applicable.

3.2 The Characterisation of Novel Proteins or Other Novel Substances

(a) A full description of the biochemical function and phenotypic effects of all novel substances (e.g. a protein or an untranslated RNA) that could potentially be expressed in the new GM organism, including those resulting from the transfer of marker genes.

<u>2mEPSPS Protein</u>

The enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) is involved in the shikimic acid pathway. This pathway produces an important branch point intermediate, chorismate, for aromatic amino acid and aromatic metabolites biosynthesis in plants and microorganisms. The shikimate pathway exists exclusively in plants and microorganisms including fungi. In contrast, mammals, fish, birds, reptiles, and insects must derive their aromatic compounds from their diet. For this reason, there has been interest over the last three decades in the shikimate pathway enzymes as potential targets for non-toxic herbicides and anti-microbial compounds.

EPSPS is the sixth enzyme of the shikimate pathway and it has been shown that EPSPS enzymes are ubiquitous in nature and are present in foods derived from plant and microbial sources. The mode of action of glyphosate [N- (phosphonomethyl) glycine], a simple amino acid analog, was determined to be the selective inhibition of EPSP synthase (EPSPS; EC 2.5.1.19), (Steinrücken and Amrhein, 1980). The reaction catalyzed by EPSPS is the reversible transfer of the phosphoenolpyruvate (PEP) to shikimate-3-phosphate (S3P), leading to the formation of 5-enolpyruvyl-3-shikimate phosphate (EPSP). Substrate binding to the enzyme is sequential, with S3P binding first, followed by PEP (Boocock and Coggins, 1983). The reaction catalyzed by EPSPS proceeds via C-O bond cleavage of PEP (Walsh *et al.*, 1996). In higher plants, EPSPS is synthesized from a nuclear gene in the form of a cytoplasmic precursor, then imported into the plastids where it accumulates in its mature form (Kishore and Shah, 1988; Forlani *et al.*, 1994; Lebrun *et al.*, 1997). Transit peptides are typically cleaved from the mature protein following delivery to the plastids (Della-Cioppa *et al*.,1986).

Since the 1980s, several attempts have been made to identify and characterize glyphosateinsensitive EPSPS enzyme variants from various organisms with the ultimate aim to engineer glyphosate tolerance in crop plants (Kishore and Shah, 1988). Lebrun *et al.* (1997) selected a double mutant gene from maize, which when fused to a chimeric optimized transit peptide, generates optimal glyphosate tolerance in various crops, with no pleiotropic effects: the *2mepsps* gene encoding the 2mEPSPS protein. The *2mepsps* gene has been introduced as the source of glyphosate tolerance in the maize transgenic event GA21 which has been approved by different agencies worldwide for environment, food and feed (OECD unique identifier MON-ØØØ21-9) (AGBIOS, 2006). Recently, glyphosate tolerance was also achieved in rice by mutagenesis of the rice *epsps* gene (Zhou *et al.*, 2006).

The *2mepsps* gene was generated by introducing mutations into the wt *epsps* gene from maize (*Z. mays* L.), leading to a double mutant EPSPS protein with two amino acid substitutions (2mEPSPS). These modifications confer to the protein a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of the herbicide. Therefore, the plants bearing this gene become tolerant to glyphosate herbicides (Lebrun *et al.*, 1997).

HPPD W336 Protein

The biochemical pathways in which HPPD is involved differ between plants and non-photosynthetic organisms. In bacteria and animals it merely serves catabolic purposes by catalyzing the first committed step in tyrosine degradation that in the end yields energetically exploitable glucogenic and ketogenic products (Brownlee *et al*., 2004; see Figure 17 below). In plants, HPPD is involved in several anabolic pathways, and its reaction product homogentisate (2,5-dihydroxyohenylacetate, HG) is the aromatic precursor of vitamin E, a membrane-associated antioxidant, and of plastoquinone (Fritze *et al*., 2004).

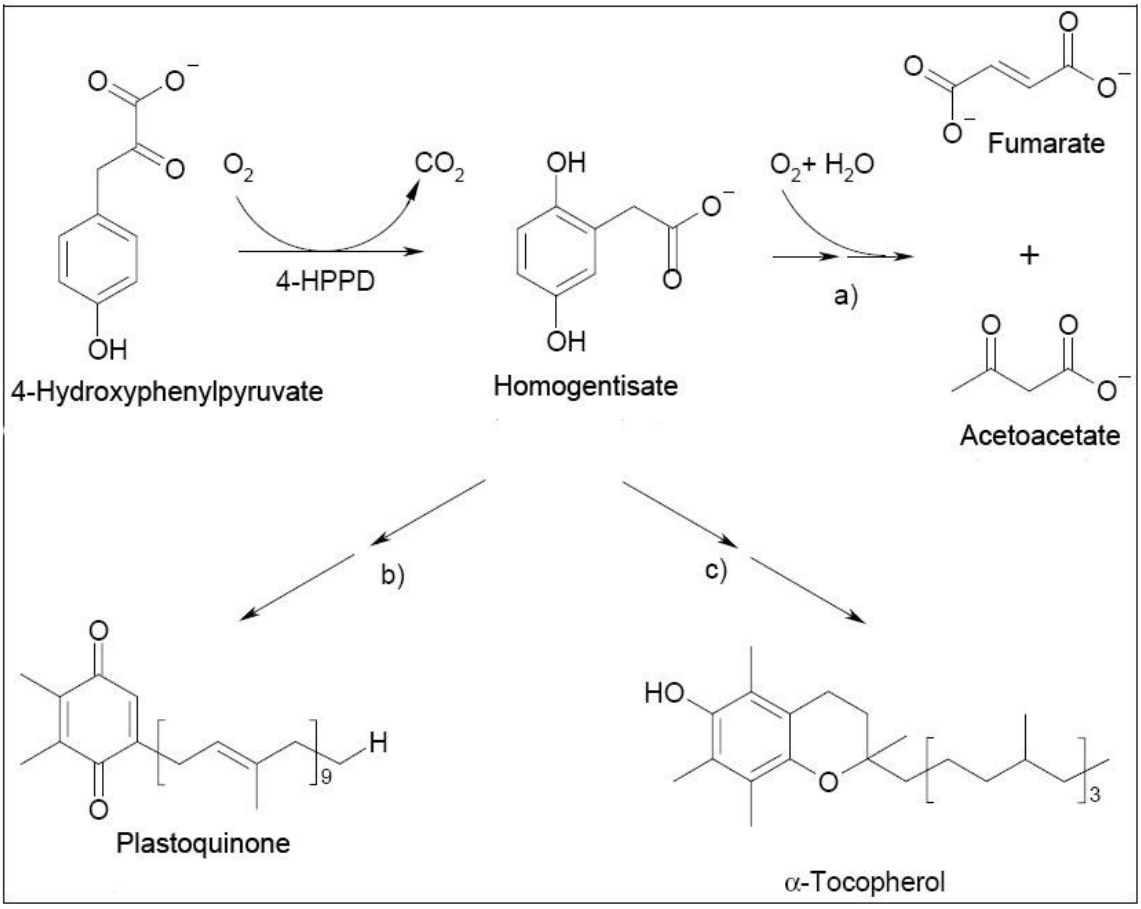


Figure 17 Biochemical pathways of HPPD proteins: a) catabolism of tyrosine, b) biosynthesis of plastoquinone (plants) c) biosynthesis of tocopherol (plants)

HPPD is ubiquitous in nature across all kingdoms: bacteria, fungi, plants and animals including mammals. HPPD amino acid sequences have been determined in bacteria such as *Streptomyces avermitilis* (Accession number Q53586), in fungi such as *Aspergillus fumigatus* (Accession number Q4WPV8), in plants such as *Arabidospis thaliana* (Accession number P93836), and in animals such as *Caenorhabditis elegans* (Accession number Q22633), mouse (*Mus musculus*, Accession number P49429), and human (*Homo sapiens*, Accession number P32754). HPPD has also been characterized in organisms present in human food, such as carrot (*Daucus carota*, Accession number O23920), barley (*Hordeum vulgare* Accession number O48604), pig (*Sus scrofa*, Accession number Q02110) and cow (*Bos Taurus*, Accession number Q5EA20). HPPD proteins from diverse origins share a common structure, with properties consistent with toxicity or allergenicity. Although HPPD from *P. fluorescens* has a low overall percentage of amino acid sequence identity with HPPD from plant and mammals (21% and 29%, respectively), sequence alignments of HPPD from a wide variety of sources show that 27 residues among more than 350 residues are completely conserved in these and all other HPPD proteins (Yang *et al*., 2004).

In the development of soybean event FG72, several different HPPD variants were tested for their activity in the presence and absence of the inhibitor isoxaflutole. The HPPD W336 protein used in the soybean event FG72 is a mutated form of the HPPD isolated from *Pseudomonas fluorescens*. The resistance to isoxaflutole was introduced by changing one amino acid, a Glycine at position 336, with a tryptophane (Boudec *et al.*, 2001). This was achieved by changing the coding sequence of the 4-hydroxyphenylpyruvate dioxygenase of *Pseudomonas fluorescens* strain A32 (Sailland *et* a/., 2001). The resulting sequence of the mutated HPPD W336 protein consists of 358 amino acids as described in Table 20. The HPPD W336 protein has a theoretical molecular weight of 40 kDa. The protein encoded by the plant transformation vector is a fusion protein between the HPPD W336 protein and the transit peptide TPotp Y. The theoretical molecular weight of this fusion protein is 48 kDa. The transit peptide sequence is cleaved from the fusion protein upon transition of the protein to the chloroplast. In comparison to the wild type protein, the mutated form shows significantly less inhibition when treated with isoxaflutole. The interaction of the HDDP W336 protein and isoxaflutole herbicide is depicted in Figure 18.

Table 20 Overview of the HPPD protein

Protein name	Expression vector	Description	N° of Nucleic acid	N° of Amino Acid	MW (kDa)
HPPD	/	Native protein derived from *Pseudomonas*	1075	358	40
HPPD W336	C56	Mutated HPPD: G at position 336 changed to W	1075	358	40
TPotp Y + HPPD W336	pSF10	HPPD W336 fused to the optimized transit peptide, as inserted in transgenic plants.	1282	427	48

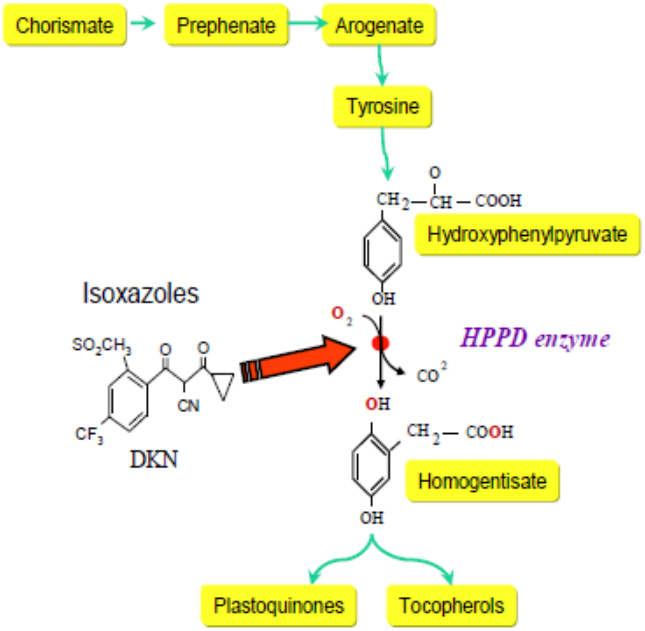


Figure 18 Interaction of HPPD and isoxazole herbicides

(b) The identification of any other novel substances (e.g. metabolites) that might accumulate on or in the GM organism as a result of the genetic modification, and their levels and site of accumulation.

As detailed in previous sections on the molecular characterisation of the soybean event FG72 (see Sections 2.3(d)(v), no novel genes apart from the *2mepsps* and *hppdPf W336* genes are expressed as a result of the genetic modification event.

(c) Data on the site of expression of all novel substances, particularly whether they are likely to be present in the edible portions of the organism, and levels of expression.

Protein quantification *in planta*

The 2mEPSPS and HPPD W336 proteins are expected to be present in all tissues during the life of FG72 plants. The expression levels of these two proteins in FG72 plants were determined in different tissues at different growth stages using ELISA analysis. Transgenic and wild type (non-transgenic) plants were grown under greenhouse conditions with leaf samples taken at three different growth stages (V4, V6 and V8), and stem and root samples were taken at two different growth stages (V4 and V8). Samples of the wild type soybean line (variety Jack) were also sampled at the same growth stages for the same tissues. Seeds were also collected from both transgenic and non-transgenic plants. This analysis is described by Habex and Debaveye (2009; Appendix 10).

The 2mEPSPS and HPPD W336 proteins were detected in all transgenic samples covering all growth stages and tissue types sampled. The protein expressions levels (mean ± SD and range), expressed per gram fresh weight and per gram dry weight, are summarised below in Table 21.

Protein quantification in processed fractions

A study was perfomed to determine the content of the 2mEPSPS and HPPD W336 proteins in processed fractions derived from soybean event FG72 (Robinson, 2009; Appendix 7). The soybean seed raw agricultural commodity was produced in one trial site in Iowa in 2008 (Kowite, 2009a; Appendix 11). The soybean plants were grown in three plots under conditions typical of production practices. One of the two transgenic event plots in the field trial was sprayed one time with the herbicides Isoxaflutole + Glyphosate (IFT+GLY). The other transgenic plot was untreated. The

soybean grain (whole soybean seeds) was harvested and processed for determination of 2mEPSPS and HPPD W336 protein content using ELISA. The protein analyte content (mean ± standard deviation), expressed in ng/g on a fresh weight and dry weight basis, are presented in Table 22 below.

Protein quantification in raw fractions

In another study, the content of 2mEPSPS and HPPD W336 proteins in the raw agricultural commodity (soybean seed) derived from FG72 plants was determined (Poe, 2009; Appendix 12). The seed used in this analysis was obtained from plants grown in ten US field trials described by Kowitt (2009b; Appendix 13) and analysed using ELISA. These plants were grown under typical conditions for soybean production. There were six plots with event FG72 at each trial site. Three of the six transgenic event plots in each field trial were sprayed one time with the test herbicide Isoxaflutole + Glyphosate (IFT + GLY). The other three plots were untreated. The protein content (range, mean ± standard deviation) expressed in ng/g and percentage on a fresh weight and dry weight bases, are presented in Table 23 below.

Table 21 HPPD W336 and 2mEPSPS contents per gram fresh and dry weight of the different FG72 soybean tissues and growth stages

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 ± 22.6	12.8 – 96.0	79.1 ± 29.6	39.2 – 136	437 ± 163	216 – 753
	V8	4.69 ± 1.87	2.00 – 8.91	27.3 ± 10.9	11.6 – 51.8	115 ± 38.2	60.5 – 203	668 ± 222	351 – 1182
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 ± 70	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 ± 23	76.1 – 152
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 – 54.6
	V8	0.84 ± 0.50	0.20 – 1.64	6.42 ± 3.82	1.51 – 12.5	5.75 ± 2.31	2.62 – 10.7	43.7 ± 17.6	20.0 – 81.2
Seed	NA	1.27 ± 0.42	0.71 – 2.68	1.41 ± 0.47	0.79 – 2.96	2.37 ± 0.75	1.34 – 3.74	2.62 ± 0.83	1.48 – 4.13

Table 22 Protein analyte content in processed fractions

Matrix	Treatment	2mEPSPS		HPPD W336	
		Fresh Weight (ng/g) ± SD	Dry Weight (ng/g) ± SD	Fresh Weight (ng/g) ± SD	Dry Weight (ng/g) ± SD
Hull Material	Unsprayed	552 ± 56	624 ± 63	941 ± 64	1064 ± 72
	Sprayed	501 ± 16	563 ± 17	957 ± 42	1077 ± 47
Protein Isolate	Unsprayed	483 ±31	493 ± 32	627 ± 50	640 ± 51
	Sprayed	1020 ± 41	1042 ± 42	1078 ± 17	1101 ± 17
Untoasted Meal	Unsprayed	Not Detected		Not Detected	
	Sprayed	<LOQ		Not Detected	
Toasted Meal	Unsprayed	Not Detected		Not Detected	
	Sprayed	Not Detected		Not Detected	
Crude Oil	Unsprayed	Not Detected		Not Detected	
	Sprayed	Not Detected		Not Detected	
Crude Lecithin	Unsprayed	Not Detected		Not Detected	
	Sprayed	Not Detected		Not Detected	
Refined Oil	Unsprayed	Not Detected		Not Detected	
	Sprayed	Not Detected		Not Detected	
Refined Bleached Deodorized Oil	Unsprayed	Not Detected		Not Detected	
	Sprayed	Not Detected		Not Detected	

Table 23 Protein analyte content in soybean grain

Protein	Treatment		Fresh Weight (ng/g)	Dry Weight* (ng/g)	% Crude Protein*
2mEPSPS	Unsprayed	Range	364 – 5790	1500	0.00040
		Mean ± SD	1360 ± 1080		
2mEPSPS	Sprayed	Range	326 – 2690	1300	0.00034
		Mean ± SD	1180 ± 589		
HPPD W336	Unsprayed	Range	455 – 1320	936	0.00024
		Mean ± SD	846 ± 183		
HPPD W336	Sprayed	Range	411 – 1313	887	0.00023
		Mean ± SD	802 ± 207		

*Dry weight and % crude protein analyte amounts were determined using the average of four individual results per sample, therefore no standard deviation or range is given for these amounts.

(d) Information on whether any newly expressed protein has undergone any unexpected post-translational modification in the new host.

Glycosylation is one of the principal co-translational and post-translational modifications of various membrane-bound and secreted proteins. Attachment of saccharides to target proteins is known to promote proper protein folding and confer enhanced protein stability. Some allergens are glycosylated, raising the possibility that the glycosyl groups may contribute to their allergenicity

(Jenkins *et al*., 1996). However, it is also well recognized that many allergens are not glycosylated, and a large number of non-allergens are glycoproteins, indicating that glycosylation is neither necessary nor sufficient for allergenicity. While N-glycosylation does not provide definitive evidence in terms of protein safety, it is one factor taken into consideration in a weight-of-evidence approach and it supports the protein equivalence between the proteins expressed in plants compared to those expressed in bacteria, the latter being used for performing the allergenicity testing.

2mEPSPS

Potential N-glycosylation sites were determined using *in silico* search of the 2mEPSPS protein sequence for the presence of the consensus epitope Asn-Xaa-Ser/Thr (N-X~P-S/T), where Xaa = any amino acid except Pro (P), and Asn-Xaa-Cys (N-X-C) (Capt, 2008a; Appendix 14). Two potential N-glycosylation sites were identified on the amino acid sequence of the 2mEPSPS protein. However, the biological relevance of those potential N-glycosylations in eliciting allergenic response is not proven. The 2mEPSPS protein is not expected to be glycosylated, since chloroplastic proteins targeted directly to the chloroplast do not transit through the Endoplasmic Reticulum (ER) where glycosylation occurs in eukaryotes (Mousdale and Coggins, 1985, Pattison and Amtmann, 2009). In bacteria, protein glycosylation is rare (Sherlock *et al*., 2006). Furthermore, in the specific case of event FG72, it has been shown that the 2mEPSPS protein is not glycosylated. Therefore, potential allergenicity triggered by the presence of N-glycosylation sites is a remote possibility.

HPPD W336

Posttranslational modifications of bacterial proteins are generally similar to that exhibited by their eukaryotic counterparts. The possible posttranslational modifications of bacterial proteins are listed below in Table 24.

Table 24 Post-translational modifications of bacterial proteins

Posttranslational Modification	Comment
Glycosylation	N-glycosylation, O-glycosylation
Lipid modification	Lipoproteins, isoprenylation, acylation, PGI anchoring
Phosphorylation	Phosphoaspartate, phosphohistidine, phosphoserine, phosphothreonine, phosphotyrosine
Disulfide bonds	Cytosolic proteins
Proteolytic processing	Signal sequence cleavage, intein excision, amino-terminal and carboxyterminal maturation
Methylation	Methylarginine, methlyaspartic acid, methylcysteine, methylglutamic acid, methylglutamine, methylhistidine, methyllysine
Acetylation	
Amino acid modification	Hypusination, thiolation

Glycosylation has long been recognized as a fundamental eukaryotic strategy for influencing and modulating protein structure and function. It is becoming increasingly evident that protein glycosylation is also abundant in prokaryotes, with many glycoproteins identified. The glycans found on prokaryotic glycoproteins are far more diverse in terms of sugar composition and structure than those found in eukaryotic organisms. Due to their different cell structure, prokaryotes employ different mechanisms to glycosylate proteins. An example of glycosylation in *Pseudomonas* is the pilin O-glycan. This protein bears a shortchain oligosaccharide (Hitchen and Dell, 2006).

Potential N-glycosylation sites on the HPPD W336 protein were determined using an *in silico* search for the presence of the consensus epitope Asn-Xaa-Ser/Thr (N-X~P-S/T), where Xaa = any amino acid except Pro (P), and Asn-Xaa-Cys (N-X-C) (Capt, 2009b; Appendix 15). This analysis did not detect any potential N-glycosylation sites on the amino acid sequence of the HPPD W336 protein.

(e) Evidence of non-expression of a gene, in the case where a transferred gene is not expected to express any novel substances (e.g. because it has a 'silencing' roles or is in a non-functional form).

Not applicable. The two transferred genes, *2mepsps* and *hppdPf W336*, are functional as intended.

(f) Information about prior history of human consumption of the novel substances, if any, or their similarity to substances previously consumed in food.

Information pertaining to the prior history of human consumption of soybean, and the EPSPS and HPPD W336 proteins found in other sources, has been discussed previously in Sections 2.1 (d), 2.2 (a)(i)-(ii), 2.2(b)(ii)-(iv) and 2.4(a).

3.3 The Potential Toxicity of Novel Proteins or Other Novel Substances

(a) A bioinformatic comparison of the amino acid sequence of each of the novel proteins to known protein toxins and anti-nutrients (e.g. protease inhibitors, lectins).

HPPD W336 Protein
In an *in silico* approach, the potential amino acid sequence similarity of the HPPD W336 protein with known toxins within public protein databases was investigated. This overall amino acid sequence identity search was carried out by using the BLASTP algorithm, and all protein sequences present in the Uniprot_Swissprot, Uniprot_TrEMBL, PDB, DAD and GenPept reference databases. The scoring matrix, BLOSUM62, was used, with a conservative criterion of threshold E value of 0.1. Matched sequence proteins were further examined for potential toxicity records in literature in order to assess their biological relevance.

As expected, the HPPD W336 amino acid sequence showed homology with other HPPD proteins from various origins. These proteins included HPPD-like protein sequences annotated as toxins or putative toxins. However, a refined bioinformatics analysis demonstrated that these specific homologies were unlikely to be relevant. These matches are due to the presence, in the matching sequences, of a typical HPPD pattern that is extremely conserved amongst all known HPPDs from plants, animals, fungi and bacteria. Since direct hemolytic activity of these matching proteins was not demonstrated, it is unlikely that these homologies are biologically relevant (Capt, 2009c; Appendix 16).

2mEPSPS Protein
Using an *in silico* approach, the potential amino acid sequence similarity between 2mEPSPS protein and known toxins within public protein databases was investigated. The overall amino acid sequence identity search was conducted using the FASTA algorithm and the Uniprot_Swissprot, Uniprot_TrEMBL, PDB, DAD and GenPept reference databases. The criterion indicating potential toxicity was a 35% identity over at least 80 consecutive amino acids with a toxin.

As expected, similarities were found between the 2mEPSPS protein and other EPSPS proteins, and to other enzymes from various organisms that have good safety records. No identity was detected with known toxins, therefore it is unlikely that the 2mEPSPS protein could exhibit toxic properties (Capt, 2008d; Appendix 17).

Open Reading Frames
One potential indirect effect of insertion of trangenic sequences is the creation of new ORFs that can result in the expression of new proteins. As detailed in Section 2.3(d)(v) above, examination of the newly created junctions between the inserted sequences and the host genomic DNA in event FG72 revealed 46 putative ORFs between stop and stop codons (ORF-1 to ORF-46), and 8 putative ORFs between start and stop codons (ORF-66 to ORF-73).

Using an *in silico* approach, comparison was made between the putative translated amino acid sequences of ORF-66 to ORF-73, and sequences of known proteins (including toxins) that are contained within the Uniprot_Swissprot, Uniprot_TrEMBL, PDB, DAD, and GenPept reference databases, using the BLASTP algorithm. The matching criteria for a significant similarity to a toxin were 35% identity over the full-length query sequence and a low E-value (< 0.1) in an overall homology search. No significant similarities were identified in this analysis between the putative ORF-66 to ORF-73 amino acid sequences and sequences from the databases, including known toxins.

(b) Information on the stability to heat or processing and/or to degradation in appropriate gastric and intestinal model systems.

Heat Stability

The 2mEPSPS and HPPD W336 proteins were tested for stability on exposure to temperatures of 60, 75 or 90°C for periods of 10, 30 or 60 minutes. The 2mEPSPS and HPPD W336 proteins were examined by Coomassie blue stained-SDS-PAGE and by Western blot analysis (Rouquie, 2007; Appenidix 19 and Rascle, 2009a; Appendix 19).

For the 2mEPSPS protein, no significant changes were observed after a 10 minute treatment at 60°C. After 30 minutes at 60°C, some 2mEPSPS protein degradation fragments were evident. These fragments continued to degrade with increasing temperature and incubation periods. A marked, but incomplete, degradation of the 2mEPSPS protein was observed after incubation at 90°C for 60 minutes. It was concluded that the 2mEPSPS protein is partially heat-stable up to 90°C for 30 minutes and markedly degraded at 90°C after 60 minutes. Immunodetection was performed using a polyclonal antibody directed against 2mEPSPS protein. The western blot analysis showed an unchanged intensity of the intact 2mEPSPS band after incubation at 60°C or 75°C for 10 up to 60 minutes. At 90°C, the intensity of the intact 2mEPSPS band was unchanged after 10 and 30 minutes, but was decreased after 60 minutes, in accordance with the results obtained by SDS-PAGE analysis after Coomassie blue staining. In conclusion, the 2mEPSPS protein is partially heat-stable up to 90°C for 60 minutes.

Highly purified HPPD W336 protein produced in *E. coli* (96% purity) was also tested for structural stability at temperatures of 60, 75 or 90°C for periods of 10, 30 or 60 minutes. The protein was examined by SDS-PAGE followed by Coomassie blue staining and by western blot analysis. The immunodetection was performed using a polyclonal antibody directed against HPPD W336 protein. A 10% loading control was included in gels, to verify the sensitivity of the staining procedures. The Coomassie blue-stained SDS-PAGE showed no significant changes in the HPPD W336 protein after heat treatment at 60, 75 or 90°C from 10 to 60 minutes, with intensities similar to the unheated sample. Similar results were obtained with western blot analyses. These findings illustrate that while the HPPD W336 protein may retain its structural stability, its enzymatic activity decreases rapidly at 45^{o}C and is completely lost at higher temperatures (60^{o}C and above) after 2.5 minutes.

Digestibility tests

2mEPSPS Protein

The 2mEPSPS protein was tested for stability in human simulated gastric fluid (SGF) using a protocol that was developed based on the SGF test of the International Life Science Institute (ILSI) ring trial (Thomas *et al*., 2004) and the United States Pharmacopea (1990). The SGF was prepared as a pepsin (Sigma) solution at pH 1.2, and the incubation times tested ranged from 0.5 to 60 minutes. Control samples included the test protein without pepsin at pH 1.2 (0 and 60 minutes time-points), SGF alone (0 and 60 minutes time-points), and a 10% loading test protein control. Reference proteins horseradish peroxidase (HRP; Sigma) and ovalbumin (OVA; Sigma), known to be digested rapidly and slowly, respectively, were tested in parallel.

Test protein and reference protein solutions were incubated with human SGF at approximately 37°C, with samples taken for analysis at time-points of 0, 0.5, 2, 5, 10, 20, 30 and 60 minutes. The samples were analysed for presence of the test protein and potential stable protein fragments by Coomassie blue-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot analysis.

In the absence of pepsin, the 2mEPSPS protein remained stable throughout the sampling period. In the SGF, the intact 2mEPSPS protein was evident at time 0, but by the 30 second time point was completely digested. The pepsin remained unchanged throughout the 60 minutes of testing demonstrating its activity. The HRP and OVA exhibited the expected activities, with rapid and slow digestion, respectively. These results for the reference proteins are consistent with the results obtained in the ILSI ring trial (Thomas *et al*., 2004). Therefore, the 2mEPSPS protein was degraded very rapidly with no residual protein within 30 seconds of incubation with SGF, in presence of pepsin, at pH 1.2 (Rouquie, 2006a; Appendix 20).

The *2mEPSPS* protein was also tested for stability in human simulated intestinal fluid (SIF). This test was conducted using a protocol that was developed based on the SGF test of the ILSI ring trial (Thomas *et al*., 2004) and the United States Pharmacopea (1990). The SIF solution was prepared using porcine pancreatin at pH 7.5, and the incubation times tested ranged from 0.5 to 60 minutes. Controls included the pancreatin solution without the 2mEPSPS protein.

Samples were incubated at approximately 37°C and analysed at the time-points of 0, 0.5, 2, 5, 10, 20, 30 and 60 minutes using Coomassie blue-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot analysis.

At the sampling times of 0 and 60 minutes, the 2mEPSPS protein was not degraded in the absence of pancreatin. In the SIF, the 2mEPSPS protein already appeared degraded at the 0 time point, suggesting that digestion began within seconds of exposure to pancreatin. By the 30 second time point, complete digestion of the 2mEPSPS protein was confirmed. Therefore, the 2mEPSPS protein was completely degraded in SIF in less than 30 seconds in the presence of pancreatin at pH 7.5 (Rouquie, 2006b; Appendix 21)

HPPD W336 Protein
Using the same methodology described above for the 2mEPSPS protein, the digestibility of the HPPD W336 protein was also tested in SGF and SIF.

The SGF was prepared as a pepsin solution at pH 1.2, and the incubation times tested ranged from 0.5 to 60 minutes. Control samples included the test protein without pepsin at pH 1.2 (0 and 60 minutes time-points), SGF alone (0 and 60 minutes time-points), and a 10% loading test protein control. Reference proteins horseradish peroxidase (HRP) and ovalbumin (OVA), known to be digested rapidly and slowly, respectively, were tested in parallel.

At time zero of incubation with SGF, the HPPD W336 protein and the pepsin were clearly visible. At 30 seconds and all subsequent incubation times, the HPPD W336 protein was completely absent, while the pepsin remained unchanged throughout the 60 minutes of testing demonstrating its activity. The HRP and OVA exhibited the expected activities, with rapid and slow digestion, respectively. These results for the reference proteins are consistent with the results obtained in the ILSI ring trial (Thomas *et al*., 2004). Therefore, the HPPD W336 protein was degraded very rapidly with no residual protein after 30 seconds of incubation with SGF, in presence of pepsin, at pH 1.2. There was no significant digestion at pH 1.2 in the absence of pepsin, showing that digestion requires pepsin (Rascle, 2009b; Appendix 22).

For the SIF test, control samples included the HPPD W336 protein in buffer without pancreatin, the corresponding 10% loading condition (to verify the sensitivity of the detection procedure) and SIF without HPPD W336 protein. At the sampling times of 0 and 60 minutes, the HPPD W336 protein was not degraded in the absence of pancreatin. In the SIF, the HPPD W336 already appeared degraded at the 0 time point, suggesting that digestion began within seconds of exposure to pancreatin. By the 30 second time point, complete digestion of the HPPD W336 protein was confirmed (Rascle, 2009c; Appendix 23).

(c) Detailed reports of all available acute or short term oral toxicity studies in animals on the novel proteins or other novel substances.

2mEPSPS

The acute toxicity of the 2mEPSPS protein was determined in two tests; oral gavage in mice, and intravenous injection in mice.

Oral gavage

A 2mEPSPS protein extract of 99.52% purity isolated from *E. coli* was used in the oral gavage test. Groups of five female OF1 mice were administered either the 2mEPSPS protein or bovine serum albumin (negative control; >97% purity, Sigma-Aldrich) at dose levels of 2000 mg/kg body weight. Animals were observed for the nature, severity, reversibility and duration of all clinical signs daily for fifteen days, and their body weights were measured at days -6, -2, 1, 8 and 15. At termination of the study period, all animals were subjected to a necropsy including macroscopic examination of abdominal and thoracic cavities, major organs and tissues. There were no mortalities, clinical signs, treatment-related macroscopic abnormalities or treatment-related effects on body weight in female OF1 mice after an acute oral administration of 2mEPSPS protein at 2000 mg/kg body weight. Therefore, the treatments did not cause systemic toxicity in the OF1 female mouse (Rouquie, 2006c; Appendix 24).

Intravenous injection

The intravenous injection test was based on the US EPA Health Effects Test Guidelines OPPTS 870.1100 adopted in 2002 (US EPA, 2002), and on the OECD Test Guideline 425 (OECD, 2001a). Groups of five female OF1 mice were administered either with the 2mEPSPS protein, the aprotinin (negative control; Sigma-Aldrich) or the melittin (positive control; Sigma-Aldrich) proteins in PBS buffer by intravenous injection at dose levels of 1 and 10 mg/kg body weight. All animals were observed for clinical signs daily for fifteen days whilst their body weights were measured weekly. At termination of the study period, animals were subjected to a necropsy including macroscopic examination of abdominal and thoracic cavities, major organs and tissues.

As expected, all the animals of the positive control group (melittin at 10 mg/kg body weight) died within 4 hours of treatment. This is consistent with the LD50 for melittin which is approximately 3 mg/kg. Animals treated at 1 mg/kg of melittin and the negative control animals treated with aprotinin at 1 or 10 mg/kg body weight or with the vehicle only, showed no visible signs of systemic toxicity. In the animals administered the the 2mEPSPS protein at 1 or 10 mg/kg body weight, there were no mortalities, clinical signs, macroscopic abnormalities or treatment-related effects on body weight in female OF1 mice after intravenous injection. Therefore, under the conditions of this study, the 2mEPSPS protein was not acutely toxic by the intravenous route (Rouquie, 2008; Appendix 25).

HPPD W336

The acute toxicity of the HPPD W336 protein was determined in two tests; oral gavage in mice, and intravenous injection in mice.

Oral gavage

A group of five female OF1 mice were administered a single dose of the HPPD W336 protein by oral gavage at the dose level of 2000 mg/kg body weight. A similarly constituted group of five female mice received bovine serum albumin (BSA) at the same dose level and acted as a control. The test or reference proteins were administered in two doses of 1000 mg/kg body weight administered within a 4 hour period on the day of treatment. All animals were observed for clinical signs daily for fifteen days whilst their body weights were measured weekly. At termination of the study period, all animals were subjected to a necropsy including macroscopic examination, and the spleen, liver, kidney and brain were weighted. Microscopic examination of the spleen revealed there were no mortalities, no clinical signs or treatment-related effects on body weight, body weight gain, organ weights, gross and microscopic examinations. In conclusion, under the conditions of this study, the HPPD W336 protein did not induce any evidence of systemic toxicity in the OF1 female mouse (Rascle, 2009d; Appendix 26).

Intravenous injection
Groups of five female OF1 mice were administered either with the HPPD W336 or the aprotinin protein in PBS buffer supplemented with 1 µM FeCl3 by intravenous injection at dose levels of 1 and 10 mg/kg body weight. All animals were observed for clinical signs daily for fifteen days whilst their body weights were measured weekly. At termination of the observation period, animals were subjected to a necropsy including macroscopic examination. There were no mortalities and no treatment-related effects in female OF1 mice after an acute intravenous administration of the HPPD W336 at 1 or 10 mg/kg body weight. Therefore under the conditions of the study, the HPPD W336 was devoid of acute toxic potential up to 10 mg/kg (intravenous route) (Rascle, 2009e; Appendix 27).

3.4 The Potential Allergenicity of Novel Proteins

(a) The source of the introduced protein.

Soybean event FG72 encodes for two novel proteins:
(i) The 2mEPSPS protein. The coding sequence for the 2mEPSPS protein was developed by introducing two point mutations to the wild-type *epsps* gene cloned from maize (*Zea mays*). Expression of the 2mEPSPS protein confers tolerance to glyphosate herbicides.
(ii) The HPPD W336 protein. The coding sequence for the HPPD protein is derived from *Pseudomonas fluorescens* strain A32 (Genebank A69533; McKellar, 1982). A single amino acid change was introduced to the wild type HPPD protein at position 336 (glycine changed to tryptophane) to encode for the HPPD W336 protein. Expression of the HPPD W336 protein confers tolerance to isoxazole herbicides.

(b) Any significant similarity between the amino acid sequence of the protein and that of known allergens.

Amino Acid Sequence Homology Searches

HPPD W336 Protein
The potential amino acid sequence similarity of the HPPD W336 protein with known allergens was investigated using several *in silico* approaches:
(i) An epitope search was conducted to identify any short amino acid sequences that might represent an isolated shared allergenic epitope. In this search, the amino acid sequence of the HPPD W336 protein was subdivided into 8 amino acid blocks and compared with all known allergens within the public allergen database AllergenOnline (www.allergenonline.com; release 9.2, 1386 sequences). The algorithm used was FindPatterns (GCG package) and the criterion indicating potential allergenicity was a 100% identity on a window of 8 amino acids with an allergenic protein.
(ii) An overall amino acid identity search was conducted using the FASTA algorithm to compare the complete amino acid sequence of the HPPD W336 protein with all protein sequences present in the AllergenOnline database. The criterion indicating potential allergenicity was a 35% identity over at least 80 consecutive amino acids with an allergenic protein.
(iii) An 80-mer allergenic identity search was performed to compare the query sequence subdivided into 80 amino acid blocks with all known allergens present in the AllergenOnline database. The criterion indicating potential allergenicity was a 35 % identity with an allergenic protein.

Potential N-glycosylation sites were also investigated by searching known consensus sequences of allergenic proteins.

The epitope homology search revealed no identity between the HPPD protein and epitopes from known allergenic proteins. The overall and 80-mer identity searches also revealed no relevant sequence similarity between the HPPD W336 protein and any known allergens from the AllergenOnline database. Furthermore, no potential N-glycosylation sites were identified in the HPPD

W336 amino acid sequence. In conclusion, it is unlikely that the HPPD W336 protein possesses allergenic properties (Capt, 2009b; Appendix 15).

2mEPSPS Protein

The potential amino acid sequence similarity of the 2mEPSPS protein with known allergens was investigated using two *in silico* approaches:

(i) An overall amino acid identity search was conducted using the FASTA algorithm to compare the complete amino acid sequence of the 2mEPSPS protein with all protein sequences present within the Uniprot_Swissprot, Uniprot_TrEMBL, PDB, DAD and GenPept reference databases and AllergenOnline database. The criterion indicating potential allergenicity was a 35% identity over at least 80 consecutive amino acids with an allergenic protein.

(ii) An 80-mer allergenic identity search was performed to compare the query sequence subdivided into 80 amino acid blocks with all known allergens present in the AllergenOnline database. The criterion indicating potential allergenicity was a 35 % identity with an allergenic protein

No identity was detected between the 2mEPSPS protein and known allergens. As expected, only high sequence similarities were detected with other EPSPS proteins and other enzymes from various organisms, which have good safety records. Therefore, it is unlikely that the 2mEPSPS protein could exhibit allergenic properties (Capt, 2008d; Appendix 17).

Open Reading Frames

One potential indirect effect of insertion of transgenic sequences is the creation of new ORFs that can result in the expression of new proteins. As detailed in Section 2.3(d)(v) above, examination of the newly created junctions between the inserted sequences and the host genomic DNA in event FG72 revealed 46 putative ORFs between stop and stop codons (ORF-1 to ORF-46), and 8 putative ORFs between start and stop codons (ORF-66 to ORF-73).

Using an *in silico* approach, comparison was made between putative translated amino acid sequences of ORF-66 to ORF-73, and sequences of known allergens contained within the Uniprot_Swissprot, Uniprot_TrEMBL, PDB, DAD, and GenPept reference databases and an allergen database (AllergenOnline) using BLASTP or FindPatterns algorithms. The matching criteria for significant similarity to an allergen were 100% identity over a linear contiguous 8 amino acid segment (epitope homology search), or 35% identity over the full-length query sequence with a known allergen and a low E-value (<0.1) (overall homology search). No similarities were found between the putative ORF-66 to ORF-73 amino acid sequence and sequences from the databases, including known allergens (Verhaeghe, 2009d; Appendix 5).

(c) Its structural properties, including but not limited to, its susceptibility to enzymatic degradation (e.g. digestion by pepsin), heat stability and/or acid and enzymatic treatment.

The structural stability of the 2mEPSPS and HPPD W336 proteins on exposure to heat, and degradation in simulated human gastric and intestinal fluids are detailed above in Section 3.3(b).

(d) Specific serum screening where a newly expressed protein is derived from a source known to be allergenic or has sequence homology with a known allergen.

Not applicable. The newly expressed proteins encoded by the FG72 event, 2mEPSPS and HPPD W336, are not from sources known to be allergenic, nor do they show any homology to known allergens.

3.5 Compositional Analyses of the GM Food

(a) The levels of key nutrients, toxicants and anti-nutrients in the GM food compared with the levels in appropriate comparator (usually the non-GM counterpart). The statistical significance of any observed differences must be assessed in the context of the range of natural variations for that parameter to determine its biological significance

A compositional analysis was performed to compare the nutritional properties of components derived from the FG72 event with non-transgenic soybean. Composition data were obtained for 120 samples of soybean seed that were generated from 10 field trials in the states of Iowa, Illinois, Indiana and Missouri (Kowite, 2009b; Appendix 13). There were 12 plots from four groups in each field trial with a single sample being harvested from each plot:

1. Non-transgenic counterpart, non-tolerant Jack soybean (3 plots per trial).
2. FG72 transgenic event soybean that was not sprayed with herbicides (3 plots per trial).
3. FG72 transgenic event soybean that received a single foliar application of a tank mix Balance® Pro at 0.062 lb ai/A, Roundup Original Max® at 0.95 lb ai/A, and ammonium sulfate at 8 lb /100 gal at post V4-V5 growth stage (3 plots per trial).
4. Commercial non-transgenic Stine® soybean lines 2686-6, 2788, and 3000-0 (1 plot of each variety per trial).

Samples were analyzed for proximates, minerals, anti-nutrients, total amino acids, total fatty acids, vitamins and isoflavones. The results of these analyses, with reference ranges calculated from three commercial soybean varieties, are presented below in Tables 25 through Table 36 (Oberdörfer, 2009; Appendix 8), with the statistical analysis detailed in Rattemeyer (2009; Appendix 28). Raw data is presented in Appendix 29 (Mackie, 2009).

Proximate and Fibre

The proximates (moisture, crude protein, total fat, ash, acid detergent fiber, and neutral detergent fiber) were measured in soybean. Total carbohydrates were calculated as 100% minus the protein %, fat %, and ash %. The mean ± the standard deviation of the proximates (expressed on a dry matter basis except for % moisture) for the various soybean regimens across all trials are presented in Table 26. The means of the proximates for the transgenic (treated and not-treated) and the non-transgenic control are inside the reference range for the commercial products and ranges in the published literatures (Table 25).

In the results of statistical evaluation, no significant differences between regimens mean values over all sites (p-value ANOVA for main effect treatment ≥ 0.05 in Table 25) were found for moisture, fat, protein, ADF and NDF indicating substantial equivalence between data sets of the non-transgenic and transgenic groups. Due to significant regimen*site interactions an overall analysis is not valid for carbohydrates. Therefore, the by site analyses for carbohydrates were performed and a summary of results is given in Table 26. It can be seen that for the carbohydrates the majority of the by site analyses did not show significant differences between the regimen, which indicates the equivalence of the data sets cannot be excluded.

Statistically significant differences were found for ash in the over-all sites analyses, so the individual treatment comparisons were evaluated. Although the outcome of the t-tests showed significant difference for ash (Table 25), the biological and nutritional relevance of the statistical findings is negligible because of the following reasons:

- The mean values of ash are inside the reference range for the commercial products and ranges in the published literatures.
- The estimated differences between regimens are very small. They are lower than the variation (SD) inside the non-transgenic control group.

Minerals

The results of the mineral analysis included calcium, phosphorus, potassium, magnesium, sodium and iron. The mean ± the standard deviation of the individual minerals (expressed on a dry matter basis) of the various soybean regimens across all trials are presented in Table 27. The mean values of all of these mineral components are within the reference range for the commercial products and

ranges in the published literatures.

In the results of statistical evaluation, no statistically significant differences were identified (p-value > 0.05) for phosphorus and iron by the ANOVA comparing the three entries, indicating substantial equivalence between data sets of the non-transgenic and transgenic groups (Table 27). Significant regimen effects were seen for calcium, magnesium and sodium (A vs B and A vs C), however the similarity in numerical differences for the mean values and ranges, along with the standard deviations demonstrates that there was very little variation in this analysis between entries. All mean values for these minerals fell within the calculated commercial variety and literature reference ranges indicating that the statistical differences observed are not considered to be of biological importance nor due to the modification of the FG72 event.

Due to significant regimen*site interactions an overall analysis is not valid for potassium. A summary of the by site analyses for potassium is given in the Table 28. It can be seen that the majority of the by site analyses did not show significant differences between regimen A and B; the results for the comparison of A and C were ambiguous.

Antinutrients

The anti-nutrients phytic acid, raffinose, stachyose, trypsin inhibitors, and lectin were measured in soybean. The mean ± the standard deviation of antinutrients expressed on a dry matter basis) of the various soybean regimens across all trials are presented in Table 29. The mean values of all of these anti-nutrients are within the reference range for the commercial products and ranges in the published literatures.

In the results of statistical evaluation, no statistically significant differences were identified (p-value > 0.05) for stachyose, phytic acid, lectin by the ANOVA comparing the three entries, indicating substantial equivalence between data sets of the non-transgenic and transgenic groups (Table 29). The level of trypsin inhibitor measured for the test and control entries was shown to be statistically different in the across location analysis. Although the ANOVA showed a significant effect, no significant differences were detected in the individual t-tests. A significant regimen difference was seen for raffinose (A vs B and A vs C). However, the levels of raffinose in the two test and control entries were within the range of the commercial lines tested, as well as the literature reference range, indicating that the observed statistical difference was not of biological relevance, nor due to the modification of the FG72 event.

Amino Acids

Eighteen amino acids (aspartic acid, threonine, serine, glutamic acid, proline, glycine, alanine, cystine, valine, methionine, isoleucine, leucine, tyrosine, phenylalanine, lysine, histidine, arginine, and tryptophan) were measured in soybean. The mean ± the standard deviation of the amino acids (expressed on a dry matter basis) of the various soybean regimens across all trials are presented in Table 30. The means of amino acids for the two transgenic (treated and not-treated) and the non-transgenic control are inside the reference range for the commercial products and ranges in the published literatures (Table 30).

Analysis of the total amino acid profile for all 18 amino acids, between the two transgenic (treated and not-treated) and the non-transgenic control entries, were found to be similar, and no statistically significant differences were identified (p-value > 0.05) by the ANOVA comparing the three entries, indicating substantial equivalence between data sets of the non-transgenic and transgenic groups (Table 30).

Due to significant regimen*site interactions an overall analysis is not valid for serine. Therefore, the by site analyses for serine were performed and a summary of results is given in Table 31. It can be seen that the majority of the by site analyses did not show significant differences between regimen A and B and between A and C, which indicates the equivalence of the data sets cannot be excluded.

Fatty Acids

Twenty eight fatty acids (caprylic, capric, lauric, myristic, myristoleic, pentadecanoic, pentadecenoic, palmitic, palmitoleic, heptadecanoic, heptadecenoic, stearic, oleic, linoleic, gamma linolenic, linolenic, octadecatetraenoic, arachidic, eicosenoic, eicosadienoic, arachidonic, eicosatrienoic, eicosapentaenoic, behenic, erucic, docosapentaenoic, lignoceric, and docosahexaenoic) were measured in soybean. The mean ± the standard deviation of the fatty acids (expressed as percent relative to the total) of the various soybean regimens across all trials are presented in Table 32. Of the 28 fatty acids assayed, only those fatty acids with concentration values above the LOQ are presented in Table 32. The sum of all detected fatty acids was 99.9% in the analysis conducted, accounting for nearly all fatty acids present in the oil.

The mean values of all of these components, except palmitic acid and oleic acid, are within the reference range for the commercial products and ranges in the published literatures. The mean values of palmitic acid and oleic acid were slightly different to the range calculated from the results of the three tested commercial soybean lines. However they are still within the reference ranges from literature.

All other fatty acids are shown to be statistically significant in the across location analysis, however the similarity in numerical differences for the mean values and ranges, along with the standard deviations demonstrates that there was very little variation in this analysis between entries. All mean values for the fatty acids fell within both the calculated commercial variety and literature reference ranges, indicating that the statistical differences observed are not considered to be of biological importance, nor due to the modification of the FG72 event.

For C24:0 an overall analysis is not valid due to significant regimen*site interactions. Therefore, the by site analyses for C24:0 Lignoceric were performed and a summary of results is given in Table 33. It can be seen that the majority of the by site analyses did not show significant differences between the regimen.

Vitamins

Results of the vitamin analyses are shown in Table 34. The analysis included alpha, beta, gamma, and delta tocopherol, as well as total tocopherols. Vitamins A, B1, B2, K, and folic acid were also measured in soybean. The mean ± the standard deviation of the vitamins and tocopherols (expressed on a dry matter basis) of the various soybean regimens across all trials are presented in Table 34.

The mean values of all of these components, except Vitamins A and K, are within the reference range for the commercial products and ranges published in literature. Since some samples had no detectable vitamin A and K levels, a range including the minimum and maximum result was built for these two nutrients. Both ranges fell slightly short of the literature ranges. However, it should be taken into account that only a few reference values were found for vitamin A and K contents in soybean seeds.

In the results of statistical evaluation, no statistically significant differences were identified (p-value > 0.05) for vitamin B2 and folic acid by the ANOVA comparing the three entries, indicating substantial equivalence between data sets of the non-transgenic and transgenic groups. A significant regimen effect was seen for vitamin B1 (A vs C), gamma tocopherol (A vs B) and total tocopherol (A vs B and A vs C). Due to significant regimen*site interactions an overall analysis is not valid for vitamin A, vitamin K, alpha tocopherol and delta tocopherol. Therefore, the by site analyses for these components was performed and a summary of the by site analyses is given in the Table 35. It can be seen that the majority of the by site analyses did not show significant differences between regimen A and B and between A and C.

Isolflavones

Soybeans contain isoflavones which are glucosides and esters of three aglycones (daidzein, genistein and glycitein). Total isoflavones were calculated by converting the isoflavone glucosides genistin, glycitin, and daidzin to aglycon equivalents and summing as follows:

Total isoflavones (aglycon equivalents) = ([genistin] x 270/432) + ([glycitin] x 284/446) + ([daidzin] x 254/416) + [genistein] + [glycitein] + [daidzein]

If the value for an individual isoflavone was <LOQ, the value of 0 was used for that analyte to determine the total isoflavones. The mean ± the standard deviation of the isoflavones (expressed on a dry matter basis) of the various soybean regimens across all trials are presented in Table 36. The mean values of all of these components except glycitin are within the reference range for the commercial products and ranges in the published literatures. The mean values of glycin were slightly different to the range calculated from the results of the three tested commercial soybean lines. However they are still within the reference ranges from literature.

In the results of statistical evaluation, no statistically significant differences were identified (p-value > 0.05) for daidzein and daidzin by the ANOVA comparing the three entries, indicating substantial equivalence between data sets of the non-transgenic and transgenic groups (Table 36). Significant regimen differences were seen for glycitin, genistin (A vs B and A vs C) and total isoflavones (A vs C). However the similarity in numerical differences for the mean values and ranges, along with the standard deviations, demonstrates that there was very little variation in this analysis between entries. All mean values for these components are within the reference ranges indicating that the statistical differences observed are not considered to be of biological importance nor due to the modification of the FG72 event.

Due to significant regimen*site interactions an overall analysis is not valid for genistein. A summary of the by site analyses for genistein and total isoflavones is given in the Table 37. It can be seen that the majority of the by site analyses did not show significant differences between regimen A and B and between A and C for these two parameters.

Table 25 Proximate and fiber compounds in soybean seed of event FG72 and the non-transgenic counterpart Jack compared to commerical soybean line (reference range) and results from ANOVA and T-test (over all sites)

Parameter	Non-Transgenic Not-treated	Transgenic Not-treated	Transgenic Treated [a]	Range Commercial Soybean Lines [b]	Reference Range [c]	ANOVA (p-values)		T-test (p-values)*	
	Mean ± SD	Mean ± SD	Mean ± SD			regimen effect	regimen*site interaction	A vs.B	A vs. C
Moisture % fw	9.51 ± 0.82	9.65 ± 0.84	9.45 ± 0.83	8.00 – 10.60	5.6-12	0.499	0.234		
Fat % dm	19.3 ± 0.9	18.9 ± 1.2	19.2 ± 1.1	15.1 – 21.4	8.1 – 24.7	0.064	0.146		
Protein % dm	38.2 ± 1.1	38.2 ± 0.8	38.1 ± 0.9	35.8 – 40.1	32 – 45.5	0.799	0.277		
Ash % dm	5.24 ± 0.31	5.07 ± 0.30	5.06 ± 0.28	4.89 – 5.73	3.9 – 7.0	<.001	0.568	0.001	<.001
Total carb. % dm [d]	37.3 ± 1.2	37.9 ± 1.0	37.6 ± 1.2	34.8 – 41.6	29.6- 50.2	NA	0.012		
ADF[e] % dm	17.8 ± 1.9	18.1 ± 2.0	17.9 ± 1.8	13.6 – 23.5	7.8 – 18.6	0.832	0.342		
NDF[f] % dm	19.8 ± 2.0	20.3 ± 2.1	20.0 ± 1.5	16.1 – 24.8	5.0 – 21.3	0.500	0.637		

% fw – Percentage Fresh Weight
% dm – Percentage of Dry Matter
[a]Treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY])
[b]Reference ranges of the 3 analysed commercial soybean lines taken from Table 1a of Appendix A to the statistical report (Rattemeyer, 2009; Appendix 28)
[c] Reference ranges from Table 2 of Appendix B in the reference (Oberdorder, 2009; Appendix 8)
[d]Total carbohydrates calculated as 100% - (protein %dm + fat %dm + ash %dm)
[e]ADF – Acid Detergent Fibre
[f] NDF – Neutral Detergent Fibre

* only in cases of no interactions (p>0.05) and significant regimen effects (p<0.05)
A = non-transgenic plants (Jack), not treated,
B = transgenic plants (FG72), not treated
C = transgenic plants (FG72), treated
NA: not applicable due to regimen*site interaction

Table 26 Results from T-tests - by-site analysis for proximate and fiber compounds

Summary t-test procedures *)	A vs B		A vs C	
	significant	not significant	significant	not significant
Carbohydrates	2	8	1	9

*) N of sites with significant ($p < 0.05$) and not significant ($p \geq 0.05$) treatment differences
A = non-transgenic seed from the control Jack
B = transgenic seed from non glyphosate treated FG72 entry
C = transgenic seed from glyphosate treated FG72 entry

Table 27 Minerals in soybean seed of event FG72 and the non-transgenic counterpart Jack compared to commercial soybean Lines (reference ranges)and results from ANOVA and T-tests (over all sites)

Parameter	Non-Transgenic			Transgenic Non-treated			Transgenic Treated [a]			Range Commercial Soybean Lines [b]	Reference Range [c]	ANOVA (p-values)		T-test (p-values)*	
	Mean	±	SD	Mean	±	SD	Mean	±	SD			regimen effect	regimen*site interaction	A vs.B	A vs. C
Calcium (% dm)	0.282	±	0.023	0.258	±	0.024	0.259	±	0.026	0.212 – 0.347	0.12 – 0.34	<.001	0.058	<.001	<.001
Phosphorus (% dm)	0.626	±	0.053	0.618	±	0.062	0.620	±	0.065	0.499 – 0.651	0.49 – 0.94	0.490	0.151		
Potassium (% dm)	1.93	±	0.08	1.85	±	0.08	1.85	±	0.09	1.84 – 2.11	1.4 – 2.3	NA	0.006		
Magnesium (% dm)	0.241	±	0.010	0.226	±	0.012	0.226	±	0.010	0.197 – 0.263	0.21 – 0.32	<.001	0.065	<.001	<.001
Sodium (%)	0.012	±	0.003	0.015	±	0.007	0.016	±	0.008	<LOQ[d] – 0.026	0.002 – 0.02	0.010	0.279	0.019	0.004
Iron (mg/kg)	93.3	±	41.8	82.6	±	13.3	84.1	±	18.9	58.8 – 175.0	55.4 - 172	0.127	0.303		

[a] Treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY])
[b] Reference ranges of the 3 analysed commercial soybean lines taken from Table 1b of Appendix A to the statistical report (Rattemeyer, 2009; Appendix 28)
[c] Reference ranges from Table 3a of Appendix B in the reference (Oberdorder, 2009; Appendix 8)
[d] some samples had no detectable vitamin K, LOQ for Vitamin K was 0.10 mg/kg dm

* only in cases of no interactions ($p>0.05$) and significant regimen effects ($p<0.05$)
A = non-transgenic plants (Jack), not treated,
B = transgenic plants (FG72), not treated
C = transgenic plants (FG72), treated
NA – not applicable due to regimen*site interaction

Table 28 Results from T-tests - by-site analysis for minerals

Summary t-test procedures *)	A vs B		A vs C	
	significant	not significant	significant	not significant
Potassium	4	6	5	5

*) N of sites with significant ($p < 0.05$) and not significant ($p \geq 0.05$) treatment differences
A = non-transgenic seed from the control Jack
B = transgenic seed from non glyphosate treated FG72 entry
C = transgenic seed from glyphosate treated FG72 entry
#) 'not significant' was also assumed if all samples of a site were equal or below the limit of quantification for the two respective treatments (N of sites in brackets)

Table 29 Anti-nutrients in soybean seed of event FG72 and the non-transgenic counterpart Jack compared to commercial soybean lines (reference ranges) and results from ANOVA and T-tests (over all sites)

Parameter	Non-Transgenic	*Transgenic Non-treated*	*Transgenic Treated* [a]	*Commercial Soybean Lines* [b]	*Literature Reference Range* [c]	ANOVA (p-values)		T-test (p-values)*	
						regimen effect	regimen*site interaction	A vs.B	A vs. C
	Mean ± SD	Mean ± SD	Mean ± SD						
Phytic Acid (%dm)	1.40 ± 0.16	1.37 ± 0.23	1.35 ± 0.23	0.96 – 1.50	0.63 – 2.74	0.140	0.122		
Raffinose (% dm)	0.36 ± 0.04	0.38 ± 0.05	0.38 ± 0.06	0.29 – 0.50	0.11 – 1.28	0.035	0.106	0.027	0.022
Stachyose (% dm)	2.49 ± 0.24	2.42 ± 0.18	2.50 ± 0.19	2.23 – 2.96	1.21 – 6.30	0.272	0.915		
Lectin (HU/mg)	1.74 ± 0.60	1.40 ± 0.50	1.54 ± 0.42	0.46 – 8.63	0.11 - 129	0.054	0.836		
Trypsin Inhibitor (TIU/mg)	33.0 ± 6.6	30.1 ± 6.1	33.9 ± 5.7	23.5 – 60.1	19.59 - 118	0.041	0.879	0.061	0.564

[a] Treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY])
[b] Reference ranges of the 3 analysed commercial soybean lines taken from Table 1d of Appendix A to the statistical report (Rattemeyer, 2009; Appendix 28)
[c] Reference ranges from Table 5 of Appendix B in the reference (Oberdorder, 2009; Appendix 8)

* only in cases of no interactions ($p>0.05$) and significant regimen effects ($p<0.05$)
A = non-transgenic plants (Jack), not treated,
B = transgenic plants (FG72), not treated
C = transgenic plants (FG72), treated
NA: not applicable due to regimen*site interaction

Table 30 Total amino acids in soybean seed of event FG72 and the non-transgenic counterpart Jack compared to commercial soybean line (reference range) and results from ANOVA and T-test (over all sites)

Parameter	% Dry matter					ANOVA (p-values)	
	Non-Transgenic	Transgenic Non-treated	Transgenic Treated [a]	Range Commercial Soybean Lines [b]	Reference Range [c]	regimen effect	regimen*site interaction
	Mean ± SD	Mean ± SD	Mean ± SD				
Alanine	1.68 ± 0.04	1.68 ± 0.04	1.68 ± 0.04	1.55 – 1.78	1.51 – 2.10	0.901	0.644
Arginine	2.94 ± 0.10	2.97 ± 0.10	2.95 ± 0.10	2.69 – 3.13	2.17 – 3.40	0.344	0.487
Aspartic acid	4.40 ± 0.12	4.38 ± 0.12	4.37 ± 0.13	4.06 – 4.67	3.81 – 5.12	0.555	0.450
Cysteine	0.58 ± 0.03	0.58 ± 0.02	0.59 ± 0.03	0.50 – 0.63	0.37 – 0.81	0.476	0.245
Glutamic acid	6.75 ± 0.21	6.77 ± 0.23	6.74 ± 0.22	6.32 – 7.23	5.84 – 8.20	0.812	0.409
Glycine	1.68 ± 0.04	1.68 ± 0.04	1.68 ± 0.04	1.53 – 1.76	1.46 – 2.27	0.960	0.575
Histidine	1.05 ± 0.03	1.05 ± 0.03	1.05 ± 0.03	0.93 – 1.07	0.84 – 1.22	0.963	0.720
Isoleucine	1.81 ± 0.05	1.80 ± 0.05	1.79 ± 0.05	1.62 – 1.96	1.54 – 2.32	0.379	0.977
Leucine	2.99 ± 0.08	2.99 ± 0.08	2.98 ± 0.08	2.71 – 3.13	2.2 – 4.0	0.671	0.575
Lysine	2.48 ± 0.05	2.48 ± 0.06	2.47 ± 0.06	2.34 – 2.64	1.55 – 2.84	0.943	0.731
Methionine	0.54 ± 0.02	0.54 ± 0.02	0.54 ± 0.02	0.50 – 0.58	0.43 – 0.76	0.916	0.461
Phenylalanine	1.97 ± 0.05	1.98 ± 0.06	1.96 ± 0.06	1.83 – 2.08	1.60 – 2.39	0.264	0.603
Proline	1.82 ± 0.07	1.83 ± 0.07	1.82 ± 0.07	1.71 – 1.94	1.69 – 2.33	0.753	0.291
Serine	1.97 ± 0.07	1.98 ± 0.08	1.99 ± 0.06	1.77 – 2.13	1.11 – 2.48	NA	0.047
Threonine	1.55 ± 0.04	1.54 ± 0.04	1.53 ± 0.04	1.44 – 1.62	1.14 – 1.89	0.254	0.156
Tryptophan	0.45 ± 0.03	0.44 ± 0.03	0.44 ± 0.03	0.39 – 0.54	0.36 – 0.67	0.119	0.445
Tyrosine	1.40 ± 0.04	1.40 ± 0.04	1.40 ± 0.04	1.32 – 1.48	0.10 – 1.61	0.582	0.225
Valine	1.89 ± 0.06	1.88 ± 0.05	1.87 ± 0.06	1.66 – 2.03	1.50 – 2.44	0.609	0.861

[a] Treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY])
[b] Reference ranges of the 3 analysed commercial soybean lines taken from Table 1f of Appendix 28 to the statistical report (Rattemeyer, 2009)
[c] Reference ranges from Table 6 of Appendix 9 in the reference (Oberdorder, 2009; Appendix 8)
NA: not applicable due to regimen*site interaction

Table 31 Results from T-tests - by-site analysis for amino acids

Summary t-test procedures *)	A vs B		A vs C	
	significant	not significant	significant	not significant
Serine	3	7	-	10

*) N of sites with significant ($p < 0.05$) and not significant ($p \geq 0.05$) treatment differences
A = non-transgenic seed from the control Jack
B = transgenic seed from non glyphosate treated FG72 entry
C = transgenic seed from glyphosate treated FG72 entry
#) 'not significant' was also assumed if all samples of a site were equal or below the limit of quantification for the two respective treatments (N of sites in brackets)

Table 32 Total fatty acids in soybean seed of event FG72 and the non-transgenic counterpart Jack compared to commercial soybean line (reference range) and results from ANOVA and T-test (over all sites)

Parameter	% relative					ANOVA (p-values)		T-test (p-values) *	
	Non-Transgenic	*Transgenic Non-treated*	*Transgenic Treated* [a]	*Range Commercial Soybean Lines* [b]	*Reference Range* [c]	regimen effect	regimen*site interaction	A vs.B	A vs. C
	Mean ± SD	Mean ± SD	Mean ± SD						
Saturated									
C16:0 (palmitic)	10.06 ± 0.22	9.34 ± 0.17	9.38 ± 0.23	9.78 – 11.40	7 – 16	< .001	0.376	< .001	< .001
C18:0 (stearic)	4.28 ± 0.16	4.52 ± 0.19	4.51 ± 0.23	3.49 – 4.81	2 – 5.9	< .001	0.358	< .001	< .001
C20:0 (arachidic)	0.312 ± 0.015	0.324 ± 0.017	0.324 ± 0.019	0.25 – 0.35	< 0.10 - 0.48	< .001	0.067	< .001	< .001
C22:0 (behenic)	0.319 ± 0.009	0.339 ± 0.012	0.327 ± 0.017	0.25 – 0.35	0.28 – 0.60	0.001	0.462	< .001	< .001
C24:0 (lignoceric)	0.113 ± 0.020	0.119 ± 0.022	0.122 ± 0.025	< 0.10 – 0.15	0.15	NA	0.033		
Sum Saturated	14.9	14.5	14.5	13.8 – 17.2	9.43 – 23.55				
Mono-unsaturated									
C18:1 (oleic)	21.97 ± 1.05	24.65 ± 0.99	24.12 ± 0.90	21.10 – 24.10	14 – 34	< .001	0.153	< .001	< .001
C20:1 (Gadoleic)	0.161 ± 0.011	0.165 ± 0.019	0.166 ± 0.012	< 0.10 – 0.18	0.14 – 0.35	0.003	0.454	0.017	< .001
Sum Mono-saturated	22.13	24.81	24.29	21.10 – 24.28	14.14 – 34.83				
Poly-unsaturated									
C18:2 (linoleic)	54.56 ± 0.90	52.65 ± 0.95	53.08 ± 0.82	51.50 – 55.40	48 – 60	< .001	0.230	< .001	< .001
C18:3 (alpha linolenic)	8.27 ± 0.50	7.94 ± 0.45	8.01 ± 0.48	7.59– 10.30	2 – 10	< .001	0.608	< .001	< .001
Sum Poly-saturated	62.83	60.59	61.09	59.09 – 65.70	50 - 70				
Sum of total fatty acids	99.93	99.91	99.92	-					

[a] Treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY])
[b] Reference ranges of the 3 analysed commercial soybean lines taken from Table 1g of Appendix A to the statistical report (Rattemeyer, 2009; Appendix 28)
[c] Reference ranges from Table 7 of Appendix B in the reference (Oberdorder, 2009; Appendix 8)

* only in cases of no interactions (p>0.05) and significant regimen effects (p<0.05)
A = non-transgenic plants (Jack), not treated,
B = transgenic plants (FG72), not treated
C = transgenic plants (FG72), treated
NA: not applicable due to regimen*site interaction

Table 33 Results from T-tests - by-site analysis for total fatty acids

Summary t-test procedures *)	A vs B		A vs C	
	significant	**not significant**	**significant**	**not significant**
C24:0 Lignoceric	2	8	1	9

*) N of sites with significant ($p < 0.05$) and not significant ($p > 0.05$) treatment differences
A = non-transgenic seed from the control Jack
B = transgenic seed from non glyphosate treated FG72 entry
C = transgenic seed from glyphosate treated FG72 entry

Table 34 Vitamins in soybean seed of event FG72 and the non-transgenic counterpart Jack compared to commercial Soybean lines (reference ranges) and results from ANOVA and T-tests (over all sites)

Parameter	Non-Transgenic	Transgenic Non-treated	Transgenic Treated [a]	Range Commercial Soybean Lines [b]	Reference Range [c]	ANOVA (p-values)		T-test (p-values)*	
	Mean ± SD	Mean ± SD	Mean ± SD			regimen effect	regimen*site interaction	A vs.B	A vs. C
Vitamin B1 (mg/kg dm)	3.59 ± 0.76	3.44 ± 0.95	3.16 ± 0.91	1.60 – 4.70	1.01 - 16.02	0.009	0.072	0.279	0.003
Vitamin B2 (mg/kg dm)	4.42 ± 0.88	4.52 ± 0.89	4.80 ± 0.84	3.36 – 6.38	1.9 – 14.5	0.253	0.956		
Folic Acid (mg/kg dm)	2.976 ± 0.353	3.068 ± 0.300	3.122 ± 0.344	2.19 – 4.33	2.4 - 4.7	0.117	0.491		
Vitamin A (mg/kg dm)	0.217 ± 0.047	0.261 ± 0.112	0.284 ± 0.117	< LOQ[d]	0.26 – 4.37	NA	<.001		
Vitamin K (mg/kg dm)	0.191 ± 0.069	0.203 ± 0.078	0.215 ± 0.087	< LOQ[e] – 0.263	0.38 – 0.51	NA	0.030		
Alpha Tocopherol (mg/kg dm)	17.4 ± 3.9	19.0 ± 5.1	20.7 ± 5.8	12.2 – 24.9	2 – 70	NA	0.003		
Gamma Tocopherol (mg/kg dm)	195 ± 16	200 ± 14	198 ± 11	153 – 237	18 – 461	0.038	0.076	0.011	0.132
Delta Tocopherol (mg/kg dm)	74.1 ± 7.4	75.2 ± 8.3	74.0 ± 11.1	41.5 – 99.2	31 – 186	NA	0.014		
Total Tocopherols (mg/kg dm)	286 ± 16	294 ± 14	293 ± 13	225 - 346	120 - 674	0.017	0.130	0.007	0.031

[a] Treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY])
[b] Reference ranges of the 3 analysed commercial soybean lines taken from Table 1c of Appendix A to the statistical report (Rattemeyer, 2009; Appendix 28)
[c] Reference ranges from Table 3a and 3b of Appendix B in the reference (Oberdorder, 2009; Appendix 8)
[d] some samples had no detectable vitamin A, LOQ (Limit of Detection) for Vitamin A was 0.20 mg/kg dm
[e] some samples had no detectable vitamin K, LOQ for Vitamin K was 0.10 mg/kg dm

* only in cases of no interactions (p>0.05) and significant regimen effects (p<0.05)
A = non-transgenic plants (Jack), not treated,
B = transgenic plants (FG72), not treated
C = transgenic plants (FG72), treated
NA: not applicable due to regimen*site interaction

Table 35 Results from T-tests - by-site analysis for vitamins

Summary t-test procedures *)	A vs B		A vs C	
	significant	**not significant**	**significant**	**not significant**
Vitamin B1	1	9	1	9
Vitamin A #)	3	2(5)	4	1(5)
Vitamin K	1	9	1	9
Alpha Tocopherol	2	8	3	7
Gamma Tocopherol	1	9	2	8
Delta Tocopherol	1	9	4	6

*) N of sites with significant ($p < 0.05$) and not significant ($p \geq 0.05$) treatment differences
A = non-transgenic seed from the control Jack
B = transgenic seed from non glyphosate treated FG72 entry
C = transgenic seed from glyphosate treated FG72 entry
#) 'not significant' was also assumed if all samples of a site were equal or below the limit of quantification for the two respective treatments (N of sites in brackets)

Table 36 Isoflavones in soybean seed of event FG72 and the non-transgenic counterpart Jack compared to commercial soybean lines(reference ranges)and results from ANOVA and T-tests (over all sites)

Parameter	Non-Transgenic	Transgenic Non-treated	Transgenic Treated [a]	Commercial Soybean Lines [b]	Literature Reference Range [c]	ANOVA (p-values)		T-test (p-values)*	
						regimen effect	regimen*site interaction	A vs.B	A vs. C
	Mean ± SD	Mean ± SD	Mean ± SD						
Daidzein (mg/kg dm)	11.0 ± 2.0	10.6 ± 1.4	10.3 ± 1.0	< LOQ[d] – 14.6	5 – 35	0.155	0.292		
Genistein (mg/kg dm)	11.5 ± 2.1	11.2 ± 1.8	10.5 ± 0.8	< LOQ – 20.6	0.3 – 46	NA	0.010		
Daidzin (mg/kg dm)	1035 ± 350	1034 ± 356	994 ± 357	568 – 2530	60.0 – 2454	0.320	0.562		
Glycitin (mg/kg dm)	365 ± 39	414 ± 43	400 ± 56	142 – 315	15.3 – 1070	<.001	0.887	<.001	<.001
Genistin (mg/kg dm)	1817 ± 482	1682 ± 465	1640 ± 446	1130 – 3290	144 – 2837	<.001	0.812	<.001	<.001
Total Isoflavones (mg/kg dm)	2010 ± 522	1953 ± 507	1891 ± 488	1160 - 3390	679 – 3733	0.030	0.816	0.201	0.008

[a] Treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY])
[b] Reference ranges of the 3 analysed commercial soybean lines taken from Table 1e of Appendix A to the statistical report (Rattemeyer, 2009)
[c] Reference ranges from Table 4 of Appendix B in the reference (Oberdorder, 2009)
[d] some samples had no detectable isoflavones, LOQ for isoflavones was 10 mg/kg dm

* only in cases of no interactions (p>0.05) and significant regimen effects (p<0.05)
A = non-transgenic plants (Jack), not treated,
B = transgenic plants (FG72), not treated
C = transgenic plants (FG72), treated
NA: not applicable due to regimen*site interaction

Table 37 Results from T-tests - by-site analysis for isoflavones

Summary t-test procedures *)	A vs B		A vs C	
	significant	not significant	significant	not significant
Genistein #)	-	6 (4)	1	5(4)
Total Isoflavones	1	9	2	8

*) N of sites with significant ($p < 0.05$) and not significant ($p \geq 0.05$) treatment differences
A = non-transgenic seed from the control Jack
B = transgenic seed from non glyphosate treated FG72 entry
C = transgenic seed from glyphosate treated FG72 entry
#) 'not significant' was also assumed if all samples of a site were equal or below the limit of quantification for the two respective treatments (N of sites in brackets)

In summary, no safety related issues were identified in the analysis of the nutrient composition of FG72 soybean grain. All components measured were comparable to either the commercial soybean varieties grown at the same locations as the test and control entries, or were within the cited literature reference ranges. All statistically significant differences observed between the non-transgenic control isoline Jack and the transgenic entries of event FG72 were determined not to be of biological relevance from a nutritional standpoint, nor due to the modification of the FG72 event.

(b) The levels of any other constituents that may potentially be influenced by the genetic modification, as a result, for example, of downstream metabolic effects, compared with the levels in an appropriate comparator.

Other than the intended presence of the HPPD W336 and 2mEPSPS proteins in soybean varieties containing event FG72, food products derived from them have been shown to be compositionally and nutritionally equivalent (see Section 3.5(a) directly above).

(c) The levels of any naturally occurring allergenic proteins in the GM food compared with the levels in an appropriate comparator. Particular attention must be paid to those foods that are required to be declared when present as an ingredient, and where significant alterations to protein content could be reasonably anticipated.

One of the specific allergy risks of transgenic crops is associated with alterations to the allergenicity of the whole product, e.g. due to over-expression of natural endogenous allergens as an unintended effect of the genetic modification (EFSA, 2006).

An investigation was undertaken to compare the expression levels of known endogenous soybean allergens (described in AllergenOnline), between the transgenic soybeans (event FG72) and its non-transgenic counterpart, using two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). In order to have a better understanding of the natural variability of these endogenous allergens, the same allergens found in the transgenic soybeans were also compared to commercial non-transgenic soybeans. Therefore, the samples investigated included seeds of the transgenic soybeans (event FG72), the non-transgenic near-isogenic counterpart (variety Jack), and three commercial non-transgenic control soybean lines (Stine® 2686-6, Stine 2788 and Stine 3000-0).

Protein extracts from all samples were separated using 2D-PAGE and quanitified using image analysis. For each, 18 highly abundant polypeptides and 19 less abundant polypeptides were analysed, representing a total of 37 allergen proteins. The 5 known soybean food allergen families (glycinin, Kunitz trypsin inhibitor, β-conglycinin, Bd 28K and Bd 30K), as identified in the AllergenOnline database, were investigated for each sample. Among the 37 polypeptides, the expression level of three allergens (two Kunitz trypsin inhibitor and one β-conglycinin polypeptides) was found to be slightly lower in the FG72 soybean than in Jack. These differences were small, ranging from -21 and -43 % in mean quantity. The level of the two Kunitz trypsin inhibitor allergens (among 7 Kunitz trypsin inhibitor allergens studied) in FG72 soybean fell within the range of natural variability observed in the 4 non-transgenic soybean lines. The level of the β-conglycinin allergen (among 4 β-conglycinin polypeptides studied) was slightly lower than the range of natural variability observed in the 4 non-transgenic soybean lines. However, it cannot be concluded that the decrease in this β-conglycinin allergen level would reduce the allergic potential of FG72 soybean seeds.

Based on this investigation of the allergen content in transgenic and non-transgenic soybean seeds, the allergenic potential of FG72 soybean seeds is similar to that of the non-transgenic soybeans sampled (Rouquié, 2009 Appendix 30)

Part 4 Information Related to the Nutritional Impact of the Genetically-Modified Food

4.1 Data to Allow the Nutritional Impact of Compositional Changes in the Food to be Assessed

Section 3.5(a) provides data from a compositional analysis that was performed to compare the nutritional properties of raw agricultural commodity derived from FG72 soybean. Analysis of the nutritional composition of the double-herbicide-tolerant soybean, transformation event FG72, was completed for soybean grain harvested from 10 locations in the soybean growing areas of North America, representing Group 2-3 soybean varieties. (Kowite, 2009b; Appendix 13). The study was conducted during the 2008 growing season using seed of the T_8 generation. The test plots were each 15 ft by 20 ft in size and contained 6 rows spaced 30 inches apart. At maturity, grain samples were harvested from the four interior rows of each plot.

Planted at each of the 10 sites were three entries:

Entry A. The control isoline variety Jack, a group 2 soybean variety which was treated with conventional herbicides registered for use on soybean.
Entry B. The test entry FG72 treated with conventional herbicides,
Entry C. The test entry FG72 treated with the intended herbicides (isoxaflutole [IFT] and glyphosate [GLY]).

Each of the three entries was planted in a randomized complete block design (RBC) with three replications per location.

Additionally, three non-transgenic commercial soybean varieties were planted along side the test and control entries at the same locations. These three commercial soybean varieties provided reference values to establish ranges of natural variation for the nutritional components analyzed in this study.

The nutritional composition analysis was based on the OECD guidance document for soybean (OECD, 2001c), and also included some minor components. The nutrient composition evaluation of FG72 consisted of measuring 62 parameters measured from 120 samples representing 10 geographic locations and environments. The nutritional parameters analyzed were proximates, fiber compounds, minerals and vitamins, anti-nutrients, isoflavones, total amino acids, and fatty acids.

The statistical analysis of the nutritional composition data was performed to check for differences between the three entries (A, B, and C) (Rattemeyer, 2009; Appendix 28). Each compositional parameter was analysed by Analysis of Variance (ANOVA) methods using a fixed model for the three entries, each location and their interaction term. Subsequent to the ANOVA, t-tests of entries A versus B and A versus C were also conducted. In cases of significant interaction between entries and location, the results of the by-site analyses were taken into account. The level of significance was fixed at 0.05 (two-sided). All analyses were performed using SAS version 8.2 (WINDOWS XP).

For comparative purposes, the values obtained for the commercial reference lines were used to establish in-study ranges in addition to the ranges reported in the published literature (OECD, 2001c; ILSI 2007). Together, these two sets of ranges were used to evaluate the nutritional composition results of the FG72 soybean. Nutrient means that fell within the limits of the commercial or literature reference ranges were considered to be within the normal variation for commercial soybeans.

4.2 Data from an Animal Feeding Study, if Available

A feeding study was conducted for soybean event FG72 using broiler chickens. The growing broiler is a very sensitive test species as a 15-fold increase in body weight occurs during the first 21 days of feeding. Therefore, the broiler is an appropriate species to detect differences in nutrient quality as well as toxic effects of genetically modified seed varieties. The objective of this study is to compare the effects upon ROSS variety chickens (*Gallus gallus domesticus*) of exposure to feed containing either transgenic FG72 toasted soybean seedmeal and the non-transgenic, non-GM counterpart over

a 42-day period. A non-transgenic, commercially available toasted soybean seedmeal variety was also administered in parallel (reference control).

The formulations used the compositional data of the meal groups and were designed to meet specific poultry nutritional requirements under typical local industry husbandry practices. Diets were prepared to ensure that for each toasted soybean type and growth phase, the diets were isoenergetic/isocaloric, isoproteic, and as similar as possible relative to limiting amino acids. For the transgenic diet, seed of generation T_8 was used. Tested diets included: (i) commercial non-genetically modified (non-GM) toasted soybean seedmeal, (ii) FG72 toasted soybean seedmeal, and (iii) non-GM counterpart toasted soybean seedmeal. Each regimen was formulated into separate, nutritionally balanced diets with an incorporation rate of 20%. Diets were formulated to create a starter, grower, and finisher diet phase using each regimen.

All chickens were monitored at least daily for health status, overt signs of toxicity, and mortality. Effects of diets on health, survival, live body weight, total weight gain, feed consumption, food conversion, marketable carcass weight and muscle tissue weight and yield (breast, thigh, leg, wing), and abdominal fat pad weight were compared among groups. Gross post-mortem examination findings were reported as appropriate.

Following 42 days of daily exposure to feed containing FG72 toasted soybean seedmeal (dietary content of approximately 20%), there were no adverse effects detected in feed consumption, feed conversion ratio, survival, body weight gain, or in weight of chilled carcass, legs, thighs, wings or breasts between broiler chickens fed the genetically modified FG72 toasted soybean seedmeal and two control groups consisting of a non-transgenic commercial variety and a non-transgenic non-GM counterpart. The growth and health of chickens on a diet containing FG72 toasted soybean seedmeal were comparable to chickens on two control diets, including a commercial variety of toasted soybean seedmeal and a non-transgenic, non-GM counterpart to the FG72 toasted soybean seedmeal. The detailed findings of this feeding study are provided in Stafford (2009; Appendix 31).

LITERATURE CITED

Copies of the documents listed below can be provided on request.

1. AGBIOS (2009) GM crop database - EPSPS. Ontario, Canada. Available at: http://cera-gmc.org/index.php?evidcode=&hstIDXCode=&gType=&AbbrCode=EPSPS&atCode=&stCode=&coIDCode=&action=gm_crop_database&mode=Submit

2. Ballmer-Weber, B.K. and Vieths, S. (2008) Soy allergy in perspective. *Current opinion in allergy and clinical immunology* 8(3):270-275.

3. Bernard, R.L. and Weiss, M.G. (1973) Qualitative Genetics. In: Caldwell, B.E. (Ed) Soybeans, Production and Uses. Agronomy Series, American Society of Agronomy, Madison, Wisconsin, USA. pp. 117-154.

4. Besler, M., Helm, R.M. and Ogawa, T. (2000) Allergen data collection update: soybean (*Glycine max*). Internet Symposium on Food Allergens, 2000, 2 (Suppl 3) 435. Available at: http://www.food-allergens.de/symposium-2-2/soybean/soy-abstract.htm

5. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L. and Boyer, H.W. (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2:95-113.

6. Boocock, M.R and Coggins, J.R. (1983) Kinetics of senolpyruvylshikimate-3-phosphate synthase inhibition by glyphosate*. FEBS Letters* 154:127-133.

7. Bossis, E., Lemanceau, P., Latour, X. and Gardan, L. (2000) The taxonomy of *Pseudomonas fluorescens* and *Pseudomonas putida*: current status and need for revision. *Agronomie* 20:51-63.

8. 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. US Patent US6245968B1 (12-JUN-2001) AVENTIS CROPSCIENCE S.A.

9. Brownlee, J.M., Johnson-Winters, K., Harrison, D.H.T and Moran, G.R. (2004) Structure of the ferrous form of (4-Hydroxyphenyl) pyruvate dioxygenase from *Streptomyces avermitilis* in complex with the therapeutic herbicide, NTBC. *Biochemistry* 43: 6370-6377.

10. Carrington, J. C. and Freed, D. (1990) Cap-independent enhancement of the translation by a plant potyvirus 5′ non translated region. *Journal of Virology* 64:1590-1597.

11. Caviness, C.E. (1966) Estimates of natural cross-pollination in Jackson soybeans in Arkansas. *Crop Science* 9(2): 211.

12. Chabouté M., 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.

13. Chaubet, N., Clement, B. and Gigot C. (1992) Genes encoding a histone H3.3-like variant in *Arabidopsis* contain intervening sequences. *Journal of Molecular Biology* 225:569-574.

14. Codex Alimentarius Commision (CAC) (2003) Codex principles and guidelines on food derived from biotechnology. Codex Alimentarius Commision, Rome.

15. Codex Alimentarius Commission (CAC) (1999) Report of the XXth session of the Codex Alimentarius Commission - adoption of the food labelling requirements for specific allergens. Conference on International Food Trade Beyond 2000: Science-Based Decisions, Harmonization, Equivalence and Mutual Recognition (Prospects for the Future: Emerging Problems- Food Allergens). Melbourne, Australia , 11-15 October,1999.

16. Compant, S., Duffy, B., Nowak, J., Clement, C. and Barka, E.A. (2005) Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Applied Environmental Microbiology* 71:4951-4959.

17. Cooporative Research Centres (CRC) (1983) Handbook of processing and utilization in agriculture, Volume 2, Part 2 - Plant Products. Department of Innovation, Industry, Science and Research, Canberra.

18. De Groot, M.J.A., Bundock, P., Hooykaas, P.J.J. and Beijersbergen, A.G.M. (1998) *Agrobacterium*-mediated transformation of filamentous fungi. *Nature Biotechology* 16:839–842.

19. Della-Cioppa, G., Bauer, S.C., Klein, B.K., Shah, D.M., Fraley, R.T. and Kishore, G.M. (1986) Translocation of the precursor of the 5-enolpyruvylshikimate-3-phosphate into chloroplasts of higher plants *in vitro. Proceedings of the National Academy of Sciences U.S.A.* 83:6873−6877.

20. 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* 1:561-573.

21. Dotto, G.P., Horiuchi, K. and Zinder, N.D. (1982) Initiation and termination of phage f1 plus-strand synthesis. *Biochemistry* 79:7122-7126.

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

23. Downing, K.J, Leslie, G. and Thomson, J.A. (2000) Biocontrol of the sugarcane borer *Eldana saccharina* by expression of the *Bacillus thuringiensis cry1Ac7* and *Serratia marcescens chiA* genes in sugarcane-associated bacteria. *Applied Environmental Microbiology* 66:2804–2810.

24. European Food Safety Agency (EFSA). 2006. Guidance document of the scientific panel on genetically modified organisms for the risk assessment of genetically modified plants and derived food and feed. May 2006. Available at: http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1211902600026.htm

25. European Union (2003) Directive 2003/89/EC of the European parliament and of the council of 10 November 2003 amending Directive 2000/13/EC as regards indication of the ingredients present in foodstuffs. Office Journal of the European Communities, Brussels.

26. European Union (2000) Directive 2000/13/EC of the European parliament and of the council of 20 March 2000 on the approximation of the laws of the Member States relating to the labelling, presentation and advertising of foodstuffs. Office Journal of the European Communities, Brussels.

27. Fliessbach, A., Winkler, M., Lutz, M.P., Oberholzer, H.R. and Mader, P. (2009) Soil amendment with *Pseudomonas fluorescens* CHA0: lasting effects on soil biological properties in soils low in microbial biomass and activity. *Microbial Ecology* 1-13.

28. Food and Agriculture Organization/World Health Organisation (FAO/WHO) (2001) Evaluation of allergenicity of genetically modified foods. Report of a joint Food and Agriculture Organisation of the United Nation (FAO)/World Health Organisation (WHO) expert consultation on food derived from biotechnology, Rome.

29. Forlani, G., Parisi, B. and Nielsen, E. (1994) 5-enol-pyruvyl-shikimate-3-phosphate synthase from *Zea mays* cultured cells. *Plant Physiology* 105:1107–1114.

30. Fretzdorff, B. and Brümmer, J.-M. (1992) Reduction of phytic acid during breadmaking of whole-meal breads. *Cereal Chemistry* 69, 266–270.

31. Fritze, I.M., Linden, L., Freigang, J., Auerbach, G., Huber, R. and Steinbacher, S. (2004) The crystal structures of *Zea mays* and Arabidopsis 4-Hydroxyphenylpyruvate dioxygenase. *Plant Physiology* 134:1388-1400.

32. Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M. (1987) The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo. Nucleic Acids Research* 15:3257-3273.

33. Gershman, M.D., Kennedy, D.J., Noble-Wang, J., Kim, C., Gullion, J., Kacica, M., Jensen, B., Pascoe, N., Saiman, L., McHale, J., Wilkins, M., Schoonmaker-Bopp, D., Clayton, J., Arduino, M. and Srinivasan, A. (2008) Multistate outbreak of *Pseudomonas fluorescens* bloodstream

infection after exposure to contaminated heparinized saline flush prepared by a compounding pharmacy. *Clinical Infectious Diseases* 47:1372-1379.

34. Gibb, A.P., Martin, K.M., Davidson, G.A., Walker, B. and Murphy, W.G. (1995) Rate of growth of *Pseudomonas fluorescens* in donated blood. *Journal of Clinical Pathology* 48:717-718.

35. Haas, D. and Defago, G. (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nature Review Microbiology* 3:307-319.

36. Hallauer, A.R., Russell, W.A. and Lamkey, K.R. (1988) Corn breeding. In: Sprague, G.F. and Dudley, J.W. (Eds). *Corn and corn improvement*, Third Edition. No 18 in the series Agronomy. American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc., Madison, Wisconsin, USA. pp: 463-564.

37. Hamid, M., Siddiqui, I.A. and Shahid Shaukat, S. (2003) Improvement of *Pseudomonas fluorescens* CHA0 biocontrol activity against root-knot nematode by the addition of ammonium molybdate. *Letters in Applied Microbiology* 36:239-244.

38. Herouet-Guicheney, C., Rouquie, D., Freyssinet, M., Currier, T., Martine, A., Zhou, J., Bates, E.E.M., Ferullo, J-M., Hendrickx, K. and Rouan, D. (2009) Safety evaluation of the double mutant 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) from maize that confers tolerance to glyphosate herbicide in transgenic plants. *Regulatory Toxicology and Pharmacology* 54:143-153.

39. Hitchen, P.G. and Dell, A. (2006) Bacterial glycoproteomics. *Microbiology* 152:1575-1580.

40. Hoheisel, J.D. (1994) Creation of a rare cutter multiple cloning site. *BioTechniques* 17(3):456, 458, 460.

41. Hothersall, J., Wu, J., Rahman, A.S., Shields, J.A., Haddock, J., Johnson, N., Cooper, S.M., Stephens, E.R., Cox, R.J., Crosby, J., Willis, C.L., Simpson, T.J. and Thomas, C.M. (2007) Mutational analysis reveals that all tailoring region genes are required for production of polyketide antibiotic mupirocin by *Pseudomonas fluorescens*: pseudomonic acid B biosynthesis precedes pseudomonic acid A. *Journal of Biological Chemistry* 282:15451-15461.

42. Hui, Y.H. (1992) Encyclopedia of food science and technology. Vol 4. John Wiley and Sons Inc., New York. pp: 2389-2396.

43. Hymowitz, T. and Newell, C.A. (1981) Taxonomy of the genus Glycine, domestication and uses of soybeans. *Economic Botony* 35:272-288.

44. International Life Science Institute (ILSI) (2007) Crop Composition Database - search results version 3.0; Average, minimum and maximum nutrient and anti-nutrient content in soybean seeds. ILSI Crop Composition database. Available at: http://www.cropcomposition.org/cgi-perl/search_ora.cgi

45. Jenkins, N., Parekh, R.B. and James, D.C. (1996) Getting the glycosylation right: implications for the biotechnology industry. *Nature Biotechnology* 14:975-981.

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

47. Kishore, G. and Shah, D. (1988) Amino acid biosynthesis inhibitors as herbicides. *Annual Reviews in Biochemistry* 57:627-663.

48. 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. Patent Application: WO9704103-A 1. Rhone Poulenc Agrochimie, France.

49. Lebrun, M., Leroux, B. and Sailland, A. (1996) Chimeric gene for the transformation of plants. US Patent US5510471. Rhone Poulenc Agrochimie, France.

50. Liener, I.E. (1994) Implications of antinutritional components in soybean foods. *Critical reviews in food science and nutrition* 34(1):31-67.

51. Liu, P.V. (1964) Pathogenicity of *Pseudomonas fluorescens* and related Pseudomonads to warm-blooded animals. *American Journal of Clinical Pathology* 41:150-153.

52. Maga, J.A. (1982) Phytate: its chemistry, occurrence, food interactions, nutritional significance, and methods of analysis. *Journal of Agricultural and Food Chemistry* 30:1-9.

53. McKellar, R.C. (1982) Factors influencing the production of extracellular proteinase by *Pseudomonas fluorescens*. *Journal of Applied Bacteriology* 53:305-316.

54. Messina, M. and Messina, V. (1991) Increasing use of soybeans and their potential role in cancer prevention. *Journal of the American Dietary Association* 91:836–840.

55. Mousdale, D.M. and Coggins, J.R. (1985) High-performance liquid chromatography of shikimate pathway intermediates. *Journal of Chromatography 329*:268-272.

56. Mühldorfer, I. and Hacker, J. (1994) Genetic aspects of *Escherichia coli* virulence. *Microbial Pathology* 16(3):171–181

57. Nickell, C.D., Noel, G., Thomas, D.J. and Waller, R. (1990) Registration of 'Jack' soybean. *Crop Science* 30(6):1365.

58. Organisation for Economic Co-operation and Development (OECD) (2001a) OECD guideline for testing of chemicals.Test guideline N°425: Acute oral toxicity –up and down procedure. December 17, 2001.

59. Organisation for Economic Co-operation and Development (OECD) (2001b) OECD guideline for testing of chemicals.Test guideline N°42: Acute oral toxicity – fixed dose procedure. December 17, 2001.

60. Organisation for Economic Co-operation and Development (OECD) (2001c) Series on the safety of novel foods and feeds N°2. Consensus document on compositional considerations for new varieties of soybean: key food and feed nutrients and anti-nutrients. November 30, 2001.

61. Organisation for Economic Co-operation and Development (OECD) (1997) Series on harmonization of regulatory oversight in biotechnology N°15. Consensus document on the biology of Glycine max (L.) Merr. (soybean). November 14, 2000.

62. Organisation for Economic Co-operation and Development (OECD) (1997) Series on harmonization of regulatory oversight N°6. Consensus document on information used on the assessment of environmental applications involving Pseudomonas. March 4, 1997.

63. Palleroni, N.J. (1981) Introduction to the family *Pseudomonadaceae*. In: Starr, M.P., Stolp, H., Truper, H.G., Balows, A. and Schegel, H.G. (Eds) *The Prokaryotes - A Handbook on Habitats, Isolation and Identification of Bacteria* Chapter 59. Springer-Verlag, Berlin. pp. 655-665.

64. Pardun, H. (1989). Vegetable lecithins - valuable auxiliary and active substances? *Fat Science Technology* 91(2):45-58.

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

66. Peng, R., Xiong, A., Li, X., Fuan, H. and Yao, Q. (2003) A delta-endotoxin encoded in *Pseudomonas fluorescens* displays a high degree of insecticidal activity. *Applied Microbiology and Biotechnology* 63:300-306.

67. Puckett, A., Davison, G., Entwistle, C.C. and Barbara, J.A. (1992) Post transfusion septicaemia 1980-1989: importance of donor arm cleansing. *Journal of Clinical Pathology* 45:155-157.

68. Raaijmakers, J.M., de Bruijn, I. and de Kock, M.J. (2006) Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: diversity, activity, biosynthesis, and regulation. *Molecular Plant-Microbe Interactactions* 19:699-710.

69. Ravindran, V., Ravindran, G. and Sivalogan, S. (1994) Total and phytate phosphorus contents of various food and feedstuffs of plant origin. *Food Chemistry* 50:133-136.

70. Sailland, A., Rolland, A., Matringe, M. and Pallet, K. (2001) DNA sequence of a gene of hydroxy-phenyl pyruvate dioxygenase and production of plants containing a gene of hydroxy-phenyl pyruvate dioxygenase and which are tolerant to certain herbicides. Patent N°: US 6,268,549 B1.

71. Sherlock, O., Dobrindt, U., Jensen, J.B., Munk Vejborg, R. and Klemm, P. (2006) Glycosylation of the self-recognizing *Escherichia coli* Ag43 autotransporter protein. *Journal of Bacteriology* 188:1798-1807.

72. Shutt, D.A. (1976) The effects of plant estrogens on animal reproduction. *Endeavour* 35**:**110–113.

73. Sicherer, S. H. and Sampson, H. A. (2006) Food Allergy. *Journal of Allergy and Clinical Immunology* 117:S470–S475.

74. Skerman, V.B.D., McGowan, V. And Sneath, P.H.A. (1980) Approved lists of bacterial names. *International Journal of Systematic Bacteriology* 30:225-420. Available at: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=bacname

75. Stachel, S.E. and Zambryski, P.C. (1986) *Agrobacterium tumefuciens* and the susceptible plant cell: a novel adaptation of extracellular recognition and DNA conjugation. *Cell* 47:155-157.

76. Steinrücken, H.C, Amrhein, N. (1980) The herbicide glyphosate is a potent inhibitor of 5-*enol*pyruvyl-shikimic acid-3-phosphate synthase. *Biochemical and Biophysical Research Communications* 94:1207-1212.

77. Tacconelli, E., Carmeli, Y., Aizer, A., Ferreira, G., Foreman, M.G. and D'Agata, E.M. (2003) Mupirocin prophylaxis to prevent *Staphylococcus aureus* infection in patients undergoing dialysis: a meta-analysis. *Clinical Infectious Disease* 37:1629-1638.

78. Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.J., Glatt, C.M., Hadfield, N. and Hatzos, C. (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.

79. Thrane, C., Nielsen, M.N., Sorensen, J. and Olsson, S. (2001) *Pseudomonas fluorescens* DR54 reduces sclerotia formation, biomass development, and disease incidence of *Rhizoctonia solani* causing damping-off in sugar beet. *Microbial Ecology* 42:438-445.

80. United States Environmental Protection Agency (US-EPA) (2008a) Frost-preventing bacteria: *Pseudomonas fluorescens* A506 (006438); *Pseudomonas fluorescens* 1629RS (006439); *Pseudomonas syringae* 742RS (006411) Fact Sheet.

81. United States Environmental Protection Agency (US-EPA) (2008b) Frost-Preventing Bacteria: *Pseudomonas fluorescens* A506 (006438); *Pseudomonas fluorescens* 1629RS (006439); *Pseudomonas syringae* 742RS (006411) Product pages.

82. United States Environmental Protection Agency (US-EPA) (2002) Prevention, pesticides and toxic substances (7101), health effects test guidelines OPPTS 870.1100, acute oral toxicology, EPA 712-C-98-190.

83. United States Environmental Protection Agency (US-EPA) (1996) Residue chemistry test guidelines OCCTS 860.1000.

84. United States Environmental Protection Agency (US-EPA) (1994) *Pseudomonas fluorescens* - Tolerance Exemption 8/94. Available at: http://pmep.cce.cornell.edu/profiles/biopestbiocont/biofung/pseudomo_fluor/pseuflu_exm_0894.html

85. United States Environmental Protection Agency (US-FDA) (2006) Guidance for industry: Recommendations for the early food safety evaluations of new non-pesticidal proteins produced by new plant varieties intended in food use.

86. United States Environmental Protection Agency (US-FDA) (2004) Guidance for industry: Questions and answers regarding food allergens, including the food allergen labelling and consumer protection act of 2004 (Edition 4); Final Guidance.

87. United States Pharmacopea (1990) Vol XXII, pp. 1788-1789. United States Pharmacopeia Convention Inc. Rockville, MD.

88. Vaidehi, M.P. and Kadam, S.S. (1989) Soybean. In: Salunkhe, D.K. and Kadam, S.S. (Eds) *Handbook of World Food Legumes: Nutritional Chemistry,Processing Technology and Utilization*. Volume III. CRC Press Inc. Boca Ratón, Florida.

89. Voisard, C., Keel, C., Haas, D. and Defago, G. (1989) Cyanide production by *Pseudomonas fluorescens* helps suppress black root rot of tobacco under gnotobiotic conditions. *EMBO Journal* 8:351-358.

90. Walsh, C.T., Benson, T.E., Kim, D.H. and Lees, W.J. (1996) The versatility of phosphoenolpyruvate and its vinyl ether products in biosynthesis. *Chemistry and Biology* 3(2):83–91.

91. Wilson, M. and Lindow, S.E. (1993) Release of recombinant microorganisms. *Annual Review of Microbiology* 47:913-944.

92. Yang, C., Pflugrath, J.W., Camper, D.L., Foster, M.L., Pernich, D.J. and Walsh, T.A. (2004) Structural basis for herbicidal inhibitor selectivity revealed by comparison of crystal structures of plant and mammalian 4-hydroxyphenylpyruvate dioxygenases. *Biochemistry* 43:10414-10423.

93. Yanisch-Perron, C., Vieira, J., Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.

94. Zhou, M., Xu, H., Zhiqiang Ye, X-W., Wei, L., Gong, W., Wang, Y. and Zhu, Z. (2006) Identification of a Glyphosate-Resistant Mutant of Rice 5-Enolpyruvylshikimate 3-Phosphate Synthase Using a Directed Evolution Strategy. *Plant Physiology* 140(1):184-195.