

## Supporting document 1

### Safety assessment – Application A1081

#### Food derived from Herbicide-tolerant Soybean Line SYHT0H2

---

## Summary and Conclusions

### Background

A genetically modified (GM) soybean line with OECD Unique Identifier SYN-0000H2-5, hereafter referred to as soybean SYHT0H2, has been developed to be tolerant to two herbicides with different modes of action, namely glufosinate-ammonium and mesotrione.

Tolerance to glufosinate ammonium is achieved through expression of the enzyme phosphinothricin acetyltransferase (PAT). PAT is encoded by a *pat* gene obtained from the soil bacterium *Streptomyces viridochromogenes*. Tolerance to mesotrione is achieved through expression of the AvHPPD-03 protein encoded by the *avhppd-03* gene from oat (*Avena sativa*).

In conducting a safety assessment of food derived from soybean line SYHT0H2, a number of criteria have been addressed including: a characterisation of the transferred gene and its origin, function and stability in the soybean genome; the changes at the level of DNA, protein and in the whole food; compositional analyses; evaluation of intended and unintended changes; and the potential for the newly expressed proteins to be either allergenic or toxic in humans.

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

- environmental risks related to the environmental release of GM plants used in food production
- the safety of animal feed or animals fed with feed derived from GM plants
- the safety of food derived from the non-GM (conventional) plant.

### History of Use

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

## **Molecular Characterisation**

Comprehensive molecular analyses of soybean line SYHT0H2 indicate that there is a single insertion site at which there are two inverted and truncated copies of the T-DNA from plasmid pSYN15954 with each copy containing a small added insertion sequence. With respect to the novel genes, there is a single copy of the *avhppd-03* gene and four copies of the *pat* gene. There are also extra copies of various regulatory elements. No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred to the plant. The introduced genetic elements are stably inherited from one generation to the next.

## **Characterisation of Novel Protein**

Soybean line SYHT0H2 expresses two novel proteins, AvHPPD-03 and PAT. Expression analyses of the two proteins showed that both were detected in the plant parts tested. The results were highly variable for both proteins but in general, it can be concluded that both AvHPPD-03 and PAT are present in highest concentration in late vegetative stage (V4) leaves and in lowest concentration (some 20 – 30 times lower than in the V4 leaves) in the seed. Analysis of a range of processed fractions obtained from the seed indicated that processing does not lead to concentration of the proteins

A number of studies were used to confirm the identity and physicochemical properties of the plant-derived AvHPPD-03 and PAT proteins. These studies demonstrated that both proteins conform in size, amino acid sequence and activity to that expected, and do not exhibit any post-translational modification such as glycosylation.

For both proteins, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens; digestibility studies suggest the proteins would be rapidly degraded in the stomach following ingestion; and thermolability studies indicate that both proteins are inactivated by heating. Taken together, the evidence indicates that neither AvHPPD-03 nor PAT is likely to be toxic or allergenic in humans.

## **Herbicide Metabolites**

There is no evidence that the spraying of soybean SYHT0H2 with mesotrione and glufosinate ammonium would result in the production of novel herbicide metabolites that are not already present in conventional crops.

## **Compositional Analyses**

Detailed compositional analyses were done to establish the nutritional adequacy of seed from soybean line SYHT0H2 sprayed with mesotrione and glufosinate ammonium herbicides. Analyses were done of proximates (moisture, crude protein, fat, ash, fibre), amino acids, fatty acids, vitamins, minerals, phytic acid, trypsin inhibitor, lectin, isoflavones, stachyose and raffinose. The levels were compared to levels in the seeds of a non-GM control line ('Jack') grown alongside the GM line.

These analyses did not indicate any differences of biological significance between the seed from soybean SYHT0H2 and the non-GM control 'Jack'.

Statistically significant differences were noted in a number of analytes. However the differences were typically small and all mean values were within both the reference range obtained for non-GM reference varieties grown at the same time and (where it exists) the literature range.

Any observed differences are therefore considered to represent the natural variability that exists within soybean. The spraying of soybean line SYHT0H2 with mesotrione and glufosinate ammonium did not have a significant effect on seed composition.

### **Conclusion**

No potential public health and safety concerns have been identified in the assessment of soybean line SYHT0H2. On the basis of the data provided in the present Application, and other available information, food derived from soybean line SYHT0H2 is considered to be as safe for human consumption as food derived from conventional soybean cultivars.

# Table of Contents

<b>SUMMARY AND CONCLUSIONS.....</b>	<b>1</b>
<b>LIST OF TABLES.....</b>	<b>2</b>
<b>LIST OF FIGURES.....</b>	<b>2</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>3</b>
<b>1. INTRODUCTION.....</b>	<b>5</b>
<b>2. HISTORY OF USE.....</b>	<b>5</b>
2.1 HOST ORGANISM.....	5
2.2 DONOR ORGANISMS.....	6
<b>3. MOLECULAR CHARACTERISATION.....</b>	<b>7</b>
3.1 METHOD USED IN THE GENETIC MODIFICATION.....	8
3.2 DESCRIPTION OF THE INTRODUCED GENES.....	8
3.3 BREEDING TO OBTAIN SOYBEAN LINE SYHT0H2.....	11
3.4 CHARACTERISATION OF THE GENES IN THE PLANT.....	12
3.5 STABILITY OF THE GENETIC CHANGE.....	15
3.6 ANTIBIOTIC RESISTANCE MARKER GENES.....	16
3.7 CONCLUSION.....	16
<b>4. CHARACTERISATION OF NOVEL PROTEINS.....</b>	<b>16</b>
4.1 POTENTIAL ALLERGENICITY/TOXICITY OF NOVEL ORFs CREATED BY THE TRANSFORMATION PROCEDURE.....	17
4.2 FUNCTION AND PHENOTYPIC EFFECTS OF THE AVHPPD-03 AND PAT PROTEINS.....	17
4.3 NOVEL PROTEIN EXPRESSION IN PLANT TISSUES.....	20
4.4 NOVEL PROTEIN EXPRESSION IN PROCESSED FOOD AND FEED.....	21
4.5 PROTEIN CHARACTERISATION STUDIES.....	22
4.6 POTENTIAL TOXICITY.....	26
4.7 POTENTIAL ALLERGENICITY.....	29
4.8 CONCLUSION.....	32
<b>5. HERBICIDE METABOLITES.....</b>	<b>32</b>
<b>6. COMPOSITIONAL ANALYSIS.....</b>	<b>34</b>
KEY COMPONENTS OF SOYBEAN.....	34
6.2 STUDY DESIGN AND CONDUCT.....	35
6.3 SEED COMPOSITION.....	36
6.4 CONCLUSION.....	42
<b>7. NUTRITIONAL IMPACT.....</b>	<b>42</b>
<b>REFERENCES.....</b>	<b>43</b>

## List of Tables

Table 1: Description of the genetic elements contained in the T-DNA of pSYN15954 .....	9
Table 2: SYHT0H2 generations used for various analyses .....	12
Table 3: Average concentration (ug/g dw) over four locations of AvHPPD-03 and PAT proteins in various plant parts from soybean SYHT0H2.....	21
Table 4: Approximate relative amounts of AvHPPD and PAT proteins in processed fractions derived from SYHT0H2 seed.....	22
Table 5: Analyses performed, and test samples used, to compare AvHPPD-03 obtained from SYHT0H2 and E.coli .....	23
Table 6: Analyses performed, and test samples used, to compare PAT obtained from SYHT0H2 and E.coli .....	24
Table 7: Mean percentage of proximates and fibre in seed from 'Jack' and SYHT0H2.....	36
Table 8: Mean percentage composition, relative to total fat, of major fatty acids in seed from 'Jack' and SYHT0H2 .....	37
Table 9: Mean percentage dry weight (dw), relative to total dry weight, of amino acids in seed from 'Jack' and SYHT0H2.....	38
Table 10: Mean weight ( $\mu\text{g/g}$ dry weight expressed as aglycon equivalents) of isoflavones in SYHT0H2 and 'Jack' seed.....	39
Table 11: Mean levels of anti-nutrients in SYHT0H2 and 'Jack' seed.....	39
Table 12: Mean values for mineral levels in seed from 'Jack' and SYHT0H2. ....	40
Table 13: Mean weight (mg/g dry weight) of vitamins in seed from 'Jack' and SYHT0H2....	40
Table 14: Summary of analyte means found in seed of SYHT0H2 treatments that are significantly (adj. $P < 0.05$ ) different from those found in seed of the control line 'Jack' .....	41

## List of Figures

Figure 1: Vector map of plasmid pSYN15954 .....	8
Figure 2: Breeding strategy for plants containing event SYHT0H2 .....	12
Figure 3: Schematic location of the nine PCR products amplified from SYHT0H2.....	13
Figure 4: Map of the SYHT0H2 insert .....	14
Figure 5: Schematic diagram of the action of endogenous HPPD in the absence and presence of mesotrione.....	19
Figure 6: Production of MNBA and AMBA from mesotrione .....	33

## List of Abbreviations

ADF	acid detergent fibre
ADI	acceptable daily intake
ai/ha	active ingredient per hectare
AMBA	2-amino-4-(methylsulfonyl)benzoic acid
AOAC	Association of Analytical Communities
ARfD	Acute Reference Dose
<i>bar</i>	bialaphos resistance
BLAST	Basic Local Alignment Search Tool
BLOSUM	Blocks Substitution Matrix
bp	base pairs
bw	body weight
CaMV	Cauliflower mosaic virus
CmYLCV	Cestrum yellow leaf curling virus
DAD	DNA Data Bank of Japan Aminoacid Database
DMPT	demethylphosphinothricin
DNA	deoxyribonucleic acid
T-DNA	transferred DNA
DTNB	5,5'-dithio-bis[2-nitrobenzoic acid] (Ellman's reagent)
dw	dry weight
ELISA	enzyme linked immunosorbent assay
EU-27	27 member states of the European Union
FAO	Food and Agriculture Organization of the United Nations
FARRP	Food Allergy Research and Resource Program
FASTA	Fast Alignment Search Tool - All
FMV	Figwort mosaic virus
FSANZ	Food Standards Australia New Zealand
fw	fresh weight
GM	genetically modified
HGA	homogentisate
HPP	<i>p</i> -hydroxyphenylpyruvate
HPPD	<i>p</i> -hydroxyphenylpyruvate dioxygenase
kDa	kilo Dalton
LC MS/MS	liquid chromatography-tandem mass spectrometry
LOQ	limit of quantitation
MNBA	4-(methylsulfonyl)-2-nitrobenzoic acid
MRL	maximum residue limit
MT	Million tonnes
MW	Molecular weight
NCBI	National Center for Biotechnology Information
NDF	neutral detergent fibre
<i>nos</i>	nopaline synthase
OECD	Organisation for Economic Co-operation and Development
OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
PAT	Phosphinothricin acetyltransferase
PCR	polymerase chain reaction
L-PPT	L-phosphinothricin
PDB	Protein Data Bank
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid

Ti	tumour inducing
TNB <sup>2-</sup>	2-nitro-5-thiobenzoate anion
TMV	Tobacco mosaic virus
U.S.	United States of America
USDA	United States Department of Agriculture

## 1. Introduction

A genetically modified (GM) soybean line with OECD Unique Identifier SYN-0000H2-5, hereafter referred to as soybean SYHT0H2, has been developed to be tolerant to two herbicides with different modes of action, namely glufosinate-ammonium and mesotrione.

Tolerance to glufosinate ammonium is achieved through expression of the enzyme phosphinothricin acetyltransferase (PAT) encoded by two similar *pat* genes (*pat-03-01* and *pat-03-02*) both derived from the soil bacterium *Streptomyces viridochromogenes*. Tolerance to mesotrione is achieved through expression of the AvHPPD-03 protein encoded by the *avhppd-03* gene from oat (*Avena sativa*). The *pat* gene has been widely used for genetic modification of a number of crop species, including soybean. An HPPD protein has been previously assessed by FSANZ in Application A1051 (FSANZ, 2011) where it was used to confer tolerance in soybean to isoxazole herbicides.

It is anticipated that soybean SYHT0H2 may be grown in the United States of America (U.S.), Argentina, Brazil, China and India subject to approval. The Applicant has not indicated that there is any intention to grow the plant line in Australia or New Zealand.

## 2. History of use

### 2.1 Host organism

The host organism is a conventional soybean (*Glycine max* (L.) Merr.), belonging to the family Leguminosae. The commercial soybean cultivar 'Jack' was used as the parental variety for the genetic modification described in this application, and thus is regarded as the near-isogenic line for the purposes of comparative assessment with soybean SYHT0H2. 'Jack' was released in the U.S. in August 1989 for its resistance to soybean cyst nematode and higher yield when compared to cultivars of similar maturity (Nickell *et al.*, 1990).

Soybean is grown as a commercial food and feed crop in many countries worldwide, with some 91 countries listed as producers in 2011 (FAOSTAT 2012), and has a long history of safe use for both humans and livestock. The major producers of soybeans, accounting for 90% of world production, are the U.S. (83.1 MT), Brazil (74.8 MT), Argentina (48.8 MT), China (14.4 MT) and India (12.2 MT) (FAOSTAT 2012). Australia, while a net importer of soybean, grows crops in latitudes extending from the tropics (16° S) to temperate regions (37° S), mainly in the eastern states and as a rotational crop (James and Rose, 2004). The seed is used mainly to produce meal for use in animal feed (Grey, 2006).

In many soybean producing countries, GM soybean (mainly with a herbicide tolerant trait) accounts for a significant proportion of the total soybean grown e.g. U.S. (91%); Argentina (99%); Brazil (63%); South Africa (87%); Uruguay (99%) (Brookes and Barfoot, 2009). Australia does not currently grow any commercial GM soybean lines<sup>1</sup>.

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

- Whole soybeans are used to produce soy sprouts, baked soybeans, roasted soybeans and traditional soy foods such as miso, tofu, soy milk and soy sauce.

---

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



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

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

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

Another possible food product that can be derived from the soybean plant is bee pollen. This substance is produced by bees during foraging and is taken back to the hive to be fed to larvae and young adult bees (Krell, 1996). It comprises pollen grains that are pelleted by the bee in the corbiculae ('pollen baskets') located on the posterior pair of legs. Beekeepers can collect the pellets by placing a screen at the entrance to a hive; as the bees pass through the screen, the pellets are dislodged and fall into a collection tray. The pellets are frozen or dried for storage and are then packaged for sale as bee pollen, which is generally consumed as the raw product without any further processing.

## **2.2 Donor organisms**

### **2.2.1 *Avena sativa***

The *avhppd-03* gene is sourced from seedlings of the common oat *Avena sativa* (Andrews *et al.*, 2007).

The cereal is grown mainly in temperate regions of the world. Total world production is around 22 MT with EU-27 (particularly Finland and Poland) being the world's largest producer (7.9 MT) followed by the Russian Federation (5.3 MT), Canada (2.9 MT), Australia (1.1 MT) and the U.S. (0.7 MT) (FAOSTAT 2012).

Oat grain is used for both livestock feed (particularly horses) as well as being processed for human consumption. While the use of oats as a human food has not been for as long as other cereals such as wheat and corn, the record appears to go back to at least the first century A.D. (Webster, 2011). Therefore, oats have a long history of safe use as a human food.

These days, oat products include those containing, or derived from, rolled oats (e.g. porridge, muesli) and crushed oats (e.g. oatmeal flour used in bakery products such as oatmeal biscuits and oat bread). The hard outer layer of the oat seed is known as the bran and is also used commonly as an additive in bakery products such as bran muffins. Oats are also used to make beverages; oat milk is available on supermarket shelves, and the beer brewing industry used oats widely although in more recent years this has become limited to specialised oatmeal stouts (Hornsey, 2003).

## 2.2.2 *Streptomyces viridochromogenes*

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

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

## 2.2.3 Other organisms

Genetic elements from several other organisms have been used in the genetic modification of soybean SYHT0H2 (refer to Table 1). These non-coding sequences are used to drive, enhance or terminate expression of the two novel genes. None of the sources of these genetic elements is associated with toxic or allergenic responses in humans. The genetic elements derived from plant pathogens (Figwort mosaic virus, Cauliflower mosaic virus, Cestrum yellow leaf curling virus, Tobacco mosaic virus, *Agrobacterium tumefaciens*) are not pathogenic in themselves and do not cause pathogenic symptoms in soybean SYHT0H2.

## 3. Molecular characterisation

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

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

### Studies submitted:

- deFramond, A. (2012) Event SYHT0H2 soybean insert sequence analysis - Volumes 1 and 2. Unpublished Syngenta Report, May 17, 2012. TK0059645.
- Burgin, K. (2012) Event SYHT0H2 soybean functional element copy number Southern blot analysis. Unpublished Syngenta Report, April 30, 2012. TK0059651.
- de Framond, A. (2012) Event SYHT0H2 flanking sequence determination final report, Volumes 1 and 2. Unpublished Syngenta Report, May 14, 2012. TK0059646.
- Harper, B. (2012) Event SYHT0H2 soybean: Allergenicity and toxicity assessment of start to stop T-DNA ORFs with a minimum size of 30 amino acids. Unpublished Syngenta Report, June 12, 2012. TK0066954.
- Harper, B. (2012). Event SYHT0H2 Soybean: Allergenicity and toxicity assessment of start to stop, genome to insert junction ORFs with a minimum size of 30 amino acids. Unpublished Syngenta Report, June 12, 2012. TK0066953.
- Burgin, K. (2012). Basic Local Alignment Search Tool for Nucleotides (BLASTN) analysis of soybean genomic sequences flanking the insert using the nucleotide collection database and a plant expressed sequence tags (EST) database. . Unpublished Syngenta Report, April 23, 2012. TK0059667.
- New, S. (2011) Event SYHT0H2 soybean: Mendelian inheritance analysis. Unpublished Syngenta Report, October 31, 2011. TK0055859.

### 3.1 Method used in the genetic modification

Explants of soybean cultivar ‘Jack’ were transformed via *Agrobacterium*-mediated transformation (Deblaere *et al.*, 1987) based on the methods of Hwang *et al.* ( 2008) and Que *et al.* ( 2008). The genes of interest were inserted into plasmid pSYN15954 (refer to Figure 1) between DNA sequences known as the Left and Right Borders. These border sequences were isolated from the tumour-inducing (Ti) plasmid of *Agrobacterium tumefaciens* and normally delimit the DNA sequence (T-DNA) transferred into the plant (Zambryski, 1988). During the transformation procedure a portion of each border is expected to be integrated into the host genome (Tzfira *et al.*, 2004).

Basically, explants from immature seeds were co-cultivated with *Agrobacterium tumefaciens* strain EHA101 (Hood *et al.*, 1986) containing the binary vector pSYN15954. Following shoot regeneration, putative transformed shoots were selected on a medium containing glufosinate ammonium as the selection agent. Surviving plantlets were screened for presence of the *pat* and *avhppd-03* genes and absence of the *spec* (spectinomycin) gene (present on the plasmid backbone). Screened shoots were then rooted and transferred to the greenhouse to mature and set seed.

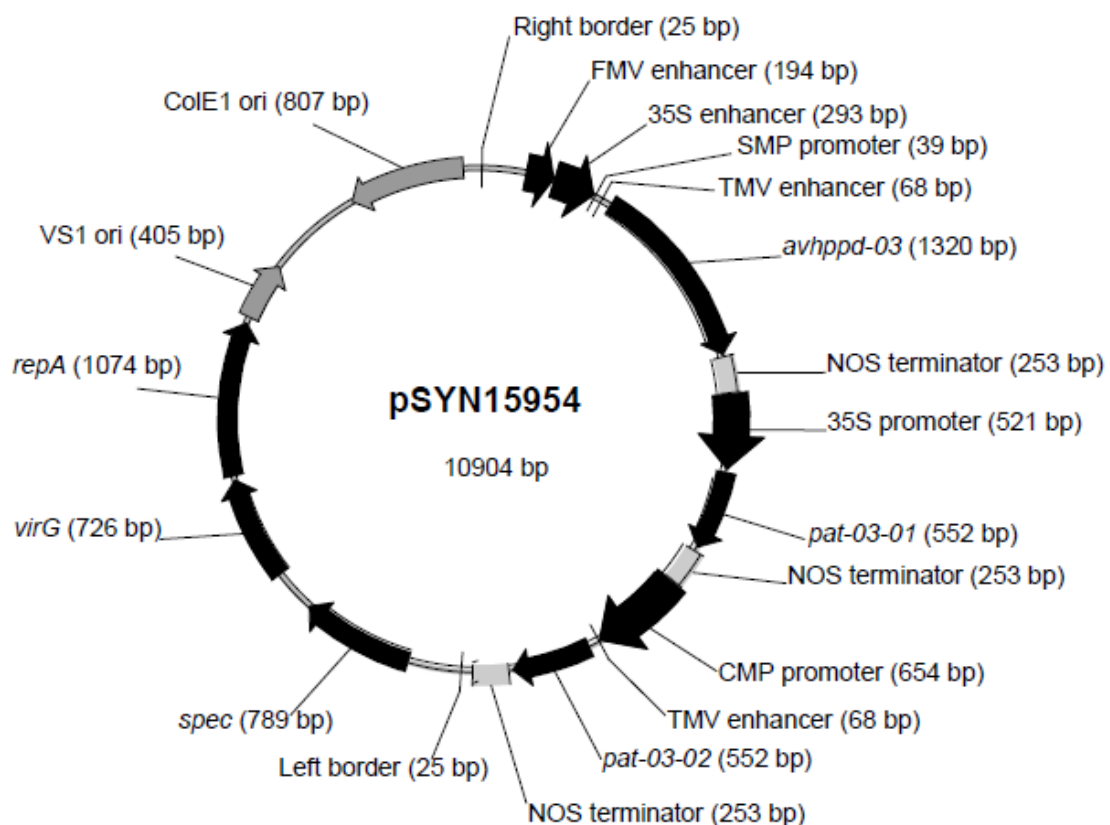


Figure 1: Vector map of plasmid pSYN15954

### 3.2 Description of the introduced genes

Information on the genetic elements in the T-DNA insert is summarised in Table1.

Table 1: Description of the genetic elements contained in the T-DNA of pSYN15954

Genetic element	bp location on pSYN15954	Size (bp)	Source	Orient.	Description & Function	References
Right Border	1 - 25	24	<i>Agrobacterium tumefaciens</i>		• Required for the transfer of the T-DNA into the plant cell	Zambryski (1988)
<b>avhppd-03 cassette</b>						
Intervening sequence	26 - 307	282			• Cloning sequence	
FMV enhancer	308 - 501	194	Figwort mosaic virus	Clockwise	• Transcriptional enhancer region • Facilitates expression of the <i>avhppd-03</i> gene	Maiti et al (1997)
Intervening sequence	502 - 507	6			• Cloning sequence	
35S enhancer	508 - 800	293	Cauliflower mosaic virus	Clockwise	• Transcriptional enhancer region • Facilitates expression of the <i>avhppd-03</i> gene	Ow et al. (1987)
Intervening sequence	801 - 820	20			• Cloning sequence	
SMP promoter	821 - 859	39	Cestrum yellow leaf curling virus	Clockwise	• Synthetic minimal plant promoter including the TATA box (an adenine-rich sequence involved in transcription initiation) linked to a sequence taken from the region that is 3' to the TATA box of the 35S promoter • Drives constitutive expression of the <i>avhppd-03</i> gene	Stavolone et al (2003)
TMV enhancer	865 - 932	68	Tobacco mosaic virus	Clockwise	• 5' non-coding leader sequence • functions as a translational enhancer in plants	Gallie et al. (1987); Gallie (2002)
Intervening sequence	933 - 935	3			•	
<i>avhppd-03</i>	936 - 2255	1320	<i>Avena sativa</i>	Clockwise	• Coding sequence of the <i>avhppd-03</i> gene providing mesotrione tolerance • Codon optimized for expression in soybean	Andrews et al (2007); Hawkes et al (2012)
Intervening sequence	2256 - 2271	16			• Cloning sequence	
<i>nos</i> terminator	2272 - 2524	253	<i>Agrobacterium tumefaciens</i>	Clockwise	• Transcriptional terminator	Depicket et al. (1982)
Intervening sequence	2525 - 2532	8			• Cloning sequence	
<b>pat-03-01 cassette</b>						
35S promoter	2533 - 3053	521	Cauliflower mosaic virus	Clockwise	• Drives constitutive expression of the <i>pat-03-01</i> gene	Ow et al. (1987)
Intervening sequence	3054 - 3077	24			• Cloning sequence	
<i>pat-03-01</i>	3078 - 3629	552	<i>Streptomyces viridochromogenes</i> strain Tü494	Clockwise	• A synthetic gene providing glufosinate ammonium tolerance. • Optimized for plant codon usage.	Wohlleben et al (1988)
Intervening sequence	3630 - 3662	33			• Cloning sequence	
<i>nos</i> terminator	3663 - 3915	253	<i>Agrobacterium tumefaciens</i>	Clockwise	• Transcriptional terminator	Depicket et al. (1982)
Intervening sequence	3916 - 3923	8			• Cloning sequence	
<b>pat-03-02 cassette</b>						
CMP promoter	3924 - 4577	654	Cestrum yellow leaf curling virus	Clockwise	• Promoter and leader sequence • Drives constitutive expression of the <i>pat-03-02</i> gene	Stavolone et al (2003)
Intervening sequence	4578 - 4582	5			• Cloning sequence	

Genetic element	bp location on pSYN15954	Size (bp)	Source	Orient.	Description & Function	References
TMV enhancer	4583 - 4650	68	Tobacco mosaic virus	Clockwise	<ul style="list-style-type: none"> <li>5' non-coding leader sequence</li> <li>functions as a translational enhancer in plants</li> </ul>	Gallie et al. (1987); Gallie (2002)
Intervening sequence	4651 - 4660	10			<ul style="list-style-type: none"> <li>Cloning sequence</li> </ul>	
<i>pat-03-02</i>	4661 - 5212	552	<i>Streptomyces viridochromogenes</i> strain Tü494	Clockwise	<ul style="list-style-type: none"> <li>Providing glufosinate ammonium tolerance.</li> <li>Optimized for plant codon usage.</li> </ul>	Wohlleben et al (1988)
Intervening sequence	5213 - 5240	28			<ul style="list-style-type: none"> <li>Cloning sequence</li> </ul>	
<i>nos</i> terminator	5241 - 5493	253	<i>Agrobacterium tumefaciens</i>	Clockwise	<ul style="list-style-type: none"> <li>Transcriptional terminator</li> </ul>	Depicket et al. (1982)
Intervening sequence	5494 - 5570	77			<ul style="list-style-type: none"> <li>Cloning sequence</li> </ul>	
					<ul style="list-style-type: none"> <li></li> </ul>	
Left Border	5571 - 5595	25	<i>Agrobacterium tumefaciens</i>		<ul style="list-style-type: none"> <li>Required for the transfer of the T-DNA into the plant cell</li> </ul>	Zambryski (1988)

### 3.2.1 *avhppd-03* expression cassette

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

In most plant species, the endogenous HPPD is sensitive to HHPD-inhibiting herbicides but in grass species such as *Avena sativa* (oat), *Brachiaria platyphylla* (signalgrass), *Cenchrus echinatus* (Mossman river grass), *Lolium rigidum* (annual ryegrass), *Festuca arundinacea* (tall fescue), *Setaria faberi* (giant foxtail), *Eleusine indica* (crowsfoot grass) and *Sorghum* sp (sorghum) the endogenous HPPD protein is relatively insensitive (Andrews *et al.*, 2007).

The *hppd* (*avhppd-03*) gene for the genetic modification described for soybean SYHT0H2 was initially isolated and cloned from oat (Andrews *et al.*, 2007; Hawkes *et al.*, 2012) but has been codon optimised for expression in soybean. The expressed protein confers tolerance to the herbicide, mesotrione.

The *avhppd-03* coding region in the T-DNA is 1,320 bp in length and is driven by the constitutive promoter from Cestrum yellow leaf curling virus (CmYLCV), with enhancers from Tobacco mosaic virus (TMV), Cauliflower mosaic virus (CaMV) and Figwort mosaic virus (FMV). The coding region is terminated by a sequence from the 3' end of the *nopaline synthase* (*nos*) gene from *Agrobacterium tumefaciens*.

### 3.2.2 *pat* gene expression cassettes

Two *pat* genes were used in pSYN15954 to ensure an acceptable level of glufosinate tolerance for use as a selectable marker and commercial trait. Both *pat* genes are 552 bp in length but differ by 2 bp. They have been optimised for expression in plants and are based on the *pat* gene from *Streptomyces viridochromogenes*. (Wohlleben *et al.*, 1988; Strauch *et al.*, 1988). Both code for an identical protein that confers tolerance to herbicides containing glufosinate ammonium (phosphinothricin). The *pat* gene has been widely used for genetic modification of food species.

The *pat-03-01* gene is synthetic and was obtained from the company AgrEvo<sup>2</sup>. It is driven by the 35S constitutive promoter from CaMV and is terminated by the 3' end of the *nos* gene sequence. A two base pair difference was introduced into *pat-03-01* to create *pat-03-02*. This change to *pat-03-02* removed two restriction enzyme sites in order to facilitate the cloning process during construction of pSYN15954. The *pat-03-02* gene is driven constitutively by a promoter region of CmYLCV with an enhancer from TMV. It is terminated by the 3' end of the *nos* gene sequence.

Since the two *pat* genes are homologues and give rise to the identical protein, any future reference to the genes in this document will refer to them collectively as a singular *pat* gene (encoding a PAT protein).

### 3.3 Breeding to obtain soybean line SYHT0H2

A breeding programme was undertaken for the purposes of:

- obtaining generations suitable for analysing the molecular and genetic characteristics of soybean SYHT0H2
- ensuring that the SYHT0H2 event is incorporated into elite proprietary breeding line(s) for commercialisation..

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

Following selection of initial transformants ( $T_0$  plants) a series of self-fertilisation and seed bulking crosses proceeded up to generation  $T_8$ . At the  $T_2$  generation, plants were crossed with two non-GM commercial lines to produce an  $F_1$  generation which was either self-fertilised to produce an  $F_2$  generation, or backcrossed to the appropriate parental elite cultivar.

---

<sup>2</sup> AgrEvo was a joint venture company in existence from 1995 – 1999. In 1999, its majority shareholder, Hoechst, merged with the French pharmaceutical and chemical company Rhône-Poulenc to become Aventis CropScience. In 2002, Aventis was acquired by Bayer (now Bayer CropScience). AgrEvo had particular interest in the herbicide glufosinate ammonium and the biotechnology associated with herbicide resistance.

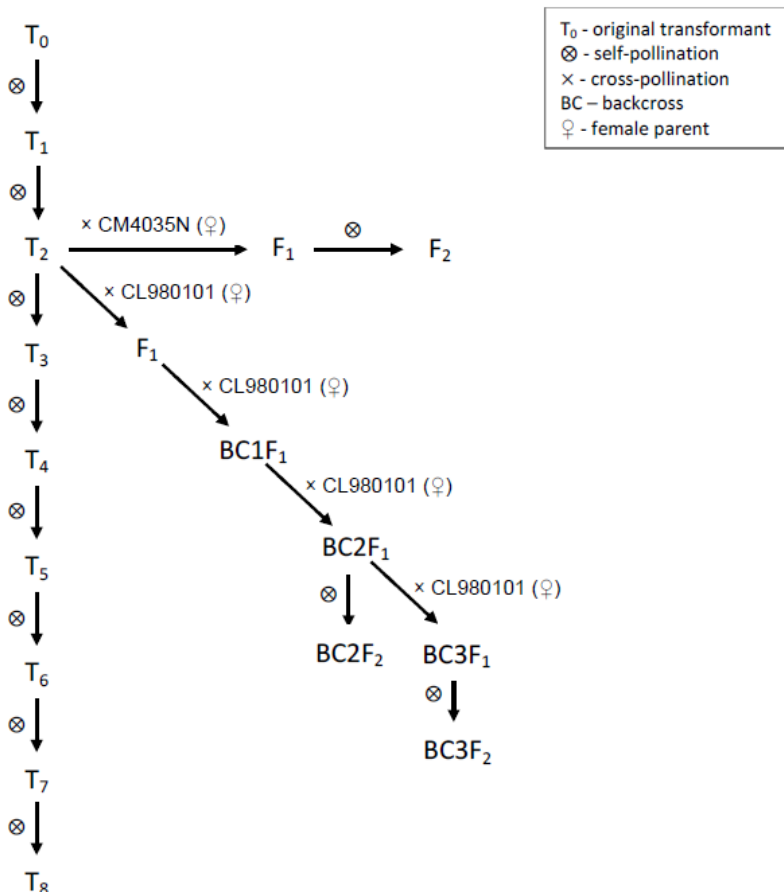


Figure 2: Breeding strategy for plants containing event SYHT0H2

Table 2 indicates the generations that were used in the various studies characterising soybean SYHT0H2.

Table 2: SYHT0H2 generations used for various analyses

Analysis	SYHT0H2 Generation used	Control used
Molecular characterisation	T <sub>4</sub>	'Jack'
Mendelian inheritance	F <sub>2</sub> , BC <sub>2</sub> F <sub>2</sub> , BC <sub>3</sub> F <sub>2</sub>	
Genetic stability	T <sub>4</sub> , T <sub>5</sub> , T <sub>6</sub>	'Jack'
Protein expression in plant parts	T <sub>6</sub>	'Jack'
Protein expression in processed fractions	T <sub>6</sub>	'Jack'
Comparison of plant and microbial proteins	T <sub>6</sub>	
Compositional analyses	T <sub>6</sub>	'Jack'

### 3.4 Characterisation of the genes in the plant

A range of analyses was undertaken to characterise the genetic modification in soybean line SYHT0H2. These included: DNA sequence, determination of insert copy number and integrity; and Open Reading Frame (ORF) analysis of inserted DNA as well as flanking and junction regions.

### 3.4.1 Insert characterisation

Genomic DNA was obtained from pooled polymerase chain reaction (PCR)-verified leaf tissue of T<sub>4</sub> generation soybean SYHT0H2. These samples were used to characterise the DNA sequence in the transgene insertion and its flanking border regions. Because of the complexity of the insert, the genomic DNA was first digested with a restriction enzyme prior to PCR.

Nine overlapping PCR fragments (see Figure 3) spanning the inserted sequences in event SYHT0H2 were amplified and then cloned into a bacterial vector. For each fragment, the DNA from three vector colonies was sequenced individually and the sequences were aligned to obtain a consensus sequence. Commercially available software (Sequencher®) was then used to assemble the nine consensus sequences to obtain a final sequence for the SYHT0H2 insert. This sequence was then compared to the T-DNA sequence in pSYN15954.

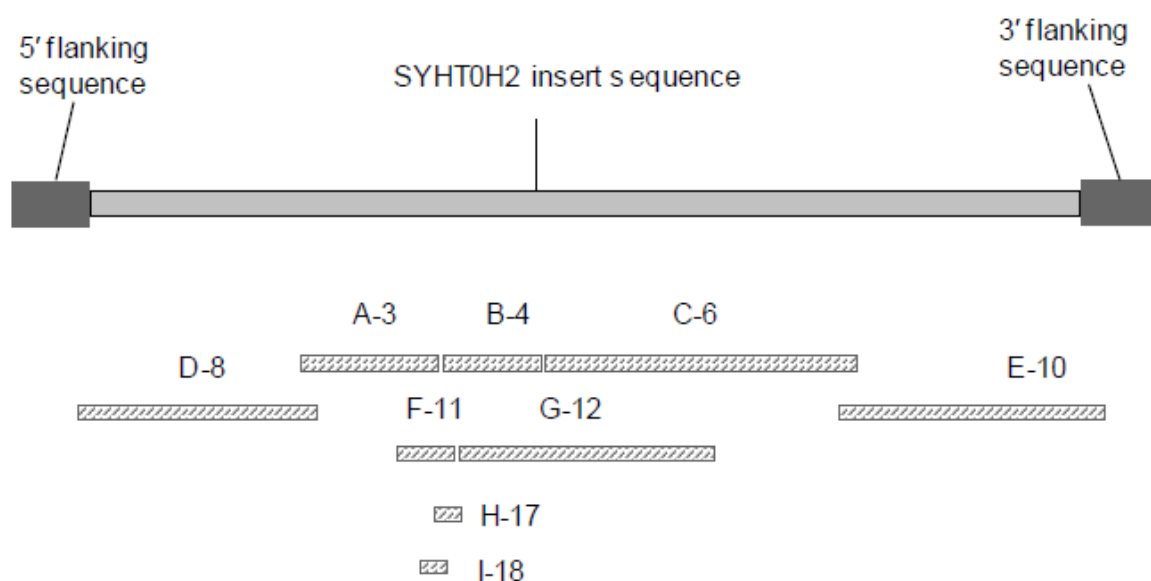


Figure 3: Schematic location of the nine PCR products amplified from SYHT0H2

The sequences at the insertion site in event SYHT0H2 were shown to comprise (see Figure 4):

- Two inverted and truncated copies (called 5' copy and 3' copy) of the pSYN15954 T-DNA
  - The 5' copy lacks a right border, the entire *avhppd-03* cassette, a portion of the CaMV35S promoter (from the *pat-03-01* cassette), and the left border
  - The 3' copy lacks the right border, the FMV enhancer and a portion of the CaMV 35S enhancer, both from the *avhppd-03* cassette, and the left border
- A 44 base pair (bp) insertion located between the two inverted copies. This has similarity to the *av-hppd-03* gene.
- A 17 bp insertion located within the CaMV35S promoter of the 3' copy. The last 15 bp of this insertion are a duplication of a sequence located just upstream.



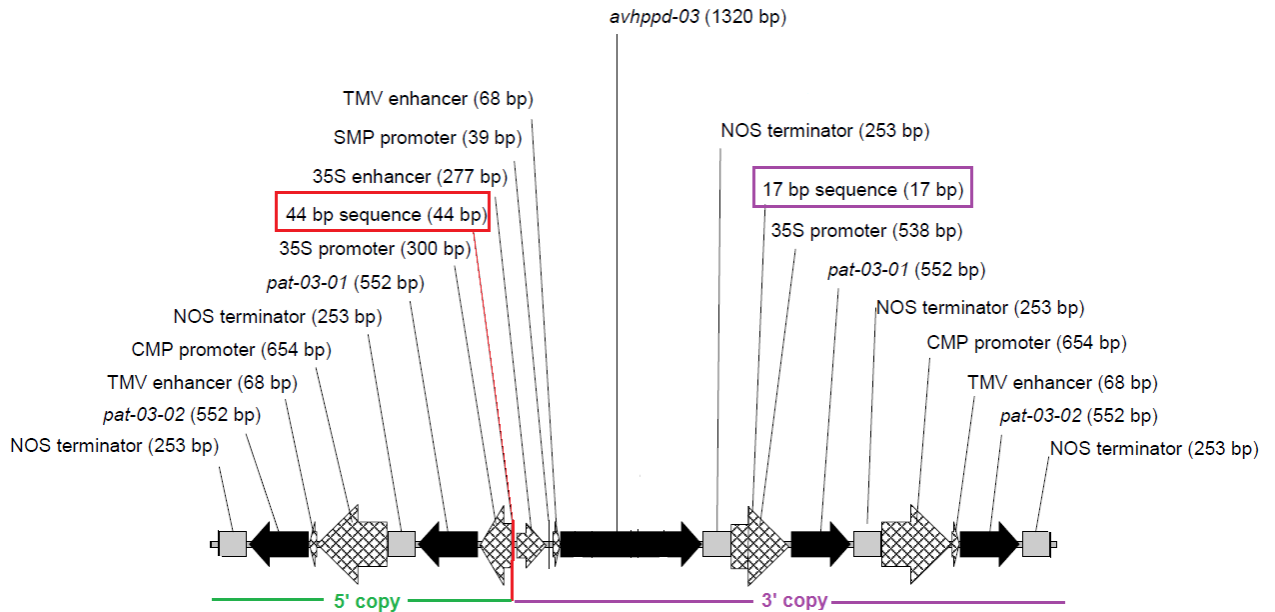


Figure 4: Map of the SYHT0H2 insert

For flanking region analysis, the SYHT0H2 genomic DNA was amplified with primers specific to the 5' and 3' regions flanking the insert. A total of 1,000 bp were then sequenced from each region.

### 3.4.2 Transgene copy number, insertion integrity and plasmid backbone analysis

Total genomic DNA from pooled leaf tissue of PCR-verified soybean SYHT0H2 (T<sub>4</sub> generation) and 'Jack' (negative control) seedlings was used for Southern blot analyses. A positive control (DNA from 'Jack' spiked with *KpnI*- and *PmeI*-digested DNA from the pSYN15954 plasmid) was also included in the Southern blot analyses.

Two restriction enzyme strategies were followed:

- To determine the number of copies of the transgenes and the presence/absence of fragments of these in other regions of the SYHT0H2 genome, DNA was digested with one of four enzymes that cut within the SYHT0H2 insert and within the soybean genome flanking the SYHT0H2 insert.
- To determine the presence/absence of extraneous DNA associated with the transgenes in the insert, the DNA was digested with a restriction enzyme within the insert to release fragments of predictable size.

The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with seven different radiolabelled probes that represented seven functional elements of the T-DNA.

The presence/absence of plasmid backbone was assessed by using a pSYN15954 backbone sequence as a probe following digestion with three restriction enzymes.

The Southern blot analyses indicated that

- SYHT0H2 soybean contains a single copy of *avhppd-03*, four copies of *pat*, a single copy of the *avhppd-03* enhancer complex sequence (consisting of TMV enhancer, synthetic minimal plant promoter, and CaMV35S enhancer sequence), two copies of the CaMV35S promoter, two copies of the CMP promoter, two copies of the TMV enhancer (contained in the *pat-03-02* cassette), and five copies of the *nos* terminator. This confirmed the results of the insert sequence analysis (Section 3.4.1 – Figure 4).

- There are no extraneous DNA fragments from any of the functional elements in other parts of the SYHT0H2 genome.
- There are no pSYN15954 backbone sequences present in the SYHT0H2 genome.

### **3.4.3 Novel open reading frame (ORF) analysis**

Analyses using commercially available software (Vector NTI Advance™) were performed to determine whether any new ORFs that may lead to the production of proteins, had been created either within the insert or in the junction regions. An ORF was defined as a region between a start (ATG) and standard stop (TAA, TAG, TGA) codon with a minimum putative translation size coding for 30 amino acids (the number required to satisfy >35% identity over at least 80 amino acids (FAO/WHO, 2001)).

Within the insert, a total of 52 ORFs were identified of which five encoded AvHPPD-03 and PAT sequences. Only one ORF, putatively coding for 103 amino acids, was identified in the junction region between the 5' end of the SYHT0H2 insert and the endogenous soybean genome. A discussion of the bioinformatic analysis of the 47 novel ORFs within the insert and 1 ORF within the 5' junction region is given in Section 4.1.

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

The Basic Local Alignment Search Tool for Nucleotides (BLASTN)(Agostino, 2013) was used to screen 1,000 bp of DNA sequence flanking each of the 5' and 3' regions of the SYHT0H2 insert for similarity to sequences in a National Center for Biotechnology Information (NCBI) non-redundant nucleotide (nr/nt) database (updated in March 2012 and containing over 15 million sequences) and the PlantGBD Viridiplantae Expressed Sequence Tags database (updated in January 2012 and containing over 21 million sequences).

The results indicated that the SYHT0H2 insert does not disrupt any known endogenous soybean gene.

## **3.5 Stability of the genetic change**

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

Mendelian inheritance was assessed using PCR-verified, greenhouse-grown plants of the F<sub>2</sub>, BC<sub>2</sub>F<sub>2</sub>, and BC<sub>3</sub>F<sub>2</sub> generations (refer to Figure 2) obtained after crossing T<sub>2</sub> SYHT0H2 plants with a non-GM line, CL98010. Genomic DNA was isolated from leaf discs and PCR analysis was done using primers and probes specific to the *avhppd-03* and *pat* genes. A chi-square (X<sup>2</sup>) analysis of the segregation data for each gene over each of the generations was used to test the hypothesis that the genes are inherited according to Mendelian principles i.e that for each generation tested, the segregation ratio for each gene was approximately 3:1 (presence:absence). The X<sup>2</sup> values obtained confirmed that the hypothesis was correct and also supported the conclusion that the SYHT0H2 insert has been stably integrated into a chromosome in the SYHT0H2 genome.

Genetic stability was assessed by Southern blot analyses of genomic DNA isolated from PCR-verified leaf tissue from T<sub>4</sub>, T<sub>5</sub> and T<sub>6</sub> generations (refer to Figure 2) of SYHT0H2 and from 'Jack'. Two probes covering the pSYN15954 T-DNA were utilized following two restriction enzyme strategies that allowed a determination of copy number of the insert, presence/absence of extraneous plasmid fragments either in/near the insert or in other areas of the genome, and intactness of the insert. The analyses showed that the hybridization bands specific to the insert were identical for DNA from the T<sub>4</sub>, T<sub>5</sub> and T<sub>6</sub> generations of SYHT0H2, and hence that the insert is stably inherited from one generation to the next.

### 3.6 Antibiotic resistance marker genes

No antibiotic marker genes are present in soybean SYHT0H2. Plasmid backbone analysis (refer to Section 3.4.2) shows that no plasmid backbone has been integrated into the soybean genome during transformation, i.e. the *spec* gene, which was used as a bacterial selectable marker gene, is not present in soybean SYHT0H2.

### 3.7 Conclusion

Comprehensive molecular analyses of soybean line SYHT0H2 indicate that there is a single insertion site at which there are two inverted and truncated copies of the T-DNA from plasmid pSYN15954 with each copy containing a small inserted sequence. With respect to the novel genes, there is a single copy of the *avhppd-03* gene and four copies of the *pat* gene. There are also extra copies of various regulatory elements. No DNA sequences from the backbone of the transformation vector, including antibiotic resistance marker genes, were transferred to the plant. The introduced genetic elements are stably inherited from one generation to the next.

## 4. Characterisation of novel proteins

In considering the safety of novel proteins it is important to recognise that a large and diverse range of proteins are ingested as part of the normal human diet without any adverse effects, although a small number have the potential to impair health, e.g. because they are allergens or anti-nutrients (Delaney *et al.*, 2008). As proteins perform a wide variety of functions, different possible effects have to be considered during the safety assessment including potential toxic, anti-nutritional and allergenic effects. To effectively identify any potential hazards requires knowledge of the characteristics, concentration and localisation of all novel proteins expressed in the organism as well as a detailed understanding of their biochemical function and phenotypic effects. It is also important to determine if the novel protein is expressed as expected, including whether any post-translational modifications have occurred.

Two types of novel proteins were considered:

- those that may be potentially generated as a result of the creation of novel ORFs during the introduction of the T-DNA of plasmid pSYN15954 (see Section 3.4.3)
- those that were expected to be produced as a result of the expression of the introduced genes. Soybean SYHT0H2 expresses two novel proteins, AvHPPD-03 and PAT.

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

### Study submitted:

Harper, B. (2012) Event SYHT0H2 soybean: allergenicity and toxicity assessment of start to stop T-DNA ORFs with a minimum size of 30 amino acids. Unpublished Syngenta Report, June 12, 2012. TK0066954.

Forty-seven novel ORFs were identified in the T-DNA insert, and one ORF in the 5' junction region, in event SYHT0H2 (refer to Section 3.4.3). The amino acid sequences corresponding to these ORFs were analysed for potential allergenicity and toxicity using an *in silico* approach. These analyses are entirely theoretical since there is no reason to expect that any of the identified ORFs would, in fact, be expressed.

### 4.1.1 Allergenicity assessment

The amino acid sequence of each identified ORF was compared with a peer-reviewed database containing 1,603 known and putative allergens, as well as coeliac-induction sequences residing in the FARRP (Food Allergy Research and Resource Program) dataset (Version 12) within AllergenOnline (University of Nebraska; <http://www.allergenonline.org/>). The allergen search utilised the Fast Alignment Search Tool - All (FASTA) search algorithm with Blocks Substitution Matrix50 (BLOSUM50) scoring matrix (see Section 4.6.2 for explanation) A separate eight-amino-acid search comparing every possible peptide of eight contiguous amino acids in the query sequence with the sequences in the FARRP AllergenOnline database was also carried out.

No similarities with known allergens that exceeded the minimum 35% shared identity over a minimum of 80 amino acids were found. No matches of eight or more contiguous amino acids were found between any sequence and any entry in the FARRP AllergenOnline database.

### 4.1.2 Toxicity assessment

The 48 ORFs were compared with protein sequences in a database containing toxins (created as a subset of the National Center for Biotechnology Information [NCBI] Entrez® Protein database), to identify pair-wise alignments with significant similarity to known putative allergens and toxins. The similarity searches used the BLASTP (Basic Local Alignment Search Tool Protein) algorithm (refer to Section 4.6.2 for an explanation).

No significant sequence similarity was observed between any of the 48 novel ORF sequences and any entry in the toxin database.

### 4.1.3 Conclusion

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

## 4.2 Function and phenotypic effects of the AvHPPD-03 and PAT proteins

### 4.2.1 AvHPPD-03 protein

As indicated in Section 3.2.1, *p*-hydroxyphenylpyruvate dioxygenase (HPPD) occurs in all five kingdoms. In plants, it is involved in the catabolism of the aromatic amino acids

phenylalanine and tyrosine (see Figure 5) and leads to the formation of vitamin E (an antioxidant) and plastoquinones (which are elements of the electron transfer chain of photosynthesis). It is a mononuclear, non-heme, iron-containing enzyme which is a member of the family of 2-oxoacid dependent dioxygenases (Ryle and Hausinger, 2002; Moran, 2005).

HPPD is a target for bleaching herbicides such as the triketone mesotrione (Mitchell *et al.*, 2001; Edmunds and Morris, 2012) which is an extremely potent competitive inhibitor of HPPD. As with all of the HPPD inhibitors that have been developed, mesotrione has a  $\beta$ -diketone moiety which mimics the  $\alpha$ -keto acid group of the substrate HPP (see Figure 5) and therefore competes for binding to HPPD (at the iron at the active site of the enzyme) and inhibits HPPD activity (Matringe *et al.*, 2005). This inhibition leads to tyrosine accumulation, plastoquinone and vitamin E depletion and an accumulation of phytoene, a non-coloured precursor of carotenoids. Since carotenoids are required for photosynthesis and protection of chlorophyll and plant cell membranes during photosynthesis, inhibition of HPPD leads to plant death.

A number of transgenic approaches to conferring tolerance to HPPD-inhibiting herbicides have been taken (see e.g. Matringe *et al.*, 2005; Hawkes *et al.*, 2012) including overexpression of the HPPD enzyme and expression of mutant HPPDs that are less sensitive to the herbicides. It is also known that wild-type monocot HPPDs (such as that of *Avena sativa*) have a lower binding affinity than dicot HPPDs (Hawkes, 2012) and their expression in a dicot host can result in increased tolerance. The use of these wild-type enzymes is preferable in transgenic plants because, in general, they exhibit considerably better stability and activity (kcat/Km) than mutant derivatives (Andrews *et al.*, 2007).

AvHPPD-03 is greater than 99.7% identical in amino acid sequence to the endogenous protein which, itself, has been shown to be more than 2.5 fold tolerant to mesotrione than the homologous protein from the dicot *Arabidopsis thaliana* (Andrews *et al.*, 2007).

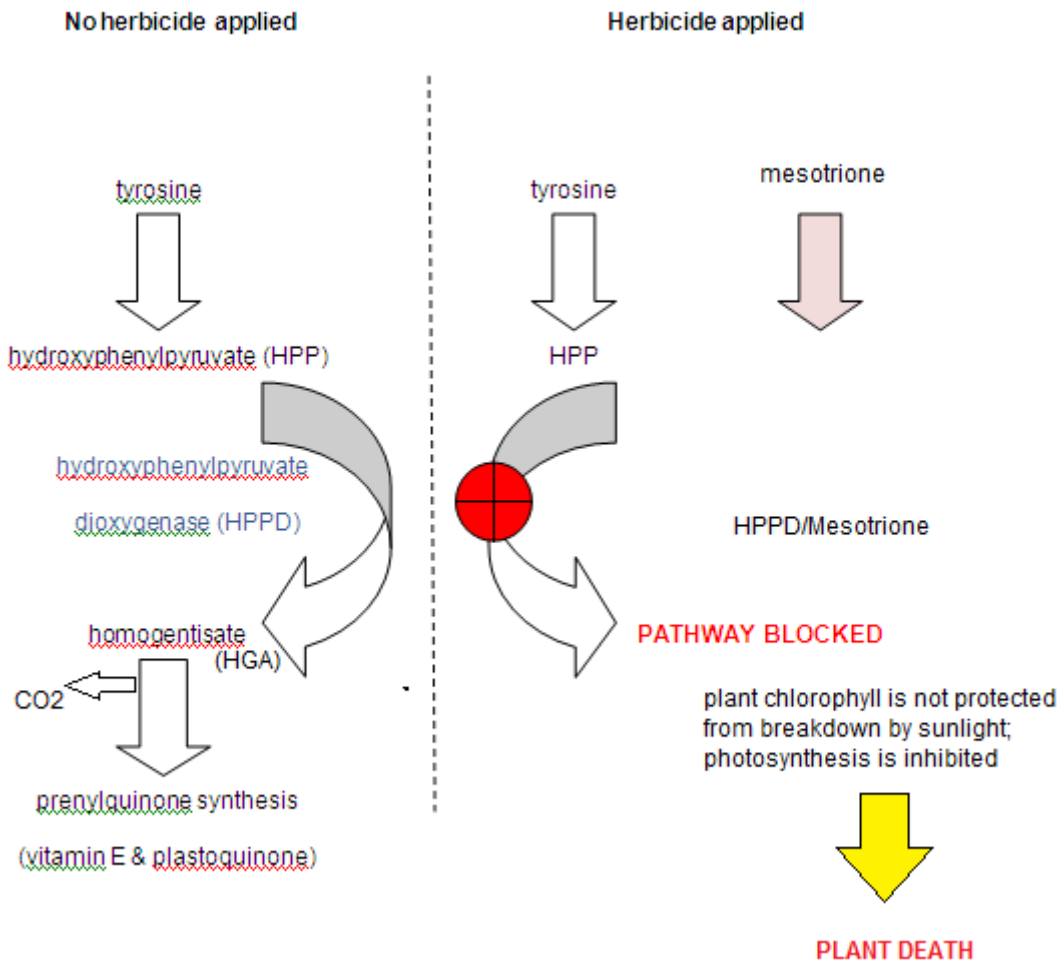


Figure 5: Schematic diagram of the action of endogenous HPPD in the absence and presence of mesotrione

#### 4.2.2 PAT protein

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

The homologous polypeptide produced by the *bar* and *pat* genes (see Section 3.2.2) is known as phosphinothricin acetyltransferase (PAT); it is an acetyl transferase with enzyme specificity for both PPT and demethylphosphinothricin (DMPT) in the acetylation reaction (Thompson *et al.*, 1987). In the presence of acetyl-CoA, PAT catalyses the acetylation of the free amino group of PPT to N-acetyl-PPT, a herbicidally-inactive compound. The kinetics and substrate specificity of the PAT enzyme are well characterised; it has a high specificity for PPT and has been shown to have a very low affinity to related compounds and amino acids; even excess glutamate is unable to block the PPT-acetyltransferase reaction (Thompson *et al.*, 1987).

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

### 4.3 Novel protein expression in plant tissues

#### Study submitted:

McDonald, J. (2012) Quantification of *p*-hydroxyphenylpyruvate dioxygenase and phosphinothricin acetyltransferase in event SYHT0H2 soybean tissues. Unpublished Syngenta Report, March 28, 2012. TK0050018.

The AvHPPD-03 and PAT proteins are expected to be expressed in all plant tissues since the genes encoding them are driven by constitutive promoters (refer to Table 1). Four sites in Argentina<sup>3</sup> (within Provincia de Buenos Aires), representing regions where soybean is commercially grown, were planted with soybean SYHT0H2 (generation T<sub>6</sub>) and 'Jack' during the 2011/2012 growing season. Two herbicide spraying treatments were applied to soybean SYHT0H2 namely, unsprayed, and a single post-emergent spray with mesotrione (0.105 kg ai/ha) and glufosinate (0.322 kg ai/ha) when plants reached the V3-V4 growth stage (see e.g. Iowa State University 2009 for description of growth stages). Samples for analysis of expression of AvHPPD-03 and PAT were taken from a number of plant parts at specific growth stages (refer to Table 3).

The AvHPPD-03 and PAT protein levels were determined by enzyme linked immunosorbent assay (ELISA) using commercial ELISA kits (Qualiplate™ ELISA Kit for HPPD in Soy, made especially for Syngenta's use; Qualiplate™ Kit for Liberty Link PAT/*pat*, EnviroLogix Catalogue No. AP 014). A commercially available software programme was used to calculate the concentrations of immunoreactive AvHPPD-03 and PAT proteins from optical density values.

No AvHPPD-03 or PAT proteins were detected in samples taken from 'Jack' plants. Additionally, no endogenous soybean HPPD was detected in 'Jack', thus confirming the specificity of the ELISA method for AvHPPD-03.

For soybean SYHT0H2 plants, AvHPPD-03 and PAT proteins were detected in all plant parts analysed (Table 3). The results were highly variable for both proteins as reflected in the large standard deviations. In general, it can be concluded that both AvHPPD-03 and PAT are present in highest concentration in V4 leaves and in lowest concentration (some 20 – 30 times lower than in the V4 leaves) in the seed.

---

<sup>3</sup> Gahan; Los Angeles; Inés Indart; and Salto

Table 3: Average concentration ( $\mu\text{g/g dw}$ ) over four locations of AvHPPD-03 and PAT proteins in various plant parts from soybean SYHT0H2

Sample source	Growth Stage	Treatment	Average protein content in $\mu\text{g/g dry weight} \pm\text{SD}$	
			AvHPPD-03	PAT
trifoliolate leaves	V4	unsprayed	242.00 $\pm$ 140.99	52.21 $\pm$ 53.28
		mesotrione + glufosinate	333.45 $\pm$ 243.09	50.29 $\pm$ 41.46
trifoliolate leaves	V8	unsprayed	212.98 $\pm$ 102.03	23.00 $\pm$ 22.84
		mesotrione + glufosinate	198.84 $\pm$ 90.26	18.48 $\pm$ 15.70
trifoliolate leaves	V10	unsprayed	165.14 $\pm$ 66.11	38.23 $\pm$ 31.10
		mesotrione + glufosinate	218.67 $\pm$ 89.43	53.09 $\pm$ 50.91
trifoliolate leaves	R6	unsprayed	105.32 $\pm$ 67.18	29.41 $\pm$ 27.51
		mesotrione + glufosinate	111.00 $\pm$ 65.13	33.89 $\pm$ 40.36
roots	V8	unsprayed	79.49 $\pm$ 47.33	21.16 $\pm$ 18.17
		mesotrione + glufosinate	70.65 $\pm$ 55.73	15.23 $\pm$ 16.78
roots	R6	unsprayed	22.50 $\pm$ 20.82	9.12 $\pm$ 18.61
		mesotrione + glufosinate	31.30 $\pm$ 23.82	11.15 $\pm$ 9.21
forage	R6	unsprayed	79.66 $\pm$ 44.43	19.17 $\pm$ 18.61
		mesotrione + glufosinate	92.99 $\pm$ 36.16	29.73 $\pm$ 21.08
seed	R8	unsprayed	8.18 $\pm$ 8.36	2.70 $\pm$ 4.04
		mesotrione + glufosinate	7.91 $\pm$ 8.30	3.89 $\pm$ 5.45

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

#### 4.4 Novel protein expression in processed food and feed

##### Study submitted:

Horner, A. and McDonald, J. (2012) Quantification of *p*-hydroxyphenylpyruvate dioxygenase and phosphinothricin acetyltransferase in processed food/feed fractions prepared from event SYHT0H2 soybean seed. Unpublished Syngenta Report, May 25, 2012. TK0059702.

PCR-verified seed of SYHT0H2 ( $T_6$  generation) was used to process into fractions (see Table 4) that were then analysed using ELISA methodology (as described in Section 4.3) to quantify the concentration of AvHPPD-03 and PAT present. Refined oil was not analysed because the non-aqueous characteristics of oil prevent accurate protein detection using ELISA. Processing was undertaken using laboratory-scale milling methodology equivalent to industry-standard processing.

Table 4 shows the results for each fraction as a comparison with the level of each protein in seed.



Table 4: Approximate relative amounts of AvHPPD and PAT proteins in processed fractions derived from SYHT0H2 seed

Fraction	Comparison with seed	
	AvHPPD	PAT
seed	20.36 µg/g dw	9.25 µg/g dw
milk	< LOD	< LOD
tofu	< LOD	< LOD
hulls	X 0.1	X 0.1
full-fat flour	X1	X 0.7
flakes	X 0.7	X 0.35
white flakes	X 0.85	X 0.65
defatted toasted meal	X 0.07	X 0.01
protein concentrate	< LOD	< LOD
protein isolate	< LOD	< LOD

Processing into flakes, white flakes, defatted toasted meal, soy protein concentrate and soy protein isolate all require at least one period of heating above 70° C. Processing into milk and tofu involve a heating step of at least 90° C. It would be expected that any protein concentration would be reduced if a heating step were involved in processing. Although hulls do not undergo a heat step during processing, it would also be expected that they contain much less protein than the seed. They are not regarded as a good protein source and animal feeds containing them may require protein supplementation from another source (Lackey, 2011). Full-fat flour is produced by cracking and milling of the seed and would therefore be expected to have similar protein content to the seed.

#### 4.5 Protein characterisation studies

##### Studies submitted:

- Shaw, L. (2012) Comparison of p-hydroxyphenylpyruvate dioxygenase (AvHPPD-03) protein produced in recombinant *Escherichia coli* and AvHPPD-03 protein produced in event SYHT0H2 derived soybean plants. Unpublished Syngenta Report, June 29, 2012. TK0031229.
- Shaw, L. (2012) Comparison of phosphinothricin acetyltransferase (PAT) protein produced in recombinant *Escherichia coli* and PAT protein produced in event SYHT0H2 derived soybean plants. Unpublished Syngenta Report, July 5, 2012. TK0059680.
- Moens, S. (2012). Characterization of the PAT/*pat* protein batch n° 1995 produced in *Escherichia coli*. Unpublished Bayer CropScience study, March 5, 2012. M-426742-01-1

The AvHPPD-03 protein produced by soybean SYHT0H2 would be expected to comprise 439 amino acids and have an approximate molecular weight of 47 kDa. This protein shows close similarity to HPPD proteins from a number of other plant sources, e.g. *Hordeum vulgare* (barley), *Triticum aestivum* (wheat), *Sorghum bicolor* (sorghum), *Zea mays* (corn), *Oryza sativa* (rice). The Applicant states that the AvHPPD-03 protein is 99.7% identical to the native oat HPPD, differing only by a single amino acid residue that is not part of the enzyme's active site.

The *pat* genes used in SYHT0H2 encode the same PAT protein produced by *Streptomyces viridochromogenes* (Uniprot Accession No. Q57146; NCBI Accession No. ZP\_07302142). The protein in SYHT0H2 would therefore be expected to comprise 183 amino acids and have an approximate molecular weight of 21 kDa.

Neither of the proteins is produced in sufficient quantity in soybean SYHT0H2 to isolate for the studies required for a safety assessment. A standard procedure to overcome this type of problem is to produce the protein in a bacterial system and, if this protein shows equivalence

to the *in planta*-produced protein, to then use the bacterially-produced protein as a surrogate for the plant-produced protein. The AvHPPD-03 and PAT proteins were therefore expressed in recombinant *Escherichia coli*. Characterisation tests were done to confirm the identity and equivalence of these bacterially-produced proteins to those produced in soybean SYHT0H2.

AvHPPD-03 and PAT proteins were purified from PCR-verified SYHT0H2 seed of generation T<sub>6</sub> and from *E. coli* bacterial expression systems and the concentrations were determined by ELISA. The following analyses of each protein were then undertaken to confirm their identity, size and biochemical function as well as to compare the plant- and bacterially-produced proteins:

- Western blot
- enzymatic activity
- glycosylation status
- peptide mass mapping
- N-terminal sequencing

The applicant supplied summaries of the test samples and purpose of analysis for each protein (Table 5 and Table 6). The 'extract' referred to in both tables is a crude extract that contained all the proteins in the plant matrix. This crude extract was immunopurified to obtain the 'purified preparation'.

*Table 5: Analyses performed, and test samples used, to compare AvHPPD-03 obtained from SYHT0H2 and E.coli*

<b>Analysis</b>	<b>Samples included in the analysis</b>	<b>Purpose of the analysis</b>
Western blot	<ul style="list-style-type: none"> <li>• microbially produced AvHPPD-03</li> <li>• SYHT0H2 extract</li> <li>• purified preparation from SYHT0H2 extract</li> <li>• nontransgenic extract</li> </ul>	Examine AvHPPD-03 apparent molecular weight, intactness, and relative immunoreactivity
Specific enzymatic activity	<ul style="list-style-type: none"> <li>• microbially produced AvHPPD-03</li> <li>• SYHT0H2 extract</li> <li>• nontransgenic extract fortified with microbially produced AvHPPD-03</li> <li>• nontransgenic extract</li> </ul>	Confirm functional equivalence of both proteins. Confirm correct folding of both proteins.
Glycosylation	<ul style="list-style-type: none"> <li>• microbially produced AvHPPD-03</li> <li>• purified preparation from SYHT0H2 extract</li> </ul>	Confirm the absence of glycosyl residues
Peptide mass mapping	<ul style="list-style-type: none"> <li>• microbially produced AvHPPD-03</li> <li>• purified preparation from SYHT0H2 extract</li> </ul>	Confirm the identity of both proteins
N-terminal sequencing	<ul style="list-style-type: none"> <li>• microbially produced AvHPPD-03</li> <li>• purified preparation from SYHT0H2 extract</li> </ul>	Confirm the N-terminal amino acid sequence of both proteins

Table 6: Analyses performed, and test samples used, to compare PAT obtained from SYHT0H2 and E.coli

Analysis	Samples included in the analysis	Purpose of the analysis
Western blot	<ul style="list-style-type: none"> <li>• microbially produced PAT</li> <li>• SYHT0H2 extract</li> <li>• purified preparation from SYHT0H2 extract</li> <li>• nontransgenic extract</li> <li>• nontransgenic extract fortified with microbially produced PAT</li> </ul>	Examine PAT apparent molecular weight, intactness, and relative immunoreactivity
Specific enzymatic activity	<ul style="list-style-type: none"> <li>• microbially produced PAT</li> <li>• SYHT0H2 extract</li> <li>• nontransgenic extract</li> <li>• nontransgenic extract fortified with microbially produced PAT</li> </ul>	Confirm functional equivalence of both proteins. Confirm correct folding of both proteins.
Glycosylation	<ul style="list-style-type: none"> <li>• microbially produced PAT</li> <li>• purified preparation from SYHT0H2 extract</li> </ul>	Confirm the absence of glycosyl residues
Peptide mass mapping	<ul style="list-style-type: none"> <li>• purified preparation from SYHT0H2 extract</li> </ul>	Confirm the identity of plant-produced PAT
N-terminal sequencing	<ul style="list-style-type: none"> <li>• purified preparation from SYHT0H2 extract</li> </ul>	Confirm the N-terminal amino acid sequence of plant-produced PAT

### Western blot

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

- For AvHPPD-03: polyclonal rabbit anti-AvHPPD-03 followed by polyclonal donkey anti-rabbit alkaline phosphatase secondary antibody.
- For PAT: polyclonal goat anti-PAT followed by polyclonal donkey anti-goat alkaline phosphatase secondary antibody.

For both immunopurified proteins from both sources there was staining of a single sharp and immunoreactive band at the expected molecular weight (i.e. approx. 48 kDa for AvHPPD-03 and approx. 20 kDa for PAT). For the PAT preparations from both sources there was also faint staining of a band at approximately 40 kDa; this is consistent with the presence of a PAT dimer. As expected for the crude extract from SYHT0H2, the bands, while occurring at the expected molecular weights, were diffuse and reflected matrix interference.

The Western blot analysis for each protein confirmed the plant- and microbial-derived proteins were of equivalent size and immunoreactivity.

### Enzymatic activity

The HPPD enzyme catalyzes the formation of homogentisate (HGA) and carbon dioxide from *p*-hydroxyphenylpyruvate (HPP) and molecular oxygen (refer to Figure 5). The HPPD enzymatic activity assay determines the amount of radiolabeled <sup>14</sup>CO<sub>2</sub> generated from a <sup>14</sup>C-labelled HPP substrate during the enzymatic reaction (Lindstedt and Odelhög, 1987).

PAT catalyzes the transfer of the acetyl group from acetyl coenzyme A (acetyl CoA) to phosphinothricin. The released free thiol (CoASH) reacts with 5,5'-dithio-bis[2-nitrobenzoic acid] (DTNB – known as Ellman's Reagent) to form 2-nitro-5-thiobenzoate anion (TNB<sup>2-</sup>) under mild alkaline conditions (pH 7-8) (see e.g. Riddles *et al.*, 1983). The molar formation of

TNB<sup>2-</sup> can be monitored by measuring absorbance increase at 412 nm and can be directly converted into the molar acetylation of phosphinothricin by PAT.

Crude extract from SYHT0H2 showed no detectable PAT activity and slight activity for HPPD (as expected from the endogenous soybean HPPD that would be present). The immunopurified AvHPPD-03 and PAT proteins from both plant and microbial sources showed the expected catalytic activity in the presence of the appropriate substrate. This confirmed the equivalent functional activity of each protein from the two sources.

### *Glycosylation*

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

Analysis of the plant- and microbial-derived proteins was done using a commercial kit (ECL™ Glycoprotein Detection Module Kit) following sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE). A glycosylated protein (transferrin) was applied to each gel as a positive control while the non-glycosylated protein, soybean trypsin inhibitor, was used as a negative control. A visible band was obtained for transferrin while the soybean trypsin inhibitor and both immunopurified novel proteins from both plant and microbial sources gave no bands. It should be noted that the non-virulent *E. coli* strains used in the laboratory for cloning and expression of novel proteins lack the necessary biochemical machinery for protein glycosylation (Wacker *et al.*, 2002; Abu-Qarn *et al.*, 2008) and therefore glycosylation would not be expected in the microbially-derived AvHPPD-03 and PAT.

These results support the conclusion that neither the microbially nor SYHT0H2-derived AvHPPD-03 and PAT proteins are glycosylated.

### *Peptide mass mapping*

Purified AvHPPD-03 proteins from both sources and purified PAT from SYHT0H2 were run on SDS-PAGE and the corresponding protein bands were excised and digested with trypsin and chymotrypsin (as well as flavastacin, in the case of PAT) followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). The masses of the detected peptides were searched (MASCOT® Software) against a protein database containing the expected amino acid sequence of either AvHPPD-03 or PAT.

In a separate study for microbially-derived PAT, the sample was digested with trypsin followed by liquid chromatography-ultra violet mass spectrometry (LC-UV-MS) peptide mapping.

For microbially-derived AvHPPD-03, sequence coverage was approximately 65% of the theoretical total peptide sequence of AvHPPD-03 while for the plant-derived protein, coverage was 55%. The analysis confirmed the identification of both the soybean SYHT0H2- and *E.coli*-derived proteins as AvHPPD-03. It also indicated that the plant-derived AvHPPD-03 lacked the first four amino acids at the N-terminus.

For plant-derived PAT, coverage was 63%. The analysis confirmed the identification of the soybean SYHT0H2-derived protein as PAT and also showed that there has been cleavage of the N-terminal methionine.

For microbially-derived PAT, coverage was 76% and the identity of the sample was confirmed.

### *N-terminal sequencing*

Purified AvHPPD-03 proteins from both sources and purified PAT from SYHT0H2 were run on SDS-PAGE and then electro-blotted onto a polyvinylidene difluoride (PVDF) membrane. The bands were excised and subjected to Edman degradation using an automated pulsed-liquid sequencer. The microbially-derived PAT was sequenced in a separate study.

For the AvHPPD-03 proteins from both sources, 12 residues were sequenced. The sequencing data were complex but could be linked to the expected sequence for AvHPPD-03. The data also confirmed that the first four amino acids in the plant-derived protein were lacking. The results of the enzymatic activity assay of AvHPPD (see above) showed that the loss of these amino acids did not affect the functioning of the enzyme.

For the plant-derived PAT protein, 10 N-terminal residues were sequenced and were consistent with the predicted N-terminal sequence of PAT except for missing the methionine at the N-terminus (in agreement with the peptide mass mapping). The terminal methionine is routinely cleaved from nascent proteins by methionine aminopeptidase (Polevoda and Sherman, 2000). Five N-terminal residues of the microbially-derived PAT were sequenced and showed consistency with the expected sequence for PAT.

### *Conclusion*

A range of characterisation studies confirmed the identity and equivalence of the AvHPPD-03 and PAT proteins produced in both a bacterial expression system and in soybean SYHT0H2. Based on weight-of-evidence, it is concluded that microbially-derived AvHPPD-03 and PAT proteins are suitable surrogates for use in safety assessment studies.

## **4.6 Potential toxicity**

While the vast majority of proteins ingested as part of the diet are not typically associated with toxic effects, a small number may be harmful to health. Therefore, if a GM food differs from its conventional counterpart by the presence of one or more novel proteins, these proteins should be assessed for their potential toxicity. The main purpose of an assessment of potential toxicity is to establish, using a weight of evidence approach, that the novel proteins will behave like any other dietary protein.

The assessment focuses on:

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

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

#### 4.6.1 History of human consumption

As outlined in Section 4.2 and Section 4.5, homologues of the AvHPPD-03 protein are found in plants, including common crop plants, and a range of other organisms and would therefore be routinely consumed as a normal part of the diet.

*Streptomyces hygroscopicus* and *S. viridochromogenes* (refer to Section 2.2.2) are common soil bacteria, therefore humans have a long history of exposure to the PAT protein through the consumption of roots and vegetables. Since 1995, humans have also been directly exposed to the PAT protein through the consumption of foods derived from GM glufosinate ammonium-tolerant canola, soybean, cotton and corn, without any evidence of toxicity (H erouet *et al.*, 2005; Delaney *et al.*, 2008).

#### 4.6.2 Amino acid sequence similarity to known protein toxins

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

##### Studies submitted:

Harper, B. (2010) AvHPPD-03: assessment of amino acid sequence similarity to known or putative toxins. Unpublished Syngenta Report, July 30, 2010. SSB-210-10.

Capt, A. (2012) Pat/pat amino acid sequence homology search with known toxins. Unpublished Bayer Report, March 5, 2012. M-266573-04-1

#### AvHPPD-03

A similarity search was done for the AvHPPD-03 protein, using the BLASTP<sup>4</sup> (Basic Local Alignment Search Tool Protein) algorithm (Altschul *et al.*, 1997), and BLOSUM62 scoring matrix, against known protein sequences present in the NCBI Entrez<sup>®</sup> Protein database (NCBI, 2010) that contained over 10 million sequences.

The BLOSUM series of matrices tabulate the frequency with which different substitutions occur in conserved blocks of protein sequences and are effective in identifying distant relationships (Henikoff and Henikoff, 1992). BLOSUM62 is the default for the BLAST programmes and is derived from blocks that are  $\leq 62\%$  identical.

BLASTP is now frequently applied for searching for similarities in protein sequences by performing local alignments of domains or short sequence similarities; this detects more similarities than would be found using the entire query sequence length (439 amino acids in the case of AvHPPD-03 and 183 in the case of PAT). The search generates a parameter known as the *E* value (see eg Baxevanis, 2005). Comparisons between highly homologous proteins yield *E*-values approaching zero, indicating the very low probability that such matches would occur by chance. A larger *E*-value indicates a lower degree of similarity. All database sequences with an *E*-value of 1 or lower were identified by default by the BLASTP program. Although a statistically significant sequence similarity generally requires a match with an *E*-value of less than 0.01 (Pearson, 2000), setting a higher threshold (*E*-value of 0.15, in this case) ensures that proteins with even limited similarity will not be excluded. Commonly, for protein-based searches, hits with *E*-values of  $10^{-3}$  or less and sequence identity of 25% or more are considered significant although any conclusions reached need to be tempered by an investigation of the biology behind the putative homology (Baxevanis, 2005).

---

<sup>4</sup> BLASTP is used to compare a protein sequence with a database of protein sequences.

The AvHPPD-03 similarity search identified 916 alignments with an E value of < 0.15, of which 775 were identified as HPPD or similar proteins (including glyoxylases and members of the dioxygenase superfamily). Fourteen sequences were related to bacterial hemolysins. Research has indicated that bacterial hemolysin proteins themselves are not directly involved in haemolytic activity but rather that it is oxidation or polymerisation of the homogentisate formed by the action of the proteins that leads to toxicity (Heoedus and Nayak, 1994; Chang *et al.*, 1997; Steinert *et al.*, 2001) i.e. HPPD is not directly haemolytic and there is no evidence that the aerobic HPPD metabolic pathway in plants leads to metabolites associated with haemolysis. There were 43 sequences associated with non-toxic proteins of various function, and 85 sequences were identified as hypothetical, unknown or with unknown function.

### *PAT*

The amino acid sequence of PAT was compared with protein sequences present in a Bayer CropScience toxin database as well as in a number of large public reference databases - Uniprot\_Swissprot, Uniprot\_TrEMBL, PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept. The similarity search used the FASTA algorithm (Pearson and Lipman, 1988) and BLOSUM50 scoring matrix. The FASTA algorithm is essentially the same as the BLAST algorithm. BLOSUM50 is the default scoring matrix used by the FASTA family of sequence comparison programs and reports substitution frequencies for residues in conserved blocks of sequences that show  $\leq 50\%$  identity. An E-value threshold of 0.1 was used for the comparison against the public protein databases, and an E-value threshold of 10 was used for the comparison against the Bayer toxin database.

As expected, the PAT protein showed a high degree of homology with other acetyltransferases, none of which is considered to be a toxin. There were also matches with putative uncharacterised proteins, as well as with complete genomes of various organisms or from complete sequences of chromosomes – which therefore do not correspond to an identified protein.

### *Conclusion*

Neither AvHPPD-03 nor PAT has significant similarity with known toxins that are harmful to human health.

#### **4.6.3 *In vitro* digestibility**

See Section 4.7.3.

#### **4.6.4 *Stability to heat***

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

#### **Study submitted**

Winslow, S. (2010) Effect of temperature on the enzymatic activity of *p*-hydroxyphenylpyruvate dioxygenase (Av-HPPD-03) protein. Unpublished Syngenta Report, December 2, 2010. TK0022136.

Microbially-derived AvHPPD-03 (refer to Section 4.5) was incubated at 4°, 25°, 37°, 65° or 95° C for 30 min. Enzymatic activity was then determined (refer to Section 4.5). The results

showed there was no loss of activity at 25° but that at 37° approximately 74% of activity was lost and at higher temperatures activity was below the limit of detection.

The Applicants did not supply a study for PAT. However, it has been established (Wehrmann *et al.*, 1996) that the PAT protein is completely inactivated after 10 min at 50° C (Hérouet *et al.*, 2005).

#### **4.6.5 Acute oral toxicity study**

##### **Study submitted**

Eapen, A.K. (2012) AvHPPD-03: single dose oral (gavage) toxicity study in mice with a 2-day or 14-day observation period. Unpublished Syngenta Report, March 8, 2012. T007563-08.

An acute oral toxicity study in mice, using bacterially-produced AvHPPD-03 protein was submitted by the Applicants but is not included in this safety assessment since no safety concerns were identified in any of the other studies. Similarly, an acute oral toxicity study is not required for the PAT protein.

#### **4.7 Potential allergenicity**

The potential allergenicity of novel proteins was evaluated using an integrated, step-wise, case-by-case approach relying on various criteria used in combination. This is because no single criterion is sufficiently predictive of either allergenicity or non-allergenicity (see e.g. Thomas *et al.*, 2009). The assessment focuses on:

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

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

##### **4.7.1 Source of the protein**

The AvHPPD-03 protein is derived from the common oat, *Avena sativa*. Oat contains no endogenous proteins that are listed in the FARRP Allergen Protein Database (<http://www.allergenonline.org/>, accessed May 2013). However, there is some debate about whether oats play a role in gluten intolerance in those suffering from coeliac disease (Health Canada, 2007; Rashid *et al.*, 2007; Coeliac Research Fund, 2008). Prolamin fractions in cereal seeds are associated with the autoimmune reaction in coeliac individuals and oat contains avenin (Kilmartin *et al.*, 2003). However, the percentage of total protein of this prolamin in oats is much lower than that of the prolamins found in wheat, barley and rye (Kilmartin *et al.*, 2003) which, traditionally, are associated with gluten intolerance. The implication is that oat therefore contains fewer immunogenic peptides (Ciccocioppo *et al.*, 2005). Oats have been tested in extensive and long-term intervention studies and the intake of moderate amounts of oats as part of a gluten-free diet has been proven as safe.



The HPPD proteins from oat or other species, have not been implicated in any food-related allergic reactions.

The PAT protein is derived from a common soil bacterium *Streptomyces viridochromogenes* which has not been reported to be a source of allergenic proteins. Bacterial proteins are rarely allergenic because of the low exposure levels and lack of allergic sensitisation (Taylor, 2002).

#### **4.7.2 Amino acid sequence similarity to known allergens**

Bioinformatic analysis is part of a 'weight of evidence' approach for assessing potential allergenicity of novel proteins introduced to GM plants (Thomas *et al.*, 2005; Goodman, 2006). It is a method for comparing the amino acid sequence of the introduced protein with sequences of known allergens in order to indicate potential cross-reactivity between allergenic proteins and the introduced protein. As with the bioinformatic analysis that looked at similarities of the novel proteins with known protein toxins (refer to Section 4.6.2), the generation of an *E* value provides an important indicator of significance of matches (Pearson, 2000; Baxevanis, 2005). The criteria used to indicate potential allergenicity were a minimum of eight-contiguous amino acid identity or 35% identity on a window of 80 amino acids within the sequence of an allergenic protein (FAO/WHO, 2001; Codex, 2003).

##### **Studies submitted:**

McClain, S. (2012) AvHPPD-03: assessment of amino sequence similarity to known or putative allergens. Unpublished Syngenta Report, April 3, 2012. TK0059704.

Capt, A. (2012) PAT/pat amino acid sequence homology search with known allergens. Unpublished Bayer Report, March 5, 2012. M-217013-05-1.

The AvHPPD-03 and PAT sequences were compared with all known allergen sequences contained in a reference allergen database, (FARRP version 12 – released in February 2012 and containing 1,603 non-redundant entries) using the FASTA algorithm and BLOSUM50 scoring matrix (refer to Section 4.6.2).

In separate searches the AVHPPD-03 sequence was screened using a proprietary programme developed by Syngenta and the PAT sequence was screened using the SeqMatchAll tool from the European Molecular Biology Open Software Suite (Rice *et al.*, 2000). Both searches compared every possible peptide of eight contiguous amino acids in the AvHPPD-03 or PAT sequence with the sequences in the FARRP AllergenOnline database.

No matches were found for either protein with known allergenic proteins or with known allergenic epitopes and neither did the AvHPPD-03 or PAT proteins share a sequence of eight or more consecutive identical amino acids with any potential allergens.

#### **4.7.3 In vitro digestibility**

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

response. However, evidence of slow or limited protein digestibility does not necessarily indicate that the protein is allergenic.

**Studies submitted:**

- Winslow, S. (2010) *In vitro* digestibility of *p*-hydroxyphenylpyruvate dioxygenase (Av-HPPD-03) protein under simulated mammalian gastric conditions. Unpublished Syngenta Report, August 24, 2010. TKR0000157.
- Winslow, S. (2009) *In vitro* digestibility of *p*-hydroxyphenylpyruvate dioxygenase (Av-HPPD-03) protein under simulated mammalian intestinal conditions. Unpublished Syngenta Report, May 4, 2009. SSB-060-09.
- Rasclé, J.B. (2009) PAT/*pat* protein *in vitro* digestibility study in human simulated gastric fluid. Unpublished Bayer CropScience Report, June 29, 2009. M-214025-04-1.
- Rasclé, J.B. (2009) PAT/*pat* protein *in vitro* digestibility study in human simulated intestinal fluid. Unpublished Bayer CropScience Report, June 24, 2009. M-214021-03-1.

For both proteins, analyses using simulated gastric fluid – SGF (containing pepsin) and simulated intestinal fluid - SIF (containing pancreatin, a mixture of enzymes including amylase, trypsin, lipase, ribonuclease and protease) were done. The SIF study by itself may not be entirely informative because ordinarily an ingested protein would first be exposed to pepsin-mediated hydrolysis in the acidic environment of the stomach before being subject to further digestion in the small intestine.

*Simulated gastric fluid (SGF)*

A pepsin digestibility assay (Thomas *et al.*, 2004) was conducted to determine the digestive stability of the AvHPPD-03 and PAT proteins purified from a bacterial system using SGF (U.S.Pharmacopeia, 1990; U.S.Pharmacopeia, 2000). For AvHPPD-03 the test material was evaluated following incubation in SGF at 37° for 0, 1, 2, 5, 10, and 60 minutes. For PAT, incubation times at 37° C were 0.5, 2, 5, 10, 20, 30 and 60 min.

The samples were then run on SDS-PAGE. Proteins were visualised by Coomassie Blue staining of the resulting gels. Western blotting of the SDS gels was also performed using an appropriate rabbit polyclonal primary antibody and either an alkaline phosphatase-linked donkey-anti-rabbit secondary antibody (for AvHPPD-03) or a peroxidase-linked goat anti-rabbit secondary antibody (for PAT).

In both the SDS-PAGE gel and the Western blot, the AvHPPD-03 protein was not detectable at 1 min or longer and no additional bands were visible at the longer incubation times. Within 0.5 minutes and at all subsequent incubation times, the PAT protein band was not visible on either the SDS-PAGE gel or Western blot and there were no bands visible at a lower molecular weight.

These results indicate that AvHPPD-03 and PAT are readily digested by pepsin under simulated gastric conditions.

*Simulated intestinal fluid (SIF)*

As for the SGF studies, AvHPPD-03 and PAT proteins derived from a microbial system were used. For AvHPPD-03 the test material was evaluated following incubation in SIF (U.S. Pharmacopeia, 1990) at 37° for 0, 1, 2, 5, 10, and 60 minutes and 2, 3, 6, 24 and 48 h. For PAT, incubation times at 37° C were 0.5, 2, 5, 10, 20, 30 and 60 min. SDS-PAGE and Western blot (using the same antibodies as for the SGF analyses) were then used to visualise the protein bands.

The SDS-PAGE gel of the AvHPPD-03 samples indicated that there was a significant reduction in the band intensity of AvHPPD-03 between 0 and 1 min followed by a gradual reduction in intensity to undetectable over the next few hours. The Western blot indicated a more rapid loss of band intensity with the AvHPPD-03 band being below the level of detection at 1 min and subsequently.

Because other proteins present in the SIF ran closely to the PAT protein, the SDS-PAGE gel did not provide unequivocal information about PAT. However, the Western blot clearly showed that PAT was not detectable after 10 min incubation in SIF.

The results support the conclusion that AvHPPD-03 and PAT will be readily digested under typical mammalian intestinal conditions.

#### **4.7.4 Stability to heat**

See Section 4.6.4.

### **4.8 Conclusion**

Soybean line SYHT0H2 expresses two novel proteins, AvHPPD-03 and PAT. Expression analyses of the two proteins showed that both were detected in the plant parts tested. The results were highly variable for both proteins but in general, it can be concluded that both AvHPPD-03 and PAT are present in highest concentration in V4 leaves and in lowest concentration (some 20 – 30 times lower than in the V4 leaves) in the seed. Analysis of a range of processed fractions obtained from the seed indicated that processing does not lead to concentration of the proteins and that in soy milk, tofu, protein concentrate and protein isolate, levels of the novel proteins are undetectable; levels in oil were not analysed.

A number of studies were used to confirm the identity and physicochemical properties of the plant-derived AvHPPD-03 and PAT proteins. These studies demonstrated that both proteins conform in size, amino acid sequence and activity to that expected, and do not exhibit any post-translational modification including glycosylation.

For both proteins, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens; digestibility studies suggest the proteins would be rapidly degraded in the stomach following ingestion; and thermolability studies indicate that both proteins are inactivated by heating. Taken together, the evidence indicates that neither AvHPPD-03 nor PAT are likely to be toxic or allergenic in humans.

## **5. Herbicide metabolites**

For GM foods derived from crops that are herbicide tolerant, there are two issues that require consideration. The first is dealt with in this safety assessment and involves assessment of any novel metabolites that are produced after the herbicide is applied, to determine whether these are present in the final food and whether their presence raises any toxicological concerns. In particular, the assessment considers whether appropriate health-based guidance values (i.e. Acceptable Daily Intake [ADI] or Acute Reference Dose [ARfD]) need to be established.

The second consideration, which is separate from the GM food approval process and therefore not included as part of this safety assessment, relates to the presence of herbicide residues on the food. Any food products (whether derived from GM or non-GM sources) sold in both Australia and New Zealand must not have residue levels greater than the relevant maximum residue limit (MRL). Where necessary, an MRL may have to be set.

As discussed in Section 4.2.1, tolerance is achieved by expression in the plant of a HPPD protein (from oat) that has lower sensitivity to the herbicide than the endogenous soybean HPPD. As a consequence, no metabolites are produced in SYHT0H2 plants sprayed with mesotrione that are not also produced in non-GM soybeans sprayed with the same herbicide.

Mesotrione is chemically derived from a natural phytotoxin obtained from the Australian native bottlebrush plant, *Callistemon citrinus* (Mitchell *et al.*, 2001; Comes, 2006). Details of the mode of action are given in Hawkes (2012). A number of crop species, particularly monocotyledons such as sorghum, corn, barley and oats exhibit natural tolerance to post-emergent spraying with mesotrione (Mitchell *et al.*, 2001; Abit and Al-Khatib, 2009; Abit *et al.*, 2009; Soltani *et al.*, 2010). The major metabolites of mesotrione (2-[4-(methylsulfonyl)-2-nitrobenzoyl]-1,3-cyclohexanedione) are the same in animal, plant and microbial systems (EPA, 2001; Batisson *et al.*, 2010) and comprise:

- 4-(methylsulfonyl)-2-nitrobenzoic acid (MNBA) and
- 2-amino-4-(methylsulfonyl)benzoic acid (AMBA) and its conjugates

In the environment, aerobic conditions favour formation of MNBA while suboxic conditions favour AMBA.

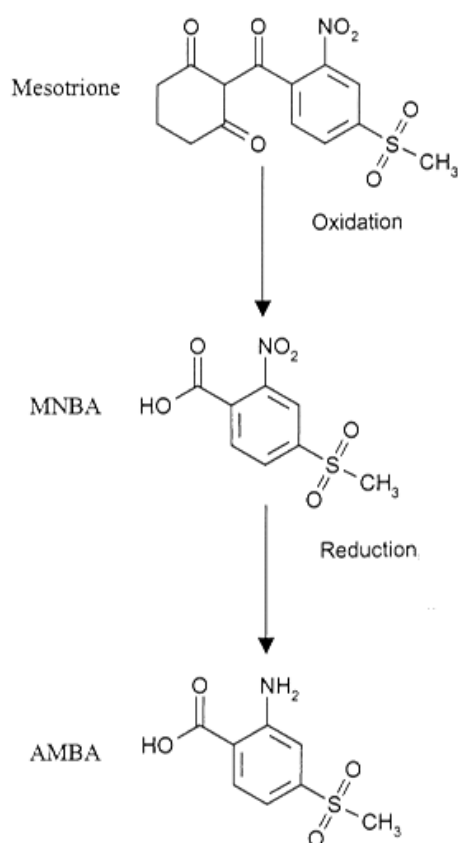


Figure 6: Production of MNBA and AMBA from mesotrione  
(Adapted from Alferness & Wiebe (2002))

As no novel herbicide metabolites are present in mesotrione-sprayed SYHT0H2, the existing health-based guidance value (i.e. ADI) for mesotrione is appropriate and relevant for assessing dietary risk with soybean SYHT0H2. In Australia, the ADI for mesotrione is 0.01 mg/kg bw/day<sup>5</sup>

In the case of PAT, the metabolic profiles resulting from the novel protein x herbicide interaction have been established through a significant history of use. There are no concerns that the spraying of soybean SYHT0H2 with glufosinate ammonium would result in the production of metabolites that are not also produced in crops sprayed with the same herbicides and already used in the food supply.

## **6. Compositional analysis**

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

The classic approach to the compositional analysis of GM food is a targeted one; rather than analysing every single constituent, which would be impractical, the aim is to analyse only those constituents most relevant to the safety of the food or that may have an impact on the whole diet. Important analytes therefore include the key nutrients, toxicants and anti-nutrients for the food in question. The key nutrients and anti-nutrients are those components in a particular food that may have a substantial impact in the overall diet. They may be major constituents (fats, proteins, carbohydrates or enzyme inhibitors as anti-nutrients) or minor constituents (minerals, vitamins). Key toxicants are those toxicologically significant compounds known to be inherently present in an organism, such as compounds whose toxic potency and level may be significant to health (eg solanine in potatoes).

### **Key components of soybean**

For soybean intended for human food use, the key components considered important for compositional analysis include the proximates (moisture, crude protein, fat, ash, fibre), amino acids, fatty acids, minerals, vitamins, isoflavones, phospholipids, sterols, saponins and the anti-nutrients phytic acid, trypsin inhibitors, stachyose, raffinose and lectins, (OECD, 2012). It is noted that the OECD recommendations for analysis of phospholipids, sterols and saponins are not emphasised in the previous version of the consensus document (OECD, 2001) and that the compositional studies done by the Applicant were based on this previous version.

Analyses for key components were done on seed and forage. In general, soybean is cultivated for the production of seed, which is used as a source of both human food and animal feed, and is only infrequently used as a forage crop for livestock. As there are no human food products derived from forage, only the results of the compositional analyses for seed and its processed fractions are presented in this report. The compositional analyses for forage focussed only on proximates, where the mean fat level for meso/glufos-sprayed SYHT0H2 was significantly lower than the mean for 'Jack'. All other analyte values were similar.

---

<sup>5</sup> ADIs are established by the Office of Chemical Safety within the Department of Health and Ageing - <http://www.health.gov.au/internet/main/publishing.nsf/Content/ocs-adi-list.htm>

**Study submitted:**

Launis, K. (2011) Compositional analysis of forage and seed from soybean event SYHT0H2 (*sic*) grown during 2010 in the USA. Unpublished Syngenta Report, November 3, 2011. TK0055213.

## 6.2 Study design and conduct

The test (PCR-verified SYHT0H2, seed of T<sub>6</sub> lineage), and control (PCR-verified 'Jack') lines were grown under similar conditions at eight field sites across North America<sup>6</sup> during the 2010 growing season. The sites were representative of where soybean is commercially grown. 'Jack' is the original transformed line and therefore represents the isogenic control line for the purposes of the comparative analyses. Six different non-GM soybean lines (PCR-verified) were also grown under the same conditions in order to generate a reference range for each analyte.

All lines were treated with conventional pesticides (termed 'untreated' for the purposes of this safety assessment). In addition, one SYHT0H2 line was treated with mesotrione (0.11 kg ai/ha) plus glufosinate (0.45 kg ai/ha) – henceforth referred to as meso/glufos - at the V3-V4 growth stage (Iowa State University 2009). All lines were planted in a randomised complete block design, with four replicated plots at each of the eight sites.

Seed and forage from soybean SYHT0H2 and 'Jack' were harvested from all replicated plots and analysed for composition. Forage was collected at the R6 plant growth stage, and seed was harvested at physiological maturity (R8 stage). Methods of composition analysis were based on internationally recognised procedures (e.g. those of the Association of Analytical Communities - AOAC), methods specified by the manufacturer of the equipment used for analysis, or other published methods.

A mean value was generated and used for each analyte comparison, and standard error and minimum and maximum values were also calculated for each analyte. The calculated means for seed are summarised in Tables 7 – 13. Analysis of Variance was used and paired contrasts were made across and within locations between untreated SYHT0H2 and 'Jack' and between meso/glufos-sprayed SYHT0H2 and 'Jack' using t-tests. In assessing the significance of any difference between the mean analyte value for soybean SYHT0H2 and 'Jack' a P-value of 0.05 was used. This means that approximately 5% of statistically significant differences are expected to occur due to chance alone.

The results for the treatments were compared to

- The reference range (min – max) compiled from the results of the six non-GM reference lines, in order to assess whether any differences were likely to be biologically meaningful.
- A combined literature range for each analyte, compiled from published literature<sup>7</sup>. Any mean value for a soybean SYHT0H2 analyte that fell within the combined literature range was considered to be within the normal variability of commercial soybean cultivars even if the mean value was statistically different from the 'Jack'

---

<sup>6</sup> The eight sites were: Richland, IA; York, NE; Fisk, MO; Stewardson and Carlyle, IL; Mebane, NC; Hamburg, PA; Rockville, IN

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

control. It is noted, however, that information in the published literature is limited and is unlikely to provide a broad reflection of the natural diversity that occurs within soybean. Therefore, even if means fall outside the published range, this is not necessarily a concern.

### 6.3 Seed composition

#### 6.3.1 Proximates and fibre

Results of the proximate and fibre analysis are shown in Table 7. The only significant difference occurred in a comparison of the acid detergent fibre (ADF) in 'Jack' and untreated SYHT0H2, where the latter had a lower mean value. However, this mean was well within both the reference range and the range reported in the literature.

Table 7: Mean percentage of proximates and fibre in seed from 'Jack' and SYHT0H2

Analyte	'Jack' (A)	SYHT0H2 untreated (B)	A vs B (P-value)	SYHT0H2 + meso/ glufos (C)	A vs C (P value)	Reference range	Combined literature range
Protein (%dw)	38.2	38.6	NS	38.4	NS	30.6 – 44.4	32 – 48.4
Fat (%dw)	20.7	20.5	NS	20.6	NS	15.8 – 25.0	8.1 – 24.7
Ash (%dw)	5.25	5.29	NS	5.17	NS	4.14 – 6.59	3.8 – 6.9
Moisture (%fw)	8.7	8.66	NS	8.8	NS	6.10 – 14.3	4.7 – 34.4
Carbohydrate (%dw) <sup>1</sup>	35.7	35.5	NS	35.9	NS	25.2 – 43.8	29.3 – 50.2
ADF <sup>2</sup> (%dw)	14.8	13.9 <sup>4</sup>	0.044	14.6	NS	8.20 – 20.6	7.81 – 26.6
NDF <sup>3</sup> (%dw)	16.7	16.0	NS	16.4	NS	11.2 – 21.9	8.53 – 23.9

<sup>1</sup> Carbohydrate calculated as 100% - (protein %dw + fat %dw + ash %dw)

<sup>2</sup> ADF = acid detergent fibre

<sup>3</sup> NDF = neutral detergent fibre

<sup>4</sup> Mauve shading represents SYHT0H2 means where a pairwise contrast t-test showed a significantly lower value than for the 'Jack' mean.

#### 6.3.2 Fatty acids

The levels of 22 fatty acids were measured. Of these, the following 13 were below the level of quantification (LOQ) in all replicates of the SHYT0H2 soybean and were therefore not statistically analysed: caprylic (8:0), capric (10:0), lauric (12:0), myristic (14:0), myristoleic (14:1), pentadecanoic (15:0), pentadecenoic (15:1), palmitoleic (16:1), heptadecenoic (17:1),  $\gamma$ -linolenic (18:3), eicosadienoic (20:2), eicosatrienoic (20:3), and arachidonic (20:4). For heptadecanoic acid (17:0) the levels were below the LOQ in >50% of replicates and were therefore not statistically analysed. Results for the remaining eight fatty acids are given in Table 8 and can be summarised as follows:

- There was no significant difference between 'Jack' and soybean SYHT0H2 in terms of the level of eicosenoic acid.
- For linoleic acid, the level in meso/glufos-sprayed SYHT0H2 was significantly lower than the level in 'Jack' but for untreated SYHT0H2 there was no difference from 'Jack'.

- The mean level of linoleic acid was significantly lower in both soybean SYHT0H2 treatments than in 'Jack'.
- The mean levels of palmitic, stearic, oleic, arachidic and behenic acids were significantly higher in both soybean SYHT0H2 treatments than in 'Jack'.
- All means for SYHT0H2 and 'Jack' were within both the reference range and the combined literature range.

*Table 8: Mean percentage composition, relative to total fat, of major fatty acids in seed from 'Jack' and SYHT0H2*

Fatty acid	'Jack' (A) % total	SYHT0H2 untreated (B) % total <sup>1</sup>	A vs B (P-value)	SYHT0H2 + meso + glufos (C) % total <sup>1</sup>	A vs C (P value)	Reference range % total	Combined literature range % total
Palmitic (16:0)	10.0	10.5	<0.001	10.6	<0.001	8.93 – 12.2	9.5 – 15.7
Stearic (18:0)	4.50	4.67	0.001	4.70	<0.001	3.75 – 6.32	2.59 – 5.88
Oleic (18:1)	23.0	24.3	0.004	24.6	0.001	18.1 – 35.2	14.3 – 45.6
Linoleic (18:2)	54.1	52.2	<0.001	51.8	<0.001	45.0 – 56.7	35.36 – 58.8
Linolenic (18:3)	7.51	7.35	<0.001	7.43	NS	5.30 – 10.1	3 – 12.52
Arachidic (20:0)	0.347	0.368	<0.001	0.373	<0.001	0.288 – 0.534	0.163 – 0.57
Eicosenoic (20:1)	0.181	0.183	NS	0.182	NS	0.153 – 0.286	<LOQ – 0.35.
Behenic (22:0)	0.357	0.372	0.001	0.373	0.001	0.304 – 0.498	0.277 – 0.595

<sup>1</sup> Mauve shading represents SYHT0H2 means where a pairwise contrast t-test showed a significantly lower value than for the 'Jack' mean, while orange shading represents SYHT0H2 means that were significantly higher than 'Jack'

### 6.3.3 Amino acids

Levels of 18 amino acids were measured. Since asparagine and glutamine are converted to aspartate and glutamate respectively during the analysis, levels for aspartate include both aspartate and asparagine, while glutamate levels include both glutamate and glutamine. Results of the analysis are given in Table 9.

- There was no significant difference between 'Jack' and soybean SYHT0H2 in terms of the levels of cysteine, glycine, isoleucine, tryptophan and valine
- For methionine, the level in meso/glufos-sprayed SYHT0H2 was significantly higher than the level in 'Jack' but for untreated SYHT0H2 there was no difference from 'Jack'.
- For alanine, serine, threonine and tyrosine the levels in untreated SYHT0H2 were significantly higher than the level in 'Jack' but for meso/glufos-sprayed SYHT0H2 there was no difference from 'Jack'.
- The mean levels of arginine, aspartate, glutamate, histidine, leucine, lysine, phenylalanine and proline were significantly higher in both soybean SYHT0H2 treatments than in 'Jack'.
- All means for SYHT0H2 and 'Jack' were within both the reference range and the combined literature range.
- It is noted that, despite over half of the amino acids being present at higher level in SYHT0H2 than in 'Jack' the mean protein levels in the two lines are no different (see Table 7).



Table 9: Mean percentage dry weight (dw), relative to total dry weight, of amino acids in seed from 'Jack' and SYHT0H2

Amino Acid	'Jack' (A) %dw	SYHT0H2 untreated (B) %dw <sup>1</sup>	A vs B (P-value)	SYHT0H2 + meso + glufos (C) % dw <sup>1</sup>	A vs C (P value)	Reference range %dw	Combined literature range %dw
Alanine	1.70	1.73	0.014	1.72	NS	1.44 – 1.85	1.43 – 2.10
Arginine	2.78	2.87	0.005	2.86	0.008	2.19 – 3.30	2.15 – 3.46
Aspartate	4.32	4.41	0.013	4.41	0.016	3.48 – 4.96	3.81 – 6.04
Cysteine	0.573	0.573	NS	0.582	NS	0.479 – 0.736	0.37 – 0.81
Glutamate	6.48	6.62	0.046	6.64	0.030	5.05 – 7.83	5.84 – 9.15
Glycine	1.66	1.69	NS	1.69	NS	1.38 – 1.85	1.41 – 2.00
Histidine	1.02	1.04	0.007	1.04	0.034	0.805 – 1.11	0.86 – 1.24
Isoleucine	1.84	1.86	NS	1.87	NS	1.49 – 2.07	1.49 – 2.08
Leucine	2.90	2.96	0.011	2.95	0.031	2.33 – 3.22	2.2 – 4.0
Lysine	2.37	2.46	<0.001	2.46	<0.001	1.98 – 2.74	2.19 – 3.32
Methionine	0.513	0.518	NS	5.31	0.011	0.422 – 0.619	0.39 – 0.68
Phenylalanine	1.94	1.99	0.01	1.98	0.017	1.55 – 2.17	1.6 – 2.44
Proline	1.90	1.96	0.002	1.95	0.008	1.54 – 2.26	1.63 – 2.28
Serine	1.90	1.94	0.048	1.93	NS	1.53 – 2.16	1.11 – 2.48
Threonine	1.54	1.57	0.021	1.56	NS	1.29 – 1.67	1.14 – 1.89
Tryptophan	0.569	0.572	NS	0.573	NS	0.488 – 0.620	0.30 – 0.67
Tyrosine	1.49	1.51	0.035	1.51	NS	1.23 – 1.64	0.79 – 1.61
Valine	1.87	1.90	NS	1.90	NS	1.50 – 2.06	1.5 – 2.44

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

### 6.3.4 Isoflavones

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

The Applicants used an extraction method (Seo and Morr, 1984) that gives total isoflavones (ie aglycons + conjugates), also referred to as aglycon equivalents. The levels are given in Table 10 and show that the mean levels of total daidzein, total genistein and total glycitein in soybean SYHT02 were not significantly different from those in 'Jack'.

Table 10: Mean weight ( $\mu\text{g/g}$  dry weight expressed as aglycon equivalents) of isoflavones in SYHT0H2 and 'Jack' seed

Isoflavone	'Jack' (A) ( $\mu\text{g/g}$ dw)	SYHT0H2 untreated (B) ( $\mu\text{g/g}$ dw)	A vs B (P-value)	SYHT0H2 + meso + glufos (C) ( $\mu\text{g/g}$ dw)	Overall treat effect (P-value) (A vs C)	Reference range ( $\mu\text{g/g}$ dw)	Combined literature range ( $\mu\text{g/g}$ dw)
Total Daidzein (aglycon equivalents)	375	391	NS	396	NS	229 - 1230	25 - 2453
Total Genistein (aglycon equivalents)	556	569	NS	548	NS	165 - 1240	28 - 2837
Total Glycitein (aglycon equivalents)	196	181	NS	193	NS	58.8 - 265	15 - 349

### 6.3.5 Anti-nutrients

Levels of key anti-nutrients are given in Table 11. Overall, there were no significant differences between the SYHT0H2 means and the control means for any of the analytes. All lectin means were within both the reference range and the combined literature range.

Table 11: Mean levels of anti-nutrients in SYHT0H2 and 'Jack' seed.

Anti-nutrient	'Jack' (A)	SYHT0H2 untreated (B)	A vs B (P-value)	SYHT0H2 + meso + glufos (C)	A vs C (P value)	Reference range	Combined literature range
Lectin (Haemagglutinat Units/mg protein dw)	25.8	26.1	NS	26.6	NS	4.19 - 61.3	0.1 - 323
Phytic acid (%dw)	1.41	1.38	NS	1.40	NS	0.766 - 2.21	0.41 - 2.74
Raffinose (%dw)	0.801	0.816	NS	0.805	NS	0.607 - 1.58	0.21 - 1.62
Stachyose (%dw)	3.72	3.76	NS	3.72	NS	3.15 - 5.13	1.21 - 6.1
Trypsin inhibitor (trypsin inhibitor units/mg)	34.4	35.9	NS	33.4	NS	18.9 - 68.3	18.14 - 118.68

### 6.3.6 Minerals

Levels of five minerals were measured. The means for these are given in Table 12 and show:

- The mean iron and potassium levels in SYHT0H2 were significantly lower than the means for 'Jack'.
- The mean potassium levels for both lines were outside (lower than) the literature range but within the reference range.
- The mean iron levels for both lines were within both the reference range and combined literature range.

Table 12: Mean values for mineral levels in seed from 'Jack' and SYHT0H2.

Mineral	'Jack' (A) (mg/100g dw)	SYHT0H2 untreated (B) (mg/100g dw) <sup>1</sup>	A vs B (P-value)	SYHT0H2 + meso + glufos (C) (mg/100g dw) <sup>1</sup>	A vs C (P value)	Reference range (mg/100g dw)	Combined literature range (mg/100g dw)
Calcium	299.0	306.2	NS	292.4	NS	205.0 – 386.0	116 - 510
Iron	8.34	8.05	0.027	7.99	0.01	4.80 – 11.0	3.73 – 10.95
Magnesium	239.1	243.3	NS	241.4	NS	182.0 – 309.0	219 - 312
Phosphorus	611.7	614.1	NS	618.2	NS	420.0 – 857.0	506 - 935
Potassium	1774.7	1726.1	0.002	1729.1	0.004	1380.0 – 2470.0	1868 - 2510

<sup>1</sup> Mauve shading represents SYHT0H2 means where a pairwise contrast t-test showed a significantly lower value than for the 'Jack' mean.

### 6.3.7 Vitamins (including Vitamin E compounds)

Levels of 13 vitamins were measured. Those of Vitamin A ( $\beta$ -carotene),  $\beta$ -tocopherol and  $\alpha$ -,  $\beta$ -,  $\gamma$ - and  $\delta$ -tocotrienol were generally below the LOQ and were not statistically analysed.

The means for the remaining seven vitamins are given in Table 13. Overall:

- There were no significant differences between the SYHT0H2 means and the control means for Vitamins B<sub>1</sub>, B<sub>2</sub> or B<sub>9</sub>.
- For Vitamin K<sub>1</sub> the mean for the meso/glufos-sprayed SYHT0H2 was significantly lower than the mean for 'Jack' but there was no difference between the mean levels in the untreated SYHT0H2 and 'Jack'.
- For  $\alpha$ -tocopherol, the mean for the untreated SYHT0H2 was significantly lower than the mean for 'Jack' but there was no difference between the mean levels in the meso/glufos-sprayed SYHT0H2 and 'Jack'.
- For  $\gamma$ - and  $\delta$ - tocopherol, the means for SYHT0H2 were higher than for 'Jack'. This increase is not unexpected since increased expression of HPPD has been shown to lead to increased synthesis of Vitamin E compounds (Falk *et al.*, 2003 and see also Fig. 5).and  $\gamma$ - and  $\delta$ - tocopherol are the major Vitamin E compounds produced in soybean seed (Kim *et al.*, 2012).
- Means for Vitamins B<sub>9</sub> and K<sub>1</sub> and  $\gamma$ - and  $\delta$ - tocopherol for both lines were within both the reference range and the combined literature range (where a range has been reported).
- For both soybean lines, means for Vitamins B<sub>1</sub> and B<sub>2</sub> were outside (higher than) the literature range but were within the reference range.

Table 13: Mean weight (mg/g dry weight) of vitamins in seed from 'Jack' and SYHT0H2

Vitamin	'Jack' (A)	SYHT0H2 untreated (B)	A vs B (P-value)	SYHT0H2 + meso + glufos (C)	A vs C (P value)	Reference range	Combined literature range
Vitamin B <sub>1</sub> (thiamin) (mg/100 g dw)	0.535	0.515	NS	0.535	NS	0.253 – 1.02	0.101 – 0.254
Vitamin B <sub>2</sub> (riboflavin) (mg/100 g dw)	0.381	0.384	NS	0.394	NS	0.270 – 0.532	0.190 – 0.321
Vitamin B <sub>9</sub> (folic acid) (mg/100 g dw)	0.415	0.440	NS	0.427	NS	0.224 – 0.680	0.238 – 0.471

Vitamin	'Jack' (A)	SYHT0H2 untreated (B)	A vs B (P-value)	SYHT0H2 + meso + glufos (C)	A vs C (P value)	Reference range	Combined literature range
Vitamin K <sub>1</sub> (phytonadione) (mg/100 g dw)	0.462	0.411	NS	0.391	0.019	0.106 – 0.886	Not reported
Vitamin E (α-tocopherol) (mg/g dw)	0.0258	0.0228	0.019	0.0241	NS	0.0115 – 0.0771	0.0011 – 0.0617
γ-tocopherol (mg/g dw)	0.201	0.226	<0.001	0.214	0.004	0.127 – 0.236	Not reported
δ-tocopherol (mg/g dw)	0.0611	0.0789	<0.001	0.076	<0.001	0.032 – 0.112	Not reported

<sup>1</sup> Mauve shading represents SYHT0H2 means where a pairwise contrast t-test showed a significantly lower value than for the 'Jack' mean, while orange shading represents SYHT0H2 means that were significantly higher than 'Jack'

### 6.3.8 Summary of analysis of key components

A total of 46 analytes were analysed. For 20 of the analytes there were no significant differences between the levels found in seed of soybean SYHT0H2 and 'Jack'. A summary of those showing statistically significant differences in the analyte is provided in Table 14. These statistically significant differences do not raise safety concerns, given that:

- there are no trends in the results,
- the magnitude of most differences were less than 10%; i.e were quite small;
- for eight of the analytes, a significant difference from the control was noted in only one of the SYHT0H2 treatments.;
- for all the analytes where a significant difference was identified, the soybean SYHT0H2 means fall within both the reference range and (where it exists) the literature range.

Table 14: Summary of analyte means found in seed of SYHT0H2 treatments that are significantly (adj.  $P < 0.05$ ) different from those found in seed of the control line 'Jack'

Analyte	'Jack'	SYHT0H2 untreated	SYHT0H2 + meso + glufos	SYHT0H2 within Reference range?	SYHT0H2 within Combined literature range?
ADF <sup>2</sup> (%dw)	14.8	13.9 <sup>4</sup>	14.6	yes	yes
Palmitic (16:0) (% total fat)	10.0	10.5	10.6	yes	yes
Stearic (18:0) (%total fat)	4.50	4.67	4.70	yes	yes
Oleic (18:1) (%total fat)	23.0	24.3	24.6	yes	yes
Linoleic (18:2) (% total fat)	54.1	52.2	51.8	yes	yes
Linolenic (18:3) (% total fat)	7.51	7.35	7.43	yes	yes
Arachidic (20:0) (% total fat)	0.347	0.368	0.373	yes	yes
Behenic (22:0) (% total fat)	0.357	0.372	0.373	yes	yes
Alanine (%dw)	1.70	1.73	1.72	yes	yes
Arginine (%dw)	2.78	2.87	2.86	yes	yes

Analyte	'Jack'	SYHT0H2 untreated	SYHT0H2 + meso + glufos	SYHT0H2 within Reference range?	SYHT0H2 within Combined literature range?
Aspartate (%dw)	4.32	4.41	4.41	yes	yes
Glutamate (%dw)	6.48	6.62	6.64	yes	yes
Histidine (%dw)	1.02	1.04	1.04	yes	yes
Leucine (%dw)	2.90	2.96	2.95	yes	yes
Lysine (%dw)	2.37	2.46	2.46	yes	yes
Methionine (%dw)	0.513	0.518	5.31	yes	yes
Phenylalanine (%dw)	1.94	1.99	1.98	yes	yes
Proline (%dw)	1.90	1.96	1.95	yes	yes
Serine (%dw)	1.90	1.94	1.93	yes	yes
Threonine (%dw)	1.54	1.57	1.56	yes	yes
Tyrosine (%dw)	1.49	1.51	1.51	yes	yes
Iron (mg/100g dw)	8.34	8.05	7.99	yes	yes
Potassium (mg/100 g dw)	1774.7	1726.1	1729.1	yes	No (but 'Jack' was also outside)
Vitamin K <sub>1</sub> (phytonadione) (mg/100 g dw)	0.462	0.411	0.391	yes	Not reported
γ-tocopherol (mg/g dw)	0.201	0.226	0.214	yes	Not reported
δ-tocopherol (mg/g dw)	0.0611	0.0789	0.076	yes	Not reported

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

## 6.4 Conclusion

The compositional analyses do not indicate any differences of biological significance between seed from soybean SYHT0H2 and the non-GM control 'Jack'. Statistically significant differences were noted in a number of constituents. However the differences were typically small and all mean values were within both the reference range obtained for non-GM reference varieties grown at the same time and (where it exists) the literature range.

Any observed differences are therefore considered to represent the natural variability that exists within soybean. The spraying of soybean line SYHT0H2 with mesotrione and glyphosate did not have a significant effect on seed composition.

## 7. Nutritional impact

In assessing the safety of a GM food, a key factor is the need to establish that the food is nutritionally adequate and will support typical growth and well-being. In most cases, this can be achieved through an understanding of the genetic modification and its consequences, together with an extensive compositional analysis of the food.

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

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

Soybean 44406 is the result of a simple genetic modification to confer herbicide tolerance with no intention to significantly alter nutritional parameters in the food. In addition, the extensive compositional analyses of seed that have been undertaken to demonstrate the nutritional adequacy of line SYHT0H2, indicate it is equivalent in composition to conventional soybean cultivars. The introduction of soybean line SYHT0H2 into the food supply is therefore expected to have little nutritional impact.

The Applicant submitted a feeding study in broiler chickens the results of which support the conclusion of the compositional analysis.

#### **Study submitted**

Tovaas, L. (2012) The effect of diets containing soybean meal from SYHT0H2, nontransgenic control and a commercial variety on broiler growth performance and carcass parameters. Unpublished Syngenta Report, July 24, 2012. TK0036575.

#### **References<sup>8</sup>**

Abit, M.J.M. and Al-Khatib, J. (2009) Absorption, translocation, and metabolism of mesotrione in grain sorghum. *Weed Science* 57:563-566.

Abit, M.J.M., Al-Khatib, J., Regehr, D.L., Tuinstra, M.R., Claassen, M.M., Geier, P.W., Stahlman, P.W., Gordon, B.W. and Currie, R.S. (2009) Differential response of grain sorghum hybrids to foliar-applied mesotrione. *Weed Technology* 23:28-33.

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

Agostino, M. (2013) Introduction to the BLAST suite and BLASTN. In: *Practical Bioinformatics*. Chapter 3. Garland Science, Taylor & Francis Group, New York, pp. 47-71.

Alferness, P. and Wiebe, L. (2002) Determination of Mesotrione residues and metabolites in crops, soil, and water by liquid chromatography with fluorescence detection. *Journal of Agricultural and Food Chemistry* 50:3926-3934.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25(17):3389-3402.

Andrews, C.J., Warner, S.A.J. and Hawkes, T.R. (2007) Methods for production of plants resistant to HPPD herbicides. (Patent No. US 7312379 B2): <http://www.patentlens.net>.

---

<sup>8</sup> All website references were current as at 5 July 2013

- Astwood, J.D. and Fuchs, R.L. (1996) Allergenicity of foods derived from transgenic plants. *Monographs in Allergy* 32 (Highlights in Food Allergy):105-120.
- Awata, H., Endo, F. and Matsuda, I. (1994) Structure of the human 4-hydroxyphenylpyruvic acid dioxygenase gene (HPD). *Genomics* 23(3):534-539.
- Batisson, I., Sancelme, M., Mallet, C. and Besse-Hoggan. (2010) Fate and environmental impact of the recently marketed herbicide Mesotrione: coupling biological and chemical studies for a global overview. In: Méndez-Vilas, A. eds. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*. available at <http://www.formatex.info/microbiology2/287-294.pdf>.
- Baxevanis, A.D. (2005) Assessing Pairwise Sequence Similarity: BLAST and FASTA. In: Baxevanis, A.D. and Ouellette, B.F.F. eds. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. Chapter 11. John Wiley & Sons, Inc., pp. 295-324.
- Berman, K.H., Harrigan, G.G., Riordan, S.G., Nemeth, M.A., Hanson, C., Smith, M., Sorbet, R., Zhu, E. and Ridley, W.P. (2009) Compositions of seed, forage, and processed fractions from insect-protected soybean MON 87701 are equivalent to those of conventional soybean. *Journal of Agricultural and Food Chemistry* 57(23):11360-11369.
- Berman, K.H., Harrigan, G.G., Riordan, S.G., Nemeth, M.A., Hanson, C., Smith, M., Sorbet, R., Zhu, E. and Ridley, W.P. (2010) Compositions of forage and seed from second-generation glyphosate-tolerant soybean MON 89788 and insect-protected soybean MON 87701 from Brazil are equivalent to those of conventional soybean (*Glycine max*). *Journal of Agricultural and Food Chemistry* :
- Bilyeu, K.D., Zeng, P., Coello, P., Zhang, Z.J., Krishnan, H.B., Bailey, A., Beuselinck, P.R. and Polacco, J.C. (2008) Quantitative conversion of phytate to inorganic phosphorus in soybean seeds expressing a bacterial phytase. *Plant Physiology* 146(2):468-477.
- Bradbury, J.F. (1986) *Guide to plant pathogenic bacteria*. Cambridge News Ltd, Aberystwyth, p190-197.
- Brookes, G. and Barfoot, P. (2009) *GM crops: global socio-economic and environmental impacts 1996-2007*. PG Economics Ltd. <http://www.pgeconomics.co.uk/pdf/2009globalimpactstudy.pdf>.
- Chang, T.M., Chuang, Y.C., Su, J.H. and Chang, M.C. (1997) Cloning and sequence analysis of a novel hemolysin gene (vIIY) from *Vibrio vulnificus*. *Applied and Environmental Microbiology* 63(10):3851-3857.
- Ciccocioppo, R., Di Sabatino, A. and Corazza, G.R. (2005) The immune recognition of gluten in coeliac disease. *Clinical and Experimental Immunology* 140:408-416.
- Codex (2003) *Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants*. Report No. CAC/GL 45-2003, Codex Alimentarius. [http://www.codexalimentarius.net/web/standard\\_list.do?lang=en](http://www.codexalimentarius.net/web/standard_list.do?lang=en).
- Coeliac Research Fund. (2008) Position statement on: the consumption of oats by individuals with coeliac disease. Coeliac Australia, <http://www.coeliac.org.au/>.
- Comes, D. (2006) Callisto: A very successful maize herbicide inspired by allelochemistry. In: Maize Association of Australia, 6th Triennial Conference 2006, available at [http://www.maizeaustralia.com.au/general\\_info.html](http://www.maizeaustralia.com.au/general_info.html).



- Deblaere, R., Reynaerts, A., Höfte, H., Hernalsteens, J.-P., Leemans, J. and van Montagu, M. (1987) Vectors for cloning in plant cells. *Methods in Enzymology* 153:277-292.
- Delaney, B., Astwood, J.D., Cunny, H., Eichen Conn, R., Herouet-Guicheney, C., MacIntosh, S., Meyer, L.S., Privalle, L.S., Gao, Y., Mattsson, J., Levine, M. and ILSI. (2008) Evaluation of protein safety in the context of agricultural biotechnology. *Food and Chemical Toxicology* 46:S71-S97.
- Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M. (1982) Nopaline synthase: transcript mapping and DNA sequence. *Journal of Molecular and Applied Genetics* 1(6):561-573.
- Edmunds, A.J.F. and Morris, J.A. (2012) Triketones. In: Krämer, W., Schirmer, U., Jeschke, P., and Witschel, M. eds. *Section 4.3 in Chapter 4, Herbicides with Bleaching Properties In: Modern Crop Protection Compounds, Volumes 1-3*. 2nd ed, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, pp. 235-262.
- EFSA. (2008) Safety and nutritional assessment of GM plants and derived food and feed: The role of animal feeding trials. *Food and Chemical Toxicology* 46:S1-S70, doi:10.1016/j.fct.2008.02.008.
- EPA (2001) *Mesotrione: Issues to be presented to the Health Effects Division (HED) Metabolism Assessment Review Committee (MARC) 10-April-01*. Report of the United States Environmental Protection Agency, Washington D.C.
- Falk, J., Andersen, G., Kernebeck, B. and Krupinska, K. (2003) Constitutive overexpression of barley 4-hydroxyphenylpyruvate dioxygenase in tobacco results in elevation of the vitamin E content in seeds but not in leaves. *FEBS Letters* 540:35-40.
- FAO/WHO (2001) *Evaluation of allergenicity of genetically modified foods*. Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology, January 22 - 25, 2001, Rome.
- FAOSTAT. (2012) Online database of the Food and Agriculture Organization of the United Nations. <http://faostat.fao.org/>.
- FSANZ (2011) *Application A1051 - Food derived from Herbicide-tolerant Soybean Event FG72*. Report prepared by Food Standards Australia New Zealand. <http://www.foodstandards.gov.au/foodstandards/applications/applicationa1051food4902.cfm>.
- Gallie, D.R. (2002) The 5'-leader of tobacco mosaic virus promotes translation through enhanced recruitment of eIF4F. *Nucleic Acids Research* 30(15):3401-3411.
- Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M.A. (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(8):3257-3273.
- Galvez-Valdivieso, G., Pineda, M. and Aguilar, M. (2010) Functional characterization and expression analysis of p-hydroxyphenylpyruvate dioxygenase from the green alga *Chlamydomonas reinhardtii* (Chlorophyta). *Journal of Phycology* 46(2):297-308.
- Goodman, R.E. (2006) Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. *Molecular Nutrition and Food Research* 50:655-660.



Grey, D. (2006) Growing soybeans in northern Victoria. State of Victoria, Department of Primary Industries,  
[http://aof.clients.squid.net/data/assets/pdf\\_file/0008/7667/Growing\\_Soybeans\\_in\\_Northern\\_Victoria.pdf](http://aof.clients.squid.net/data/assets/pdf_file/0008/7667/Growing_Soybeans_in_Northern_Victoria.pdf).

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

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

Harrigan, G.G., Ridley, W.P., Riordan, S., Nemeth, M.A., Sorbet, R., Trujillo, W.A., Breeze, M. and Schneider, R.W. (2007) Chemical composition of glyphosate-tolerant soybean 40-3-2 grown in Europe remains equivalent with that of conventional soybean (*Glycine max* L.). *Journal of Agricultural and Food Chemistry* 55(15):6160-6168.

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

Hawkes, T.R. (2012) Hydroxyphenylpyruvate Dioxygenase (HPPD): The Herbicide Target. In: Krämer, W., Schirmer, U., Jeschke, P., and Witschel, M. eds. *Section 4.2 in Chapter 4, Herbicides with Bleaching Properties In: Modern Crop Protection Compounds, Volumes 1-3*. 2nd ed, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, pp. 225-235.

Hawkes, T.R., Langford, M.P., Viner, R.C., Vernooij, B.T.M. and Dale, R. (2012) Mutant hydroxyphenylpyruvate dioxygenase polypeptides and methods of use. (Patent No. US 8269068 B2): <http://www.patentlens.net>.

Health Canada (2007) *Celiac disease and the safety of oats*. Health Canada, Ottawa.  
[http://www.hc-sc.gc.ca/fn-an/alt\\_formats/hpfb-dgpsa/pdf/securit/oats\\_cd-avoine-eng.pdf](http://www.hc-sc.gc.ca/fn-an/alt_formats/hpfb-dgpsa/pdf/securit/oats_cd-avoine-eng.pdf).

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

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

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

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

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

Hornsey, I.S. (2003) *A history of beer and brewing*. The Royal Society of Chemistry, Cambridge.

Hwang, Y.-J., Dawson, J., Sigareva, M. and Que, Q. (2008) Transformation of immature soybean seeds through organogenesis. (WO 2008/112044 A1):  
<http://www.patentlens.net/daisy/patentlens/patentlens.html>.

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

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

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

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

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

Kilmartin, C., Lynch, S., Abuzakouk, M., Wieser, H. and Feighery, C. (2003) Avenin fails to induce a Th1 response in coeliac tissue following in vitro culture. *Gut* 52:47-52.

Kim, E.-H., Ro, H.-M., Kim, S.-L. and Chung, I.-M. (2012) Analysis of isoflavone, phenolic, soyasapogenol, and tocopherol compounds in soybean [*Glycine max* (L.) Merrill] germplasms of different seed weights and origins. *Journal of Agricultural and Food Chemistry* 60:6045-6055.

Kimber, I., Kerkvliet, N.I., Taylor, S.L., Astwood, J.D., Sarlo, K. and Dearman, R.J. (1999) Toxicology of protein allergenicity: prediction and characterization. *Toxicological Sciences* 48(2):157-162.

Krell, R. (1996) *Value-Added Products from Beekeeping*. Chapter 3: Pollen, FAO Agricultural Services Bulletin No. 124, Food and Agriculture Organization of the United Nations, available online at <http://www.fao.org/docrep/w0076e/w0076e00.htm#con>.

Kützner, H.J. (1981) The Family Streptomycetaceae. In: Starr, M.P., Stolp, H., Trüper, H.G., Ballows, A., and Schlegel, H.G. eds. *The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*. Springer Verlag, Berlin, pp. 2028-2090.

Lackey, R. (2011) Distillers grains and soybean hulls - a complementary combination for cattle performance. *OMAFRA Virtual Beef* 10(31):2-6 available online at  
<http://www.omafra.gov.on.ca/english/livestock/beef/news.html>.

Lindstedt, S. and Odelhög, B. (1987) 4-hydroxyphenylpyruvate in *Pseudomonas*. *Methods in Enzymology* 142:143-148.

Lundry, D.R., Ridley, W.P., Meyer, J.J., Riordan, S.G., Nemeth, M.A., Trujillo, W.A., Breeze, M.L. and Sorbet, R. (2008) Composition of grain, forage, and processed fractions from second-generation glyphosate-tolerant soybean, MON 89788, is equivalent to that of

conventional soybean (*Glycine max* L.). *Journal of Agricultural and Food Chemistry* 56(12):4611-4622.

Maiti, I.B., Gowda, S., Kiernan, J., Ghosh, S.K. and Shepherd, R.J. (1997) Promoter/leader deletion analysis and plant expression vectors with the figwort mosaic virus (FMV) full length transcript (FLt) promoter containing single or double enhancer domains. *Transgenic Research* 6:143-156.

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

McCann, M.C., Liu, K., Trujillo, W.A. and Dobert, R.C. (2005) Glyphosate-tolerant soybeans remain compositionally equivalent to conventional soybeans (*Glycine max* L.) during three years of field testing. *Journal of Agricultural and Food Chemistry* 53:5331-5335.

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

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

Mitchell, G., Bartlett, D.W., Fraser, T.E.M., Hawkes, T.R., Holt, D.C., Townson, J.K. and Wichert, R.A. (2001) Mesotrione: a new selective herbicide for use in maize. *Pest Management Science* 57:120-128.

Moran, G.R. (2005) 4-hydroxyphenylpyruvate dioxygenase. *Archives of Biochemistry and Biophysics* 433:117-128.

NCBI (2010) *National Center for Biotechnology Information. Entrez Nucleotide Database, including GenBank*. National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>.

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

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

OECD (2003) *Considerations for the Safety Assessment of Animal Feedstuffs Derived from Genetically Modified Plants*. Series on the Safety of Novel Foods and Feeds, No. 9. Organisation for Economic Cooperation and Development, Paris. [http://www.oecd.org/olis/2003doc.nsf/LinkTo/NT0000426A/\\$FILE/JT00147696.PDF](http://www.oecd.org/olis/2003doc.nsf/LinkTo/NT0000426A/$FILE/JT00147696.PDF).

OECD (2012) *Revised consensus document on compositional considerations for new varieties of soybean [*Glycine max* (L.) Merr.]: Key food and feed nutrients, antinutrients, toxicants and allergens*. Report No. ENV/JM/MONO(2012)24, Organisation for Economic Co-operation and Development. <http://www.oecd.org/env/chemicalsafetyandbiosafety/biosafety-biotrack/consensusdocumentsfortheworkonthesafetyofnovelfoodsandfeedsplants.htm>.

Ow, D.W., Jacobs, J.D. and Howell, S.H. (1987) Functional regions of the cauliflower mosaic virus 35S RNA promoter determined by use of the firefly luciferase gene as a reporter of promoter activity. *Proceedings of the National Academy of Sciences* 84:4870-4874.

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

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

Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences* 85(8):2444-2448.

Polevoda, B. and Sherman, F. (2000) N $\alpha$ -terminal Acetylation of Eukaryotic Proteins. *Journal of Biological Chemistry* 275(47):36479-36482.

Que, Q., Dawson, J. and Sigareva, M. (2008) Transformation of immature soybean seeds through organogenesis. (WO 2008/112267 A2):  
<http://www.patentlens.net/daisy/patentlens/patentlens.html>.

Rashid, M., Butzner, D., Burrows, V., Zarkadas, M., Case, S., Molloy, M., Warren, R., Pulido, O. and Switzer, C. (2007) Consumption of pure oats by individuals with celiac disease: A position statement by the Canadian Celiac Association. *Canadian Journal of Gastroenterology* 21(10):649-651.

Rice, P., Longden, I. and Bleasby, A. (2000) EMBOSS: The European Molecular Biology Software Suite. *Trends in Genetics* 16(6):276-277.

Riddles, P.W., Blakely, R.L. and Zerner, B. (1983) Reassessment of Ellman's reagent. *Methods in Enzymology* 91:49-60.

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

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

Seo, A. and Morr, C.V. (1984) Improved high-performance liquid chromatographic analysis of phenolic acids and isoflavonoids from soybean protein products. *Journal of Agricultural and Food Chemistry* 32(3):530-533.

Shurtleff, W. and Aoyagi, A. (2007) History of soybean crushing: Soy oil and soybean meal. In: *History of soybeans and soyfoods: 1100 B.C. to the 1980s*. SoyInfo Center, available online at [http://www.soyinfocenter.com/HSS/soybean\\_crushing1.php](http://www.soyinfocenter.com/HSS/soybean_crushing1.php), Lafayette, California.

Soltani, N., Cowan, T., Shropshire, C. and Sikkema, P.H. (2010) Tolerance of spring cereals to mesotrione. In: 2010 North Central Weed Science Society Conference, Lexington KY, 13 - 16 December 2010, available at [www.weedpro75.com/public\\_trial\\_files.php?trialID=23](http://www.weedpro75.com/public_trial_files.php?trialID=23).  
[www.weedpro75.com/public\\_trial\\_files.php?trialID=23](http://www.weedpro75.com/public_trial_files.php?trialID=23).

Stavolone, L., Kononova, M., Pauli, S., Ragozzino, A., de Haan, P., Milligan, S., Lawton, K. and Hohn, T. (2003) Cestrum yellow leaf curling virus (CmYLCV) promoter: a new strong constitutive promoter for heterologous gene expression in a wide variety of crops. *Plant Molecular Biology* 53:703-713.

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

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

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

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

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

Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.-J., Glatt, C.M., Hadfield, N., Hatzos, C., Hefle, S.L., Heylings, J.R., Goodman, R.E., Henry, B., Herouet, C., Holsapple, M., Ladics, G.S., Landry, T.D., MacIntosh, S.C., Rice, E.A., Privalle, L.S., Steiner, H.Y., Teshima, R., Van Ree, R., Woolhiser, M. and Zawodny, J. (2004) A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regulatory Toxicology and Pharmacology* 39:87-98.

Thomas, K., Bannon, G., Hefle, S., Herouet, C., Holsapple, M., Ladics, G., MacIntosh, S. and Privalle, L. (2005) In silico methods for evaluating human allergenicity to novel proteins: International Bioinformatics Workshop Meeting Report February 23 - 24, 2005. *Toxicological Sciences* 88:307-310.

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

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

Tzfira, T., Li, J., Lacroix, B. and Cytosky, V. (2004) *Agrobacterium* T-DNA integration: molecules and models. *Trends in Genetics* 20(8):375-383.

U.S.Pharmacopeia. (1990) *United States Pharmacopeia, Volume 22*. United States Pharmacopeia Convention, Inc, Rockville, MD, p1788-1789.

U.S.Pharmacopeia. (2000) *The United States Pharmacopeia 24. The National Formulary 19th ed.* United States Pharmacopeia Convention, Inc, Rockville, MD.



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

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

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

Webster, F.H. (2011) Oat utilization: Past, present and future. In: *Oats: Chemistry and technology*. 2 ed, Chapter 17. AACC International, pp. 347-361.

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

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

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