

Supporting document 1

Safety assessment – Application A1089

Food derived from Herbicide-tolerant Canola Line DP-073496-4

Summary and conclusions

Background

Pioneer Hi-Bred Australia Pty Ltd has developed genetically modified (GM) canola line DP-073496-4 (OECD Unique identifier DP-073496-4) tolerant to the broad spectrum herbicide, glyphosate. The enzyme GAT4621 used to confer herbicide tolerance in this product catalyses the acetylation of glyphosate to produce *N*-acetyl glyphosate, which is herbicidally inactive. The GAT4621 enzyme is not new to the food supply.

In conducting a safety assessment of food derived from DP-073496-4 canola (herein referred to as canola line 73496), a number of criteria have been addressed including: characterisation of the transferred genes including their origin, function and stability in the canola genome; the nature of the introduced protein and its potential to be either allergenic or toxic in humans; compositional analyses and any resultant changes in the whole food. This approach evaluates the intended and any unintended changes in the plant.

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

- any potential impact on the environment arising from the release of GM food crops into the environment, and
- the safety of animal feed, or food produced from animals that consume GM feed.

Food derived from the non-GM (conventional) plant with an accepted history of safe use is used as the benchmark for the comparative analysis.

The Applicant anticipates that canola 73496 will be commercially cultivated in major canola-producing regions, including North America and Australia. Food products derived from canola line 73496 would therefore be expected to enter the Australian and New Zealand food supply via domestic production in Australia, and via imported products.

History of Use

Canola is rapeseed (*Brassica napus*, *B. rapa* or *B. juncea*) which has been conventionally bred to contain less than 2% erucic acid and less than 30 micromoles of glucosinolates per gram of seed solids, by definition. Rapeseed is the second largest oilseed crop in the world behind soybean, although annual production is around 25% of that of soybean.

Canola seeds are processed into two major products, oil and meal. The oil is the major product for human consumption, being used directly for cooking and as an ingredient in a variety of manufactured food products. Canola oil is the third largest source of vegetable oil in the world after soybean oil and palm oil. Whole canola seeds are being used increasingly in products such as breads.

Molecular Characterisation

The novel gene in canola line 73496, *gat4621*, resulted from a multigene shuffling process to derive an enzyme with optimised catalytic activity on glyphosate. The original *gat* gene sequences were isolated from three strains of *Bacillus licheniformis*. The relevant genes were identified using a mass spectrometry method that detected the product of interest, *N*-acetylglyphosate.

Comprehensive molecular analyses of canola line 73496 indicated that one intact copy of the *gat4621* gene expression cassette is present at a single insertion site in the plant. Plasmid backbone analysis shows no extraneous sequences derived from the plasmid were incorporated into the canola genome. The introduced genetic elements are stably inherited from one generation to the next. Bioinformatic analyses of open reading frames within the insert and junction regions in canola 73496 demonstrate no putative polypeptides with relevant homology to proteins that are known to be toxic, allergenic or have other biologically adverse properties.

Characterisation of Novel Proteins

The newly expressed protein, GAT4621, was measured in whole plant, roots and seeds at low levels. Based on the percentage of total protein in the seeds, GAT4621 corresponds to approximately 0.002% of seed protein in canola line 73496.

The identity and physicochemical and functional properties of the newly expressed protein were confirmed via a number of laboratory studies. These studies demonstrated that the GAT4621 protein conforms in size and amino acid sequence to that expected, is not glycosylated in the plant, and exhibits the expected functional enzyme activity. Enzymes with acetyl transferase activity are ubiquitous in nature and are not associated with known toxicity or allergenicity. Bioinformatic studies confirmed the lack of any significant amino acid sequence similarity between GAT4621 and known protein toxins and allergens. In addition, digestibility studies demonstrated that GAT4621 would be rapidly degraded in the gastric environment if ingested. As determined in previous assessments of this protein, the evidence supports the conclusion that GAT4621 is unlikely to be toxic or allergenic in humans.

Herbicide Metabolites

Expression of GAT4621 in canola line 73496 results in the acetylation of glyphosate in the plant and produces the metabolite *N*-acetyl glyphosate (NAG), which has no herbicidal activity. The results of metabolism studies of crops expressing GAT4621 demonstrated that NAG was formed as the main metabolite, and this was confirmed in the residue analysis of seeds from herbicide-treated canola line 73496 grown at test sites in Canada. After processing of the seeds, total herbicide residues (glyphosate and metabolites) in canola oil were below the limit of detection (<0.02 ppm). In the absence of any significant exposure to either parent herbicide or metabolites in canola oil, the risk to public health and safety is negligible.

Compositional Analyses

Detailed compositional analyses were conducted on seed from canola line 73496 and a non-GM control line grown in field trials in Canada and the United States. Canola line 73496 plants were sprayed with glyphosate herbicide at particular stages of growth, and all lines were grown under normal agricultural conditions.

Analyses included proximates (crude protein, crude fat, ash and total carbohydrates), fibre components, fatty acids, amino acids, micronutrients (minerals and vitamins) and anti-nutrients (glucosinolates, phytic acid, sinapine and tannins). The levels of these key constituents in the GM line were compared with those in the near isogenic control line, as well as to the normal ranges found in conventional canola varieties grown under similar conditions.

A small number of statistically significant differences were found in individual seed analytes between canola line 73496 and the control, however the composition of canola can vary significantly with the site and the prevailing agricultural conditions, and the differences reported are attributable to normal biological variation. The mean analyte levels in canola 73496 seed were within the range established for commercial canola varieties. Overall, the compositional data support the conclusion that there are no biologically significant differences in the levels of key constituents in seed from canola line 73496 when compared with conventional canola varieties grown under similar conditions.

Nutritional Impact

The expression of GAT4621 in the plant results in significant increases in the levels of five acetylated amino acids, in particular *N*-acetylaspartate and *N*-acetylglutamate, in seed from canola line 73496. Acetylated amino acids are naturally present in a variety of animal- and plant-derived foods and are normal constituents in the human diet. Notwithstanding a history of safe human consumption, analysis of processed canola seed demonstrated that acetylated amino acids are not detectable in refined canola oil. The approval of canola line 73496 as a source of food would not therefore have a significant nutritional impact on existing levels of acetylated amino acids in the diet.

Conclusion

No potential public health and safety concerns have been identified in the assessment of canola line 73496. On the basis of the data required from the Applicant, and other available information, food derived from herbicide tolerant canola line 73496 is as safe for human consumption as food derived from conventional canola varieties.

TABLE OF CONTENTS

SUMMARY AND CONCLUSIONS	1
ABBREVIATIONS AND ACRONYMS	2
1. INTRODUCTION	3
2. HISTORY OF USE	3
2.1 Host organism	3
2.2 Donor organism.....	4
3. MOLECULAR CHARACTERISATION	4
3.1 Method used in the transformation	4
3.2 Introduced gene construct	5
3.3 Breeding process and analyses	6
3.4 Characterisation of the genes in the plant	7
3.5 Stability of the genetic changes.....	11
3.6 Antibiotic resistance marker genes.....	11
3.7 Conclusion	11
4. CHARACTERISATION OF NOVEL PROTEINS	12
4.1 Function of GAT4621	12
4.2 Novel protein expression in plant tissues.....	15
4.3 Potential toxicity of the newly expressed proteins.....	17
4.4 Potential allergenicity of the newly expressed proteins.....	19
4.5 Herbicide metabolites	21
4.6 Conclusion	24
5. COMPOSITIONAL ANALYSIS	24
5.1 Key components	25
5.2 Study design and conduct for key components	25
5.3 Analyses of key components.....	26
5.4 Conclusion from compositional analysis	40
6. NUTRITIONAL IMPACT	40
REFERENCES	41

Abbreviations and Acronyms

ADF	acid detergent fibre
ADI	acceptable daily intake
AOAC	Association of Official Analytical Chemists
AOF	Australian Oilseeds Federation
ARfD	Acute Reference Dose
ATCC	American Type Culture Collection
BLAST	Basic Local Alignment Search Tool
bp	base pairs
DNA	deoxyribonucleic acid
dw	dry weight
EFSA	European Food Safety Authority
ELISA	enzyme linked immunosorbent assay
EPA	Environmental Protection Agency – United States of America
FAO	Food and Agriculture Organization of the United Nations
FARRP	Food Allergy Research and Resource Program
FSANZ	Food Standards Australia New Zealand
<i>gat4621</i>	gene derived from <i>B. licheniformis</i> encoding GAT4621 enzyme
GAT	glyphosate <i>N</i> -acetyltransferase
GAT4621	GAT enzyme optimised to use glyphosate as substrate
GM	genetically modified
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
ILSI	International Life Sciences Institute
kb	kilobase
kDa	kilo Dalton
LC/MS	high performance liquid chromatography/electrospray mass spectrometry
MALDI-TOF	Matrix-assisted laser desorption/ionisation-time of flight
NDF	neutral detergent fibre
OECD	Organisation for Economic Co-operation and Development
OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
PCR	polymerase chain reaction
RBD	refined, bleached, deodorised
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
U.S.	United States of America
WHO	World Health Organisation

1. Introduction

Canola line 73496 (OECD Unique Identifier DP-073496-4) has been genetically modified (GM) for tolerance to glyphosate, a broad spectrum herbicide used to control broadleaf weeds from pre-emergence to seven days pre-harvest. The Applicant claims herbicide-tolerant canola varieties have been useful in the management of weeds in canola crops, and their use reduces the overall volume of herbicides applied to the crop for weed control. From an agronomic perspective, the lack of competition with weeds for soil nutrients and moisture can also mean higher crop yields.

Tolerance to glyphosate in canola 73496 is achieved through constitutive expression of GAT4621 (glyphosate acetyltransferase), an enzyme that catalyses the acetylation of glyphosate, rendering it non-phytotoxic. The *gat4621* gene is a variant of three *gat* genes isolated from the common soil bacterium *Bacillus licheniformis*. The GAT4621 protein has been assessed previously by FSANZ in dual herbicide-tolerant corn line DP-98140-6 (approved in 2010) and herbicide-tolerant soybean line DP-356043-5 (approved in 2009), therefore is not new to the food supply.

The Applicant intends to market canola line 73496 under the trade name Optimum GLY™ Canola to Australian growers as an alternative to currently available commercialised canola varieties. There are no plans to grow this canola variety in New Zealand.

Canola is a high value crop grown for its seeds which are used as a source of vegetable oil for human consumption, and canola meal (dry matter after oil extraction) as a high protein livestock feed supplement.

2. History of use

2.1 Host organism

Canola is the name used for rapeseed (*Brassica napus*, *Brassica rapa* or *Brassica juncea*) crops that have less than 2% erucic acid (a fatty acid)¹ and less than 30 micromoles of glucosinolates per gram of seed solids (OECD 2011). Canola varieties were first developed in Canada in the 1950s, using traditional breeding techniques, in response to a demand for food-grade rapeseed products and animal feed with improved palatability. Rapeseed-derived products that do not meet the compositional standard cannot use the trademarked term, *canola*. Since the release of canola in Australia in 1980, it has become an important oilseed crop in most grain growing regions of Australia.

Rapeseed is the second largest oilseed crop in the world behind soybean. In 2012/13, the major oilseed rape producers globally were European Union (18.8m mt), Canada (13.3m mt) and China (12.6m mt) (USDA Foreign Agricultural Service; www.fas.usda.gov). While Canada is the largest exporter of canola, Australia regularly exports over one million tonnes of canola seed to Japan, Europe, China, Pakistan and other markets. This represents 15-20% of the world's trade in canola (AOF, 2007).

Canola seeds are processed into two major products, oil and meal. The oil is the major product for human consumption, being used in a variety of manufactured food products including salad and cooking oil, margarine, shortening and a range of prepared foods such as mayonnaise, sandwich spreads, creamers and coffee whiteners. The meal provides a good protein source in stockfeed for a variety of animals, primarily pigs, poultry and dairy cattle. Whole canola seeds are being used increasingly in products such as breads.

¹ Codex Standard for edible Vegetable Oils.

2.2 Donor organism

2.2.1 *Bacillus licheniformis*

The *gat4621* gene is derived from the gram-positive bacterium *Bacillus licheniformis*, part of the subtilis group along with *B. subtilis* and *B. pumilus*. The bacterium is commonly found in soil and bird feathers. It can exist in spore form to resist harsh environments, or in a vegetative state when environmental conditions are more favourable.

B. licheniformis is an approved bacterial source for the production of a number of enzymes used as food processing aids, such as α -amylase, hemicellulase, pullulanase (a glucanase) and serine protease, and has been used in the USA, Canada and Europe in the fermentation industry (Rey *et al* 2004). The U.S. Environmental Protection Agency has determined that this organism presents a low risk to human health and the environment when used under specific conditions for general commercial use (EPA, 1996). However, while *B. licheniformis* is widespread in the environment and people are regularly exposed to it without any associated adverse effects, non-proteinaceous toxins produced by isolates of *B. licheniformis* have been associated with food in food poisoning incidents (see Salkinoja-Salonen *et al.*, 1999).

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 process itself;
- the genetic stability of the inserted DNA and expressed traits.

Studies submitted:

PHI-2009-134: Characterisation of DP-073496-4 Canola: Insertion Integrity, Stability, Copy Number and Backbone Analysis (2011/2013)

PHI-2010-086/040: Sequence Characterisation of Insert and Flanking Genomic Regions of Canola Event DP-073496-4 (2011) **CCI**

PHI-2012- 283: ORF Analysis at the Insertion Site of Canola Event DP-073496-4 (2012)

PHI-2010-089/010: Segregation Analysis Using Endpoint PCR of Five Generations and Qualitative ELISA of Three Generations of Glyphosate Resistant Canola Events DP-073496-4 and DP-061061-7 (2010)

3.1 Method used in the transformation

Canola line 73496 was developed by microprojectile bombardment of microspores prepared from the parental canola line 1822B, according to Chen and Tulsieram, 2007. Gold particles coated with a specific fragment (PHP28181A) from the plasmid PHP28181 (see Figure 3) were used in the transformation. This method does not require the use of bacteria as intermediate host organisms. After transformation, embryogenic microspores were grown in tissue culture over several weeks, selecting for glyphosate-tolerant transformants in the presence of glyphosate (0.1mM). Germinated shoots were transferred to growth medium supplemented with glyphosate for further selection.

Plants regenerated from transformation and tissue culture (designated T0 plants) were selected for molecular characterisation, analysis of herbicide tolerance and agronomic

evaluations. The selection process resulted in the identification of line 73496 as the event with the most desired phenotypic and molecular characteristics.

3.2 Introduced gene construct

The DNA fragment used in the transformation of canola to produce line 73496 was PHP28181A (2112 bp), derived from plasmid PHP28181 (4770 bp). The required fragment was cleaved from the full-length plasmid using restriction endonucleases HindIII and NotI, and contains the gene of interest, *gat4621*, and regulatory elements necessary for expression in plant cells. Information on the genetic elements in the transforming DNA fragment is summarised in Table 1.

Table 1: Description of the genetic elements contained in PHP28181A (2112 bp)

Genetic element	Location on transforming DNA fragment (size, bp)	Source	Description & Function	References
<i>Polylinker sequence</i>	1 – 7 (7)	Plasmid PHP28181	<ul style="list-style-type: none"> Sequence containing restriction sites used in DNA cloning 	
<i>UBQ10 promoter</i>	8 – 1312 (1305)	<i>Arabidopsis thaliana</i>	<ul style="list-style-type: none"> Version of the promoter region from the plant <i>UBQ10</i> polyubiquitin gene leading to constitutive expression of transgenes Directs transcription of genes in plant cells 	Norris <i>et al.</i> , 1993 (developed by E.I. duPont de Nemours and Company)
<i>Intervening sequence</i>	1313 – 1335 (23)			
<i>gat4621</i>	1336 – 1779 (444)	<i>Bacillus licheniformis</i>	<ul style="list-style-type: none"> Optimised (synthetic) gene encoding glyphosate N-acetyltransferase (GAT) 4621 protein Confers glyphosate tolerance 	Castle <i>et al.</i> , 2004 Siehl <i>et al.</i> , 2007
<i>Intervening sequence</i>	1780 – 1796 (17)			
<i>pinII terminator</i>	1797 – 2106 (310)	<i>Solanum tuberosum</i>	<ul style="list-style-type: none"> Terminator region from the plant (potato) proteinase inhibitor II gene Terminates transcription and directs polyadenylation 	An <i>et al.</i> 19895; Keil <i>et al.</i> , 1986
<i>Polylinker sequence</i>	2107 – 2112 (6)	Plasmid PHP28181	<ul style="list-style-type: none"> Sequence used in DNA cloning 	

As PHP28181A was cleaved and purified away from the rest of the plasmid, other elements on the plasmid backbone such as the selectable marker gene *bla* (β -lactamase) were not expected to be transferred to the plant. This was confirmed later by Southern blot analysis (see below). The antibiotic resistance gene is useful only for cloning purposes in *E. coli* to maintain the whole plasmid.

3.2.1 *gat4621* gene cassette

The *gat4621* gene was the result of a fragmentation-based, multigene shuffling process with the aim of producing an optimised gene sequence encoding an enzyme with higher efficiency and increased specificity for glyphosate. The *gat4621* gene was derived from native *gat* sequences isolated from three strains of *B. licheniformis*. The relevant gene in these *B. licheniformis* strains was identified using a mass spectrometry method to detect the product of interest, N-acetylgliphosate (Castle *et al.*, 2004).

The process of fragmentation and recombination involved selection of progeny with improved properties to use as a source of genes in the next round of shuffling.

In the case of the *gat4621* gene, the optimisation process was repeated eleven times using a combination of multi-gene shuffling and the introduction of genetic diversity via the polymerase chain reaction (PCR).

In canola line 73496, expression of the *gat4621* gene is under the control of regulatory elements from other plants - the promoter from the polyubiquitin gene (*UBQ10*) from *Arabidopsis thaliana*, and 3' terminator from the proteinase inhibitor II gene (*pinII*) of *Solanum tuberosum* (common potato).

3.3 Breeding process and analyses

As described in section 3.1, plants that were regenerated after the transformation of the parental line 1822B and tissue culture (designated T0 plants) were selected for further characterisation by molecular and phenotypic analyses. The subsequent breeding of canola line 73496 proceeded as outlined in Figure 1, to produce specific generations for characterisation analyses, as presented in Table 2.

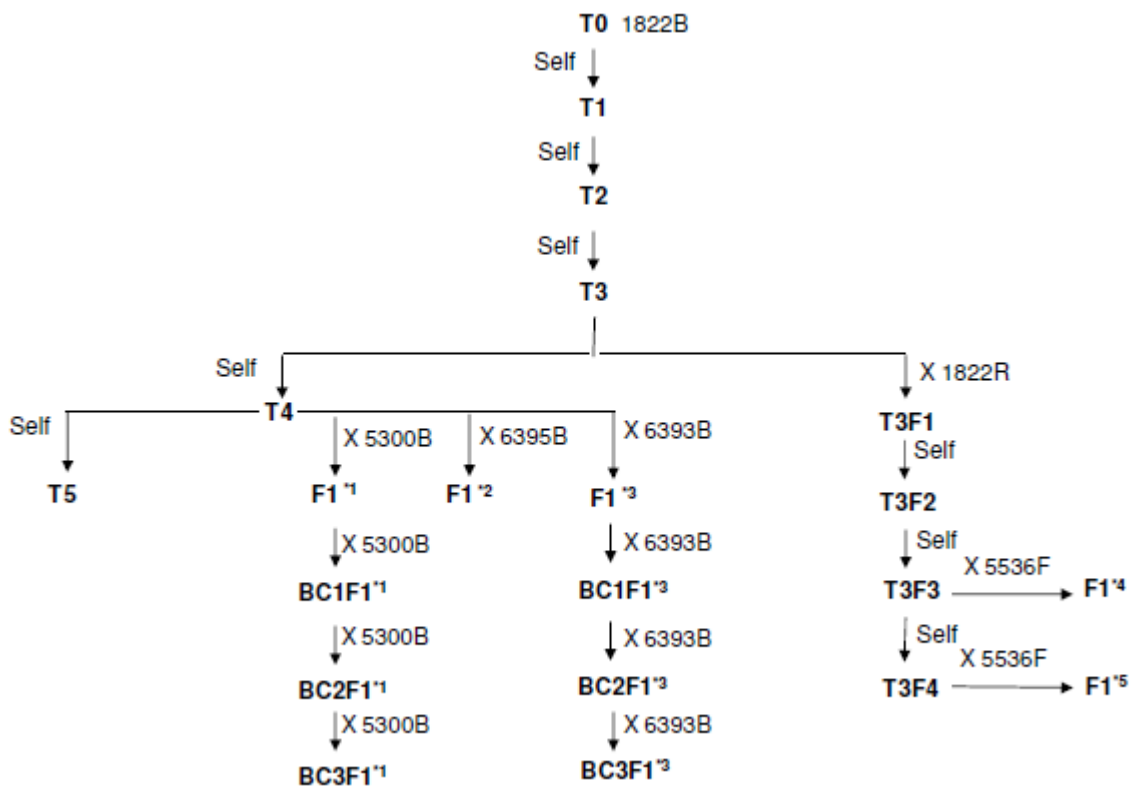


Figure 1 Breeding Diagram for Canola line 73496 and Generations Used for Analyses (Key to symbols in Table 2)

Table 2: Generations and Comparators Used for Analysis of Canola line 73496

Analysis	Generation (from Figure1)	Comparators and Commercial Reference Lines
Molecular characterisation	T2, T3, F1 ^{*2} , T3F2, T3F3	1822B, 1822R, 6395B
Genetic inheritance	F1 ^{*1,*3} , BC1 F1 ^{*1,*3} , BC2 F1 ^{*1,*3} , BC3 F1 ^{*1,*3} , T3F2	Not applicable
Concentrations of GAT4621 & acetylated amino acids	F1 ^{*4}	5536F x 1822R (near isogenic line)
Compositional analyses	F1 ^{*4}	5536F x 1822R (near isogenic line) and Pioneer commercial lines 46A65, 45H72, 45H73, 46H02 & 44A89.

3.4 Characterisation of the genes in the plant

Analyses were undertaken in order to fully characterise at the molecular level the genetic modification in canola line 73496. These analyses focussed on the exact nature of the inserted genetic elements, and whether any unintended gene re-arrangements may have occurred as a consequence of the transformation procedure. Several techniques including Southern blots, PCR and DNA sequence analysis were used for the characterisation.

3.4.1 *Transgene copy number and insertion integrity*

Southern blots were used to determine the number of T-DNA insertions and the sequence integrity of the introduced DNA in canola line 73496, and test for the presence or absence of plasmid vector backbone sequences that could have been transferred to the plant. For these analyses, seeds from the T2, T3, T3F2, T3F3 and F1^{*2} generations of GM canola plants (see Figure 1), and from conventional canola lines 1822B, 1822R and 6395B were germinated in growth chambers to generate plant tissue for genomic DNA extraction. Due to the number of traditional breeding steps after transformation of the parent 1822B, the additional conventional lines were necessary to more closely represent the genetic background of canola 73496. Leaf tissue was harvested from 10 individual plants at approximately three weeks, four weeks and five weeks after planting of the seeds of the relevant generations.

Plasmid DNA from PHP28181 was used as a positive control for Southern analysis to verify hybridisation of the probes and as a molecular size reference for fragments internal to the inserted DNA. The six probes used in this study were prepared by PCR amplification using plasmid PHP28181 as the template, and labelled via incorporation of a digoxigenin (DIG) labelled nucleotide. Probes A, B and C (as shown on Figure 2) were used to detect the inserted DNA. For the *UBQ10* promoter probe (A), two non-overlapping segments (674 and 567 bp) were generated in separate PCR reactions and combined for hybridisation with genomic DNA.

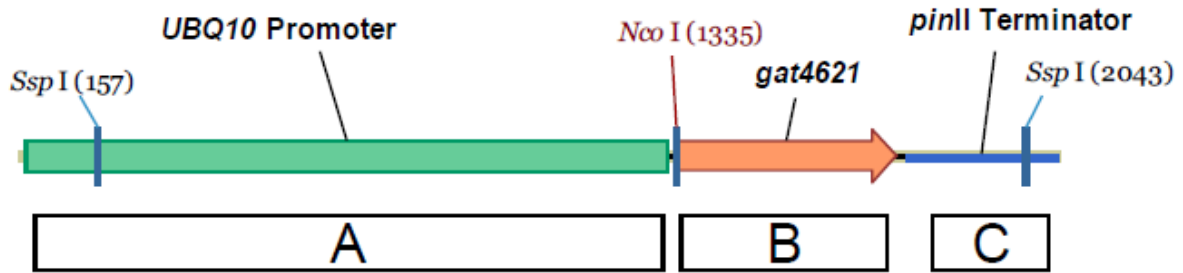


Figure 2 Schematic map of PHP28181A (2112 bp) used for transformation of canola. The boxes show the relative size and location of three molecular probes (A, B, C) used in Southern analysis.

Probes 1-3 (as shown in Figure 3) were used to check for the presence of plasmid backbone in the canola genome. Backbone probes 1, 2 and 3 were generated as separate overlapping segments (643, 831 and 1268 bp) and were combined for the hybridisation step. These three probes cover the entire plasmid backbone region.

The same sets of restriction enzymes were used to cleave DNA from canola 73496 and control lines, in addition to the reference plasmid PHP28181. This allowed direct analysis of the banding pattern obtained for each sample/enzyme combination.

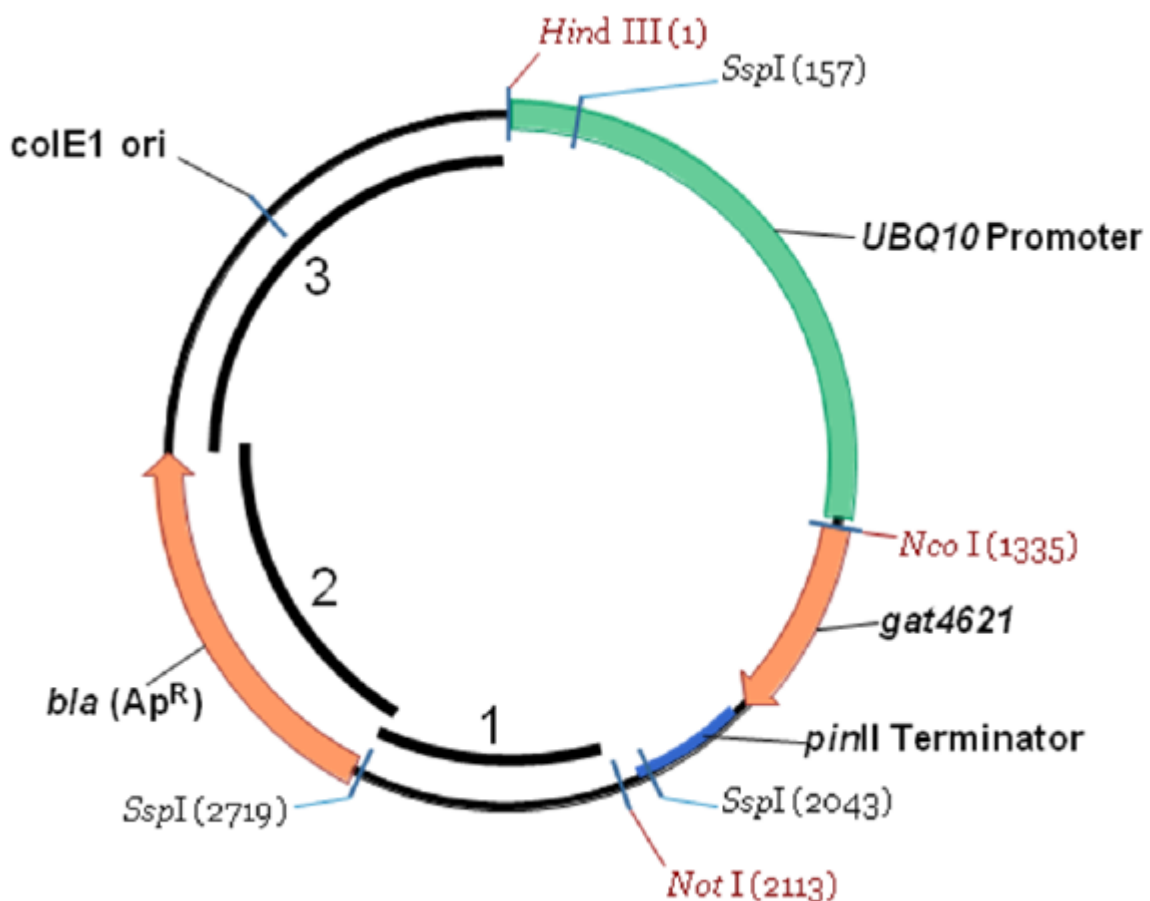


Figure 3 Circular map of PHP28181 (4770 bp) indicating the relative size and location of three plasmid backbone probes (1-3) used in Southern blots.

Analysis of multiple Southern blots using DNA obtained from five generations of canola line 73496 demonstrated identical hybridisation patterns for all generations.

The results indicated that one copy of the *gat4621* expression cassette was integrated as a single insertion event in this line. No additional DNA fragments were observed that would indicate insertion of an incomplete expression cassette, and no sequences derived from the plasmid backbone were detected in canola 73496 genomic DNA.

3.4.2 DNA sequence analysis

Analysis of event-specific PCR products and DNA sequencing complement the results obtained with Southern blots, and allow the exact nature of the inserted DNA to be described. This confirms the organisation and sequence of the genetic elements transferred into canola 73496, relative to the sequences in the plasmid fragment PHP28181A used in the transformation.

Genomic DNA was extracted from leaf tissue obtained from plants grown from seed from the T2 generation of canola 73496, and from control seed 1822B. PCR primer pairs were designed to amplify three overlapping DNA regions spanning the entire length of the insert and the adjacent plant genomic DNA flanking the 5' and 3' ends of the insert in canola 73496, as represented in Figure 4. The PCR products were separately cloned, and both sense and anti-sense strands were sequenced.

PCR was also performed on conventional control DNA using a primer pair that hybridised to genomic sequence in the 5' and 3' flanking regions, to obtain an amplicon corresponding to the site of insertion in the untransformed parental control.

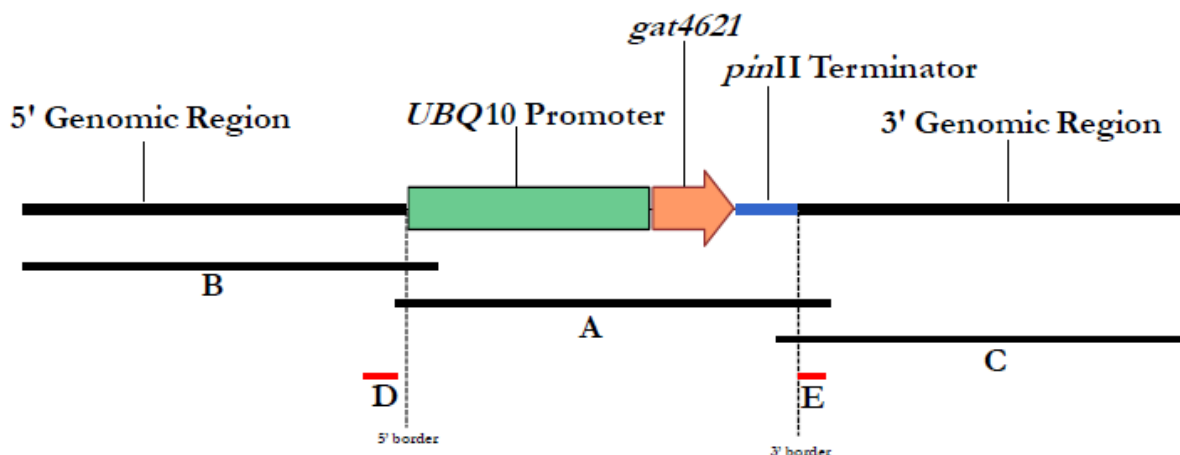


Figure 4: Schematic representation of the insert and the 5' and 3' flanking genomic regions of canola line 73496. The illustration shows the approximate position of three PCR amplicons (Products A, B, C) used for DNA sequencing. PCR amplicons labelled D and E correspond to the original DNA sequence in non-transformed canola.

Fragment A (2452 bp) corresponded to the entire cassette introduced into canola 73496 (2109 bp), in addition to 90 bp of the flanking genomic DNA at the 5' end of the insert, and 253 bp at the 3' end of the insert. Fragment B (2204 bp) consisted of 2003 bp of 5' flanking genomic sequence, and 201 bp of the *UBQ10* promoter region. Amplification of fragment C resulted in an amplicon of 2154 bp, consisting of 116 bp of the *pinII* terminator region and 2038 bp of genomic DNA at the 3' end of the insert. In total, 6150 bp of DNA sequence in canola line 73496 was fully characterised.

The DNA sequence analysis determined that the insert in canola line 73496 is identical to the sequence of the transforming plasmid fragment PHP28181A except for the first three

nucleotides at the 5' end of the transformation fragment. The small deletion does not affect the function of the promoter and raises no safety concerns.

PCR was performed on genomic DNA from canola 73496 and control canola DNA to amplify products labelled D (275 bp) and E (255 bp) in Figure 4. The DNA sequence in the transgenic and control canola samples was identical, confirming it to be original canola sequence at the site of insertion.

3.4.3 Open reading frame (ORF) analysis

Bioinformatics assessment of any putative ORFs inherent to the inserted DNA or contiguous with the adjacent plant genomic DNA is used to identify whether any might encode a peptide with homology to known toxins or allergens, or otherwise indicate a need for further characterisation if translated. The bioinformatics analysis is entirely theoretical and does not inform on whether any of the ORFs are actually transcribed into RNA and translated into protein. Putative ORFs in all six reading frames are considered (that is, three forward reading frames and three in the reverse orientation).

A total of forty ORFs within the defined inserted DNA (2327 nucleotides) in canola line 73496 were identified and translated *in silico* into amino acid sequence for further investigation. Putative peptides consisting of a minimum length of eight amino acids were compared to allergen and all protein databases using typical bioinformatics tools. In terms of potential allergenicity, eight contiguous and identical amino acids were defined as the minimum requirements for a possible immunologically relevant epitope (Silvanovich *et al.* 2006). The FASTA35 sequence alignment tool was used to assess structural relatedness between the query sequence and any protein sequences in the FARRP12² dataset (University of Nebraska, Release 12, February 2012). A 35% or greater identity over any sequence of 80 or more amino acids between the query sequence and an allergen sequence was used to indicate the threshold potential for cross-reactivity (Codex 2004).

The extent of structural relatedness between query sequences and the NCBI³ Protein dataset (Release 190.0 on 15 June 2012) was evaluated by detailed visual inspection of the alignment, the calculated percent identity and alignment length. For the reading frames spanning the junction regions, a conservative *E*-score cutoff value of 1.0 was used to find biologically meaningful alignments. The higher *E*-score cutoff finds sequences with more limited similarity and is therefore likely to generate higher numbers. Default BLAST parameters were used and the number of alignments was set to return the maximum number 2000.

The results of comparative searches of the allergen and protein databases revealed no relevant structural similarity to known allergens or toxins for any of the ORFs occurring across the inserted DNA and 5' and 3' junction regions. No short peptide matches of eight amino acids were found between canola 73496 sequences and proteins in the allergen database. Interrogation of the protein dataset using the same process, identified sequences from the acetyltransferase superfamily of proteins, as could be expected from the function of the GAT4621 protein.

Overall, the results from the bioinformatics analyses demonstrate that the presence of the transgene cassette in canola 73496 is unlikely to give rise to novel polypeptides with

² Food Allergy Research and Resource Program

³ The National Center for Biotechnology Information (NCBI) is part of the United States National Library of Medicine (NLM), a branch of the National Institutes of Health (NIH)

significant similarity to known proteins showing toxic, allergenic or other biologically adverse properties.

3.5 Stability of the genetic changes

Data demonstrating the stability of the introduced trait over a number of successive plant generations must be provided. Stability can be assessed both analytically and phenotypically. The molecular analyses include techniques such as Southern blots to probe specifically for the inserted DNA in seeds or other plant tissues from each generation. Phenotypic analysis refers to the observed expression of the introduced trait that is carried over to successive generations. Genetic stability can be quantified by a trait inheritance analysis (chemical, molecular and visual) of progeny to determine Mendelian heritability.

3.5.1 Patterns of inheritance

During development of canola 73496, phenotypic and genotypic segregation data were recorded to assess the inheritance and stability of the inserted DNA using Chi-square (χ^2) analysis. Four segregating generations (T3F2, BC1F1^{*1*3}, BC2F1^{*1*3} and BC3F1^{*1*3}) and one non-segregating generation (F1^{*1*3}) were evaluated. The breeding history of these five generations is shown in the breeding diagram in Figure 1. As shown in the diagram, *1 or *3 represents populations with two different genetic backgrounds, as a result of traditional breeding with a proprietary canola line that did not contain the *gat4621* coding sequence. The χ^2 analysis is based on comparing the observed segregation ratio to the expected segregation ratio (1:1) according to Mendelian principles.

The presence of the *gat4621* coding sequence was determined by event specific PCR analyses performed on leaf punches from 100 seedlings (one sample per plant) of each generation. After sampling, plants were assessed phenotypically with application of glyphosate and by visually evaluating each plant for the presence of herbicide injury. A positive plant exhibited no visible herbicidal injury.

The results of the χ^2 analysis showed the observed segregation ratio was consistent with the expected segregation ratio for four generations of plants. These results support the conclusion that the *gat4621* coding sequence in canola line 73496 is at a single locus within the canola genome and is inherited according to Mendelian principles of inheritance.

3.6 Antibiotic resistance marker genes

No antibiotic resistance marker genes are present in canola line 73496. The molecular analysis confirmed that no plasmid backbone was transferred to the canola genome during transformation. Phenotypic analysis of the glyphosate-tolerance trait was used to select the major lines of interest.

3.7 Conclusion

Canola line 73496 contains *gat4621*, a gene synthesised from *Bacillus licheniformis* sequences via a gene shuffling process to express glyphosate N-acetyltransferase, an enzyme optimised to use the herbicide glyphosate as preferred substrate. GAT4621 catalyses the acetylation of glyphosate, producing N-acetylglyphosate, a compound with no herbicidal activity. The *gat4621* gene was introduced via biolistic transformation.

Detailed molecular analyses indicate one complete copy of the gene expression cassette is present at a single insertion site in the genome of canola line 73496.

No extraneous sequences derived from the plasmid were incorporated into the plant genome. The introduced genetic elements and the expression of the new protein were

shown by phenotypic analysis and molecular techniques to be stable in canola line 73496, and inherited across multiple generations of conventional plant breeding. The pattern of inheritance supports the conclusion that the herbicide-tolerance trait is a single locus within the canola genome and is passed on in accordance with Mendelian principles of inheritance.

4. Characterisation of novel proteins

In considering the safety of newly expressed proteins, it is important to consider that ingestion of a large and diverse repertoire of proteins is part of a normal human diet. Almost all of the vast numbers of proteins in foods are consumed without any adverse effects, although a small number have the potential to affect health, for example, because they are allergenic, or they have anti-nutritional properties (Delaney *et al.*, 2008). Proteins that are toxic in mammals are relatively rare; some examples include ricin from the castor oil plant and amatoxin oligopeptides from poisonous mushrooms.

As proteins perform a wide variety of biochemical functions in living organisms, their characteristics and possible effects are considered during the safety assessment of GM foods. This includes the potential of a newly expressed protein to be toxic, allergenic or exhibit anti-nutritional properties if present in the diet. To effectively identify any potential hazards requires knowledge of the characteristics of the newly expressed protein and its localisation and levels in plant tissues, particularly the food-producing parts of the plant. The evaluation includes a detailed understanding of the biochemical function and phenotypic effects of the newly expressed protein. It is also necessary to determine if any post-translational modifications are present, particularly any that were not evident in the source organism.

The newly expressed protein in canola line 73496 is the GAT4621 enzyme. Laboratory analyses and bioinformatics studies were done to determine the identity, physicochemical properties, *in planta* expression, bioactivity and potential toxicity and allergenicity of the protein. Because the expression of transgenic proteins *in planta* is usually too low to allow purification of sufficient quantities for these studies, a bacterial expression system was used to generate larger amounts of the protein. The equivalence of the protein produced in *E. coli* to the plant-produced protein was determined as part of the protein characterisation.

4.1 Function of GAT4621

Studies submitted:

PHI-2010-020: Characterisation of GAT4621 Protein Derived from Canola Containing Event DP-073496-4 and Equivalency Assessment with the GAT4621 Protein Derived from a Microbial Expression System (2010)

PHI-2007-020: Characterisation of GAT4621 Protein Derived from a Microbial Expression System (2007)

Tolerance to glyphosate in canola line 73496 is conferred by the expression in the plant of GAT4621, a novel GAT enzyme that catalyses the *N*-acetylation of glyphosate, resulting in the formation of a herbicidally inactive compound, *N*-acetylglyphosate (see Figure 5). The GAT4621 protein in canola line 73496 is equivalent to the protein expressed in corn line DP-098140-6 which was assessed by FSANZ in 2009 under Application A1021 and subsequently approved (FSANZ 2010). As in GAT corn, aminomethylphosphonic acid (AMPA) and *N*-acetyl AMPA are also formed as minor metabolites during metabolism of glyphosate in 73496 canola plants.

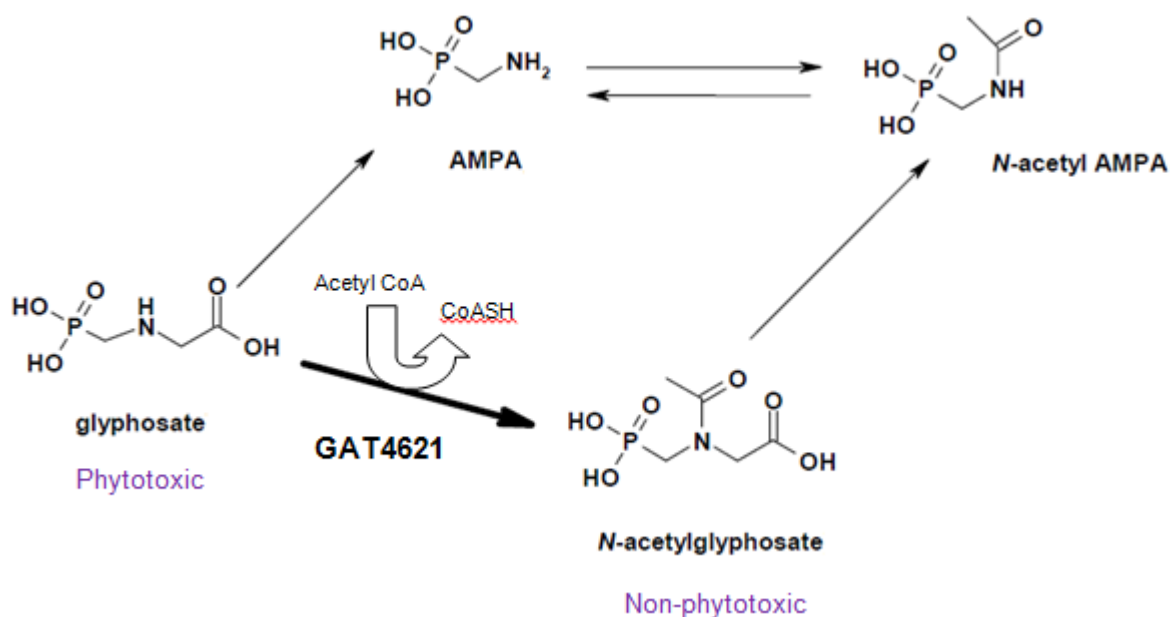


Figure 5: Enzymatic detoxification of glyphosate by GAT4621 expressed in canola line 73496

4.1.1 Mode of action of glyphosate tolerance

Glyphosate is a non-selective, broad-spectrum, foliar-applied herbicide first commercialised in 1974 and now widely used for the management of annual, perennial and biennial herbaceous species of grasses, sedges and broadleaf weeds in agricultural settings, as well as woody brush and tree species (Baylis 2000; Bradshaw et al 1997).

In plants, the primary target of glyphosate is 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), an essential enzyme in the biosynthesis of aromatic amino acids (L-phenylalanine, L-tyrosine and L-tryptophan). Glyphosate binds to and blocks the activity of EPSPS, disrupting protein synthesis and leading to plant cell death.

Mechanisms for conferring tolerance to glyphosate in GM crops have included the introduction of microbial variants of EPSPS that are insensitive to glyphosate (eg. CP4 EPSPS from *Agrobacterium tumefaciens*) or a form of the endogenous EPSPS from corn which has been altered at the glyphosate binding site to reduce binding of the herbicide (eg. modified EPSPS from *Zea mays*). In both cases, the introduced variant of EPSPS continues to function in the presence of the herbicide, despite inactivation of the endogenous EPSPS.

In contrast, the GAT proteins provide tolerance to glyphosate by enzymatically converting the herbicide to a non-phytotoxic form, N-acetylglyphosate. GAT enzymes catalyse the addition of an acetyl group to the secondary amine of glyphosate using acetyl coenzyme A (acetyl CoA) as the donor (Castle et al 2004).

4.1.1.1 Derivation of GAT4621

GAT proteins are members of the GCN5-related family of N-acetyl transferases, also known as the GNAT superfamily (Dyda et al., 2000). This large enzyme superfamily contains over 10,000 representatives and is found in plants, animals and microbes. Members of the GNAT superfamily contain a conserved GNAT motif, but are otherwise highly sequence divergent.

Expression of GAT proteins in several transgenic plant species (eg. corn and soybean) has been shown to confer functional glyphosate tolerance.

A detailed description of the stepwise process of developing GAT4621 by selecting optimal microbial GAT enzyme activity in the laboratory is included in Section 4.1 of Supporting Document 1 to the Approval Report for Application A1021. From the starting point of first screening for glyphosate acetyltransferase activity in a collection of several hundred *B. licheniformis* isolates, GAT activity was eventually improved approximately 7000-fold over the activity of the native enzymes (Castle *et al*, 2004; Siehl *et al*, 2005).

4.1.2 Characterisation of GAT4621 and equivalency studies

In order to accept the relevance of the safety data generated using *E. coli*-produced GAT4621, it was necessary for the Applicant to demonstrate equivalence with the GAT4621 protein produced in canola line 73496. Based on the gene sequence, the GAT4621 protein is 147 amino acids, and has a molecular mass of approximately 17 kDa.

Several different approaches outlined below were used to characterise the GAT4621 protein expressed in canola line 73496, and demonstrate that it exhibits equivalent physicochemical properties to a bacterially-produced GAT4621 protein used as reference material in later studies. A small quantity of the GAT4621 protein was purified from approximately 80g of leaf tissue derived from 73496 canola plants grown in a greenhouse facility.

Equivalency data for the microbially-produced GAT4621 from studies completed in 2007 were resubmitted for this part of the assessment. In some procedures, the microbially-produced protein was reanalysed in the same experiment with the canola-produced GAT4621 protein. Microbially-produced GAT4621 was evaluated by FSANZ as part of Application A1021 (herbicide-tolerant corn line DP-098140-6).

The techniques used were:

- SDS-PAGE
- Western blot analysis
- MALDI-TOF Tryptic Mass Fingerprint Analysis (MALDI-TOF MS)
- N-terminal amino acid sequencing
- Glycosylation analysis

4.1.2.1 SDS PAGE

SDS-PAGE was used to determine the apparent molecular weight (mobility) and purity of the plant-produced GAT4621 protein for comparison with that produced in *E. coli*. On staining of the plant-derived sample, the prominent band corresponded to a protein with apparent molecular weight of 16.5 kDa. The microbially expressed GAT4621 protein migrated on a separate gel as a single band with a molecular weight of approximately 16 kDa. These data are consistent with the expected theoretical size of the GAT4621 protein, based on its known amino acid sequence.

4.1.2.2 Western blot analysis

Western blot analysis was performed to confirm the identity of the GAT4621 protein and compare the immunoreactivity of the protein as extracted from canola line 73496 with the microbially-produced GAT4621 protein on the same gel system. The membrane was probed with a GAT-specific mouse monoclonal antibody. The identity of the GAT4621 protein expressed in canola line 73496 was confirmed. The plant- and *E. coli*-produced GAT4621 proteins produced identical bands on the Western blot.

4.1.2.3 MALDI-MS Tryptic Mass Fingerprint Analysis

The identity of the GAT4621 protein was confirmed using MALDI-MS analysis of peptides derived from the protein digested with trypsin. This standard technique, also referred to as peptide mapping, matches a sufficient number of observed tryptic peptide fragment masses with those predicted from the amino acid sequence of the protein. Protein identification is considered reliable with this method where $\geq 40\%$ of the protein sequence is identified by matching the observed masses with the expected masses of the tryptic peptide fragments.

Overall, the identified peptides accounted for 130 of the 147 residues of the deduced GAT4621 amino acid sequence, which corresponds to 88.4% of the protein sequence. In the earlier characterisation study of microbially-produced GAT4621, the same method identified peptides that accounted for 112 of the 147 amino acid residues, or 76% of the protein. These results serve to confirm the identity of the GAT4621 protein expressed in canola line 73694, and demonstrate the equivalence of the microbially-produced GAT4621 protein to that produced in the canola plants.

4.1.2.4 N-terminal sequencing

N-terminal amino acid sequencing was performed to confirm the molecular identity of the GAT4621 protein derived from canola 73694 plants. As expected, the N-terminal methionine residue was not present, which is consistent with MALDI-MS results. The primary sequence of the plant-derived protein matched residues 2-14 of the deduced GAT4621 amino acid sequence, and also matched the N-terminal sequence of the GAT4621 protein produced in the *E.coli* expression system.

4.1.2.5 Glycosylation analysis

A commercial glycoprotein staining kit was used to determine whether the GAT4621 protein produced in canola line 73496 undergoes post-translational modification with covalently bound carbohydrate moieties. Horseradish peroxidase was used as a positive control in the assay, and soybean trypsin inhibitor was analysed as a negative control. No glycosylation signal was observed in the equivalent sample amount of GAT4621 derived from canola line 73496. In a previous experiment, as expected, the *E. coli* - produced GAT4621 was also negative in this assay system. These results are consistent with the GAT4621 protein produced in canola line 73496 being non-glycosylated.

4.2 Novel protein expression in plant tissues

Study submitted:

PHI-2009-039/010: Quantification of GAT4621 Protein in Tissues of Herbicide-Treated Canola Lines Containing Event DP-073496-4: U.S. and Canada Test Sites (2010)

Expression of GAT4621 is expected in all plant tissues since the *gat4621* gene is coupled with a plant promoter that gives rise to constitutive expression (refer to Table 1).

Tissues from canola line 73496 and whole plant samples were collected at various stages of growth from up to six trial site locations (four in Canada and two in the U.S.) in 2009. Concentrations of GAT4621 were determined in whole plant (the entire above-ground portion of the plant), roots and seed samples from canola 73496 following treatment with herbicide, and the near isoline control canola using a validated, quantitative enzyme-linked immunosorbent assay (ELISA). Samples were assayed in duplicate and a standard curve (analysed in triplicate) was included on each ELISA plate.

Sample concentrations obtained from the software analysis were converted from ng/ml to ng/mg tissue on a dry weight basis. The results are presented in Table 3.

Samples were harvested at the following developmental stages, as described by Lancashire *et al.* (1991):

BBCH15 – five true leaves unfolded (whole plant sample)

BBCH33 – three visibly extended internodes) (whole plant sample)

BBCH65 – full flowering; 50% of flowers open on main raceme, older petals falling (whole plant and root samples)

BBCH90 – senescence (seed sample)

Table 3: Mean Concentrations of GAT4621 Protein Levels in Tissues from Herbicide-treated Canola line 73496 Grown at Multiple Sites.

Tissue	GAT4621 Concentration [ng/mg dry weight ± SD (range)] ^a			
	Plant Growth Stage ^b			
	BBCH15	BBCH33	BBCH65	BBCH90
Whole Plant (n=24)	6.9 ± 1.3 (3.9-10)	5.3 ± 1.2 (3.1-8.4)	5.2 ± 0.88 (3.9-7.6)	NC
Root (n=24)	NC ^c	NC	6.6 ± 2.4 (3.9-13)	NC
Seed (n=20)	NC	NC	NC	6.2 ± 0.94 (4.8-8.4)

^a Values are expressed as the mean of four replicate tissue samples collected from each of six locations except seed samples, which were only collected from five locations. SD = standard deviation; range, in parentheses, denotes the lowest and highest individual value across sites.

^b Plant growth stages: BBCH15 – five true leaves unfolded; BBCH33 – three visibly extended internodes; BBCH65 – full flowering; BBCH90 – senescence (Lancashire *et al.*, 1991).

^c NC = Not collected.

The results show that GAT4621 is expressed in all plant tissues at similar levels in canola 73496. As expected, the non-GM control plants did not contain detectable amounts of GAT4621; all control samples were less than the lowest limit of quantitation (<0.22 ng/mg tissue dry weight for seed) (data not shown).

4.2.1 Dietary exposure to GAT4621

Canola seeds are the source of food for human consumption. Seeds are primarily used to produce refined (canola) oil, although it has been noted that whole seeds have minor uses in some food products. The mean level of GAT4621 in 73496 canola seed was 6.2 ng/mg tissue dw (Table 3). The average percentage of crude protein in GM canola line 73496 seeds is 25.9% on a dry weight basis (see Table 6 in Section 5.3.1). Therefore GAT4621 is approximately 0.002% of total seed protein.

Taken together, the low levels of GAT4621 in canola seed from line 73496 in conjunction with the removal of protein during the oil extraction process, means that dietary exposure to GAT4621 via the consumption of canola oil would be virtually zero.

4.2.2 Conclusion on identity and function of newly expressed proteins

The studies described above confirmed the identity and expected function of the newly expressed GAT4621 protein and determined that it is not glycosylated in canola line 73496. The studies also demonstrated that the protein produced in a bacterial expression system is structurally and functionally equivalent to that expressed in the plant. GAT4621 is expressed in all plant tissues tested including roots and seeds at relatively low levels, around 6 ppm.

4.3 Potential toxicity of the newly expressed proteins

While the vast majority of proteins ingested as part of the diet are not typically associated with toxic effects, a small number may be harmful to health. Therefore, if a GM food differs from its conventional counterpart by the presence of one or more newly expressed 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 newly expressed protein will not cause adverse effects, and when ingested will be metabolised like most other dietary proteins.

The assessment focuses on: whether the newly expressed protein has a prior history of safe human consumption, or is similar to proteins that have been safely consumed in food; amino acid sequence similarity with known protein toxins or anti-nutritional proteins; structural properties of the newly expressed protein including whether it is resistant to heat or processing and/or digestion. Where results from the biochemical, bioinformatics, digestibility or stability studies indicate a reason to further investigate the protein, appropriate acute oral toxicity studies in animals can sometimes provide further assurance of safety.

4.3.1 History of human consumption

As described in Section 2.2.1, the organism from which the *gat4621* gene is derived is *Bacillus licheniformis*, which has been used in the United States, Canada and Europe in the fermentation industry for production of food enzymes, such as alpha-amylase, cyclodextrin glycosyltransferase, hemicellulase, proteases and pullulanase (Rey *et al.* 2004).

GAT4621 was constructed by consecutively selecting from three native gene sequences for a product with improved acetylation activity using glyphosate as substrate. The GAT4621 amino acid sequence is 75-78% identical and 90-91% similar to the three GAT proteins from which it is derived. As summarised above in Section 4.1.1.1, GAT proteins are members of the GCN 5-related family of *N*-acetyltransferases (also known as the GNAT superfamily). The GNAT superfamily is one of the largest group of enzymes with over 10,000 representatives from plants, animals and microbes distributed in many universal metabolic pathways. Members of this superfamily all contain highly conserved GNAT motifs but otherwise have high sequence diversity (Vetting *et al.* 2005). Humans therefore have extensive prior dietary exposure to proteins in this enzyme superfamily through direct ingestion with no adverse effects.

4.3.2 Similarities with known protein toxins

Studies submitted:

PHI-2007-009/073: Evaluation of the Amino Acid Sequence Similarity of the GAT4621 Protein to the NCBI Protein Sequence Datasets (2012)

Bioinformatic analyses are used to assess whether newly expressed proteins share any amino acid sequence similarity with proteins that may be harmful to human health, for example known protein toxins, whose sequences are stored in a number of available protein sequence databases.

Proteins that share a high degree of similarity throughout the entire sequence are often homologous, and often have common secondary structures, common three-dimensional configuration and consequently, may share similar functions.

The GAT4621 protein was evaluated against protein sequences present in the National Center for Biotechnology Information⁴ (NCBI) Protein database, using a BLASTP 2.2.25 algorithm with a conservative cut off expectation (*E*) score of less than or equal to 1.0.

As described in previous assessments, the *E* score is a measure of the probability that a particular alignment is due to random chance. When examining alignments between two protein sequences, a very low *E* score is more likely to reflect a true similarity, while a high *E* score is more likely to indicate a similarity by chance alone, therefore representing matches of less biological relevance. A statistically significant sequence similarity generally requires a match with an *E* score of less than 0.01 (Pearson 2000). By using a conservatively high cut off *E* score of 1.0, even proteins with more limited similarity will be captured for evaluation.

Sequence annotations of proteins with significant similarity to the GAT4621 protein sequence were manually inspected to identify proteins demonstrated to be toxic to humans or animals, or that otherwise raised a food safety concern, for example through the production of metabolites that could have an impact on the nutritional quality of the food or feed.

Within the parameters of the search, 797 proteins were identified with alignment similarity to the GAT4621 protein sequence. The highest scoring matches were to modified and native acetyltransferase sequences from the source organism, *Bacillus licheniformis*, followed by matches to similar sequences from related *Bacillus* species, including *B. subtilis*, *B. cereus* and *B. thuringiensis*. The majority of the remaining accessions represented both known and putative acetyltransferase proteins from various bacterial, archaeobacterial and eukaryotic species. Most of the protein matches returned were classified as GNAT or GCN5-like acetyltransferases, based upon the presence of conserved protein motifs. None of the protein sequences identified in the bioinformatics search raised a potential concern about the safety of the GAT4621 protein in GM crop plants.

The results of the overall homology search with the GAT4621 protein showed no similarity with known toxins; similarity only with other acetyltransferase proteins was revealed however none of these proteins is associated with any toxicity.

4.3.3 Stability to thermal treatment

Studies on the thermolability of a protein provide an indication of the stability of the protein under cooking or food processing conditions.

Studies submitted:

PHI-2006-184/018: Characterisation of the Thermal Stability of Glyphosate Acetyltransferase Enzyme Activity: GAT4621 (2007)

Aliquots of the GAT4621 protein produced in *E. coli* were heated for 15 minutes at a designated temperature ranging from 36 to 60 degrees (°C) in a gradient thermocycler. Enzymatic integrity was directly monitored with a continuous spectrophotometric enzyme activity assay, using glyphosate as substrate at saturating concentrations. Each assay was run in quadruplicate.

⁴ <http://www.ncbi.nlm.nih.gov/> - incorporating non-redundant entries from all GenBank and RefSeq nucleotide translations along with protein sequences from SWISS-PROT, PIR, PRF, and PDB

Enzyme activity was detected and converted to reaction rate in μM CoA per minute. The residual enzymatic activity remaining after the heat treatment was determined by comparing to the activity of non-heat treated enzyme.

The activity of the GAT4621 enzyme was reduced by half when incubated in the range of 49-50°C for 15 minutes. Enzyme activity levels following incubation for 15 minutes at temperatures above 53°C were reduced by an order of magnitude compared to the reaction rate with unheated enzyme. These results indicate that the functional activity of GAT4621 is significantly diminished with heating the protein at temperatures above approximately 50°C.

4.3.4 Acute oral toxicity studies

Studies submitted:

PHI-2005-110: GAT4621: Acute Oral Toxicity Study in Mice (2012). This study was originally submitted and assessed within Application A1021 in 2009.

An acute oral toxicity study in mice using *E.coli*-produced GAT4621 as the test substance was resubmitted by the Applicant. This study was previously considered by FSANZ within the assessment of herbicide tolerant corn line DP-098140-6 (Application A1021). The conclusion from this study was that the GAT4621 protein is not acutely toxic. The details of the study are not presented again in this report, however they are available in the safety assessment for Application A1021⁵, which is available on the FSANZ website.

While acute oral toxicity studies can be informative in certain cases, they are not absolutely necessary where the results of the biochemical, bioinformatics, digestibility and stability studies identify no safety concerns. In this case, the data assembled from these studies provide sufficient evidence that the newly expressed protein is not toxic.

4.4 Potential allergenicity of the newly expressed proteins

The potential of newly expressed proteins to be allergenic in humans is evaluated using an integrated, step-wise, case-by-case approach that relies on various criteria used in combination. This weight of evidence approach is used because no single criterion is sufficiently predictive of either allergenicity or non-allergenicity, and there are no reliable animal models for allergenicity assessment. Instead, 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 newly expressed protein is derived from a source known to be allergenic or has amino acid sequence similarity with a known allergen. In some cases, such as where the newly expressed protein has sequence similarity to a known allergen, additional *in vitro* and *in vivo* immunological testing may be warranted. Applying these criteria systematically provides reasonable evidence about the potential of the newly expressed protein to act as an allergen in humans.

The potential allergenicity of GAT4621 was assessed by consideration of:

- the source of the gene encoding the protein and history of use or exposure
- bioinformatics – a comparison of the amino acid sequence of the GAT4621 protein with that of known protein allergens
- the susceptibility of the *E. coli*-produced GAT4621 to *in vitro* digestion using simulated gastric and intestinal digestion models

⁵ <http://www.foodstandards.gov.au/code/applications/Pages/applicationa1021food4210.aspx>

4.4.1 Source of the protein

As described in Section 4.2, the GAT4621 protein is not derived from a bacterial source associated with allergenicity in humans.

4.4.2 Similarity to known allergens

Studies submitted:

PHI-2007-008/073: Comparison of the Amino Acid Sequence Identity Between the GAT4621 protein and Known Protein Allergens (2012)

Bioinformatics contributes to the weight of evidence approach for assessing potential allergenicity of novel proteins introduced to GM plants (Goodman, 2006; Thomas *et al.*, 2005). As with the bioinformatics analysis that looked at similarities with known protein toxins (refer to Section 4.3.2), this analysis compares the amino acid sequence of the newly expressed proteins with sequences of known allergens in order to identify primary structural similarities that may indicate a potential for cross-reactivity with allergenic proteins. The Codex guideline for the evaluation of potential allergenicity of newly expressed proteins indicates the possibility for cross-reactivity if the newly expressed protein is found to have at least 35% amino acid identity with an allergen over any segment of at least 80 amino acids (Codex, 2009). The FASTA35 (2009) sequence alignment tool was used for this purpose. A sliding window search of eight linear contiguous amino acids is also used to identify short peptides of possible immunological relevance in otherwise unrelated proteins.

The sequence of the GAT4621 protein was analysed against the FARRP12 database (Food Allergy Research and Resource Program, Release 12 – February 2012). This dataset contains 1603 sequences representing both known and putative food, environmental and contact allergens as well as proteins implicated in coeliac disease. No alignment of 35% identity over 80 amino acids was present, indicating that the GAT4621 protein does not share meaningful similarity with sequences in the allergen database. In addition, there were no matches of eight contiguous amino acids detected in the comparison of GAT4621 sequence to proteins in the allergen database. Taken together, these data show that the GAT4621 protein lacks both structurally and immunologically relevant similarities to known or putative allergens. The conclusion that GAT4621 is not likely to be allergenic in humans is consistent with the previous assessment conducted in 2009.

4.4.3 In vitro digestibility

Typically, food proteins that are allergenic tend to be stable to digestive enzymes such as pepsin and the acidic conditions of the stomach, and when presented to the intestinal mucosa, can lead to a variety of gastrointestinal and systemic manifestations of immune-mediated allergy in susceptible individuals (Astwood and Fuchs, 1996; Kimber *et al.*, 1999; Metcalfe *et al.*, 1996). As a consequence, one of the criteria for assessing potential allergenicity is to examine the stability of newly expressed proteins in conditions mimicking human digestion. Dietary proteins that are rapidly degraded to component amino acids in such conditions are considered less likely to be involved in eliciting an allergic response. Evidence of slow or limited protein digestibility however does not necessarily indicate that the protein is allergenic.

In vitro digestibility of GAT4621 was assessed in assays using simulated gastric fluid (SGF) containing pepsin, and simulated intestinal fluid (SIF) containing pancreatin (an enzyme mixture). The SGF assay protocol has been standardised based on results obtained from an international, multi-laboratory ring study published by Thomas *et al.* in 2004.

The published report demonstrated that the results of *in vitro* pepsin digestion assays were reproducible when a standard protocol was followed utilising a physiologically relevant acidic pH to simulate conditions in the stomach. The SIF study is considered less relevant because an ingested protein would first be exposed to pepsin-mediated hydrolysis in the acidic environment of the stomach, before being exposed to further digestion in the small intestine.

Studies submitted:

PHI-2006-120: Characterization of the *In Vitro* Pepsin Resistance of Glyphosate N-acetyltransferase 4621 Protein (GAT4621) (2007). This study was previously submitted with Application A1021.

PHI-2006-122: Characterization of the *In Vitro* Pancreatin Resistance of Glyphosate N-acetyltransferase 4621 Protein (GAT4621) (2007). This study was previously submitted with Application A1021.

SGF

The digestibility of the *E. coli*-produced GAT4621 protein (>95% purity on a total protein basis) was determined by analysing digestion mixtures incubated with SGF containing pepsin at 37°C for specific time intervals ranging from 30 seconds to 60 minutes. The pH of the reaction mixture was 1.2, which is highly acidic. A separate assay was performed under identical conditions using reaction mixtures with bovine serum albumin (BSA) and β -lactoglobulin as positive and negative control substances respectively. The products were analysed visually by protein staining following SDS-PAGE.

Visual examination of the gels showed that the GAT4621 protein was completely digested within 30 seconds of incubation in SGF. The presence of the intact GAT4621 protein band in the sample without pepsin indicates the protein was stable to the acid pH of the assay conditions and that the observed degradation over the time course was due to the activity of pepsin. These data support the conclusion that if ingested as a component of the diet, it is unlikely that intact GAT4621 protein would be in contact with the intestinal mucosa.

SIF

The digestibility of *E. coli*-derived GAT4621 protein in SIF containing pancreatin was analysed using SDS-PAGE and Western blot. Samples were incubated with SIF at pH 7.5 and 37°C, for specified time intervals (0, 0.5, 1, 2, 5, 10, 20, 30 and 60 minutes). Western blot analysis was included to clearly differentiate between protein bands from GAT4621 and the pancreatin enzyme mix. The control proteins (BSA and β -lactoglobulin) were incubated in SIF for 0, 1 and 60 minutes, and reaction products detected by protein staining.

The GAT4621 protein was hydrolysed in less than five minutes in SIF. No intact GAT4621 protein band was visible on the gel at the two minute timepoint. The results of the Western blot confirmed the observations from the protein gel. The Western blot showed that degradation of GAT4621 was almost complete at two minutes and complete by the five minute timepoint. The β -lactoglobulin positive control was also hydrolysed, with a faint band visible on a protein stained gel after one minute incubation, but no band visible after 60 minutes. The BSA negative control was only partially hydrolysed after 60 minutes.

Taken together, these studies demonstrate that the GAT4621 protein is readily susceptible to complete digestion in a simulated gastrointestinal system. These data also support the conclusion that if ingested as a component of the diet, it is unlikely that intact GAT4621 protein would be in contact with the intestinal mucosa.

4.5 Herbicide metabolites

For GM foods derived from crops that are herbicide tolerant, there are two issues that require consideration. The first is relevant to this safety assessment and involves assessment of any novel metabolites that are produced after the herbicide is applied to the crop plants to determine whether these are present in food products 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 issue, 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 pertaining to usage of the herbicide on the particular crop (both non-GM and GM) may have to be set. For imported commodities only, Australia and New Zealand can recognise residue limits set internationally by Codex or in certain cases adopt the limits set by the country in which the commodity is grown, subject to dietary exposure assessments for Australia and New Zealand.

Glyphosate is the ISO common name for 2-[(phosphonomethyl)amino] acetic acid (IUPAC) or *N*-(phosphonomethyl)glycine (CAS 1071-83-6). In Australia, glyphosate has been registered for use in a wide range of agricultural and food production systems by the Australian Pesticides and Veterinary Medicines Authority (APVMA), leading to over 60 entries in the Food Standards Code⁶. The existing MRL for use of glyphosate on rapeseed (canola) is 20 mg/kg (20 ppm), comprised of the sum of glyphosate and the metabolite aminomethylphosphonic acid (AMPA).

GAT4621 is a bacteria-derived enzyme that catalyzes the *N*-acetylation of glyphosate to the non-herbicidal compound *N*-acetylglyphosate (NAG), allowing a GM plant to function in the presence of the herbicide. Metabolism studies in GM corn and GM soybean expressing the *gat4621* gene demonstrated that *N*-acetylglyphosate was formed as the main metabolite. Glyphosate, *N*-acetyl aminomethylphosphonic acid (*N*-acetyl AMPA) and AMPA were found in low concentrations in the edible parts of the crops (FAO 2005). In herbicide-tolerant crops such as canola line 73496, residue data are needed to confirm the concentration of GM trait-specific metabolites, relative to the parent herbicide and its major metabolites on conventional crops.

4.5.1 Herbicide residues on canola line 73496

Study submitted:

PHI-2009-057: Magnitude and Decline of Glyphosate Related Residues in Forage and Seed of Genetically Modified Canola Event DP-073496-4 and Magnitude of Glyphosate Related Residues in Canola Event DP-073496-4 Seed Process Fractions Following Applications of Touchdown Total® Herbicide-Locations in the United States and Canada, Season 2009 (2010, revised 2013)

The commercial herbicide tested in this large study was Touchdown® Total Herbicide which consists of 36.5% glyphosate as the monopotassium salt. The active substance in this formulation is glyphosate. Due to the genetic modification, residue analysis quantified the parent compound, *N*-acetylglyphosate, *N*-acetyl AMPA and AMPA using liquid chromatography-tandem mass spectrometry (LC-MS/MS). The analytical method was validated down to a lower limit (LOD) of 0.02 ppm, with a limit of quantitation (LOQ) of 0.05 ppm for all commodities.

Applications of the herbicide were made to canola line 73496 grown at 24 field test sites across the United States and Canada in 2009 to determine residue levels in samples of

⁶ *Australia New Zealand Food Standards Code*, Standard 1.4.2 Maximum Residue Limits, Schedule 1

forage (used as animal feed), seed and processed commodities (refined oil and canola meal). The sites were typical of canola producing regions of those countries.

The Touchdown® Total herbicide labels are different for Canada and the United States. Both labels have a pre-emergent application and an application at about the 6-leaf growth stage. The Canadian label also allows for an application seven days before harvest. The application rates also differ between the two labels, and ammonium sulphate is added to the mixture for trials conducted according to the United States label.

The maximum labelled rate and timing for applications of the herbicide to canola line 73496 grown in Canada therefore allowed for three applications - one pre-emergent and two post-emergent applications in-crop. This treatment option was expected to provide the maximum residues under the proposed label rates. The final application to the Canadian field sites was at the rate of 900g active ingredient per hectare, as outlined in Table 4. Data from the U.S. field sites is not presented here as only two applications of herbicide were made on those crops (final application at the 6 leaf stage, 60-98 days before harvest) and in general the residues were lower overall.

Table 4: Applications of Glyphosate-based Herbicide to canola line 73496 in Canadian field trials (2009)

Appl. No.	Test substance	Rate (g a.e./hectare)	Appl. Method	Appl. Timing (± 1 DAY)	Spray Volume	Adjuvant
A1	Touchdown Total ®	675	Broadcast	Preemergent	50-100 L/ha	None
A2	Touchdown Total ®	675	Broadcast	At but no later than 6 leaf stage	50-100 L/ha	None
A3	Touchdown Total ®	900	Broadcast	7 days before harvest	50-100 L/ha	None

The results of the residue analysis in canola seed for each analyte (glyphosate and metabolites) in line 73496 are presented in Table 5. Each statistic was calculated from the individual residue values across all 16 Canadian test sites. The average total residues for seed and forage (glyphosate + N-acetylglyphosate + AMPA + N-acetyl AMPA) were 3.9 ppm and 0.49 ppm respectively. Total residues are expressed as glyphosate equivalents.

Table 5: Summary of Glyphosate residues in Canola line 73496 (Canadian field trials, 2009)

Seed Residues in ppm (mg/kg) ^a				
Statistic	Glyphosate	N-acetylglyphosate	AMPA	N-acetyl AMPA
Mean	2.5	1.3	0.03	0.04
Median	1.9	0.41	0.02	0.03
Minimum	0.29	0.21	ND	ND
Maximum	9.9	17	0.089	0.35

^a The designation 'ND' is used for a sample for which no peak was observed with a signal-to-noise ratio greater than 3:1

Seed samples were collected from separate plots at three sites where the last application of herbicide was made at five times the Canadian specifications, seven days before harvest.

Separate aliquots of each sample were processed by two different procedures – cold press and solvent extraction. Refined oil and meal were collected for residue analysis. Compared with the unprocessed seed, the results show a significant decrease of residues in refined oil regardless of the procedure used in the processing. Both procedures produced canola oil where total residues were below the limit of detection (<0.02 mg/kg) at all test sites.

4.5.2 Safety of herbicide metabolites

As outlined above, metabolism studies in corn and soybean crops expressing the *gat4621* gene following treatment with glyphosate demonstrated that *N*-acetylglyphosate is the predominant metabolite, and *N*-acetyl AMPA is a minor product. The toxicological profile of these two metabolites has been previously evaluated by FSANZ as part of Applications A1021 (corn) and A1006 (soybean).

N-acetylglyphosate and *N*-acetyl AMPA were concluded to be less toxic than glyphosate, which itself has low toxicity potential. On this basis, the establishment of a separate ADI for *N*-acetylglyphosate and *N*-acetyl AMPA was considered unnecessary. Previous assessments also concluded the current Australian ADI for glyphosate of 0.3 mg/kg bw remains appropriate for dietary risk assessment purposes.

Given that glyphosate is the only toxicologically-significant residue present on/in the seeds or grain from the GM commodities, its measurement in material derived from canola line 73496 would be considered adequate for safety assessment purposes. In addition to the absence of toxicological concerns, the studies on processed fractions derived from 73496 canola seeds have clearly demonstrated there would be no human exposure to glyphosate or its metabolites from the consumption of processed oil.

4.6 Conclusion

The newly expressed protein in canola line 73496 is GAT4621, an enzyme which confers tolerance to the herbicide glyphosate. The protein is expressed in canola seed at low levels. Herbicide residues and metabolites are also found in 73496 canola seed, however the amounts are low and do not raise a safety concern. In addition, there are no detectable levels of herbicide or metabolites in canola oil processed from this crop.

The identity and physicochemical and functional properties of the GAT4621 protein were examined in detail to confirm the expression in canola plants as intended. The characterisation studies confirmed that the GAT4621 protein in canola line 73496 conforms in size and amino acid sequence to that expected, does not exhibit any post-translational modification including glycosylation, and exhibits the expected enzyme activity.

The GAT4621 protein has been investigated previously for potential toxicity and allergenicity and was found to be innocuous. Bioinformatics studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens and *in vitro* digestibility studies demonstrated rapid degradation of GAT4621 in simulated gastric and intestinal fluids. The enzyme activity was destroyed by heating. Taken together, the evidence supports the conclusion that GAT4621 is not toxic, nor likely to be allergenic in humans, and would be as susceptible to normal digestive processes as other dietary proteins.

5. Compositional analysis

The 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 the genetic modification has resulted in a deliberate change to one or more nutrients in the food. In this case, canola line 73496 is herbicide tolerant and there was no intention to alter the nutrient composition of food derived from this plant line.

The focus of the compositional analysis is on 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 relevant 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/anti-nutrients) or quantitatively more minor constituents (minerals, vitamins). Key toxicants are those that have a level of toxicity and occur in amounts that may be significant to health (eg solanine in potatoes).

5.1 Key components

Canola oil is the primary food product used for human consumption. The key components to be analysed for a comparison between transgenic and conventional canola are outlined in the OECD Consensus Document on Compositional Considerations for New Varieties of Low Erucic Acid Rapeseed, and include proximates, amino acids, fatty acids (seed or oil), vitamins E and K, glucosinolates, tannins, sinapine and phytic acid (OECD, 2011).

5.2 Study design and conduct for key components

Studies submitted:

PHI-2009-039/020: Expressed Protein Concentration and Nutrient Composition of Canola Lines Containing Events DP-073496-4 and DP-061061-7: U.S. and Canada Test Sites (2010)

The compositional analyses were performed on seed from the F1^{*4} generation (see Figure 1) of canola line 73496. A conventional (non-GM) near isogenic line was used for the direct comparison of analytes.

The field production was conducted at six separate sites across commercial agricultural areas in Canada (four sites) and the United States (two sites). However, due to an early killing frost, canola seed at one of the Canadian test sites was unable to fully mature and was therefore excluded from the compositional analysis. Both canola 73496 and the control line were grown in each block at each site under typical agronomic conditions. Each plot of canola line 73496 was treated with glyphosate (411-464 grams acid equivalent per hectare); the control canola was left untreated.

Canola seed samples were harvested and analysed for key nutritional components: proximates (ash, crude protein, crude fat, crude fibre, carbohydrates by calculation), acid detergent fibre (ADF), neutral detergent fibre (NDF), amino acids (18), fatty acids (C8-C24), vitamins (B and tocopherols), minerals (calcium, copper, iron, magnesium, manganese, phosphorus, potassium, sodium and zinc), glucosinolates, secondary metabolites (phytosterols and tannins) and anti-nutrients (phytic acid and sinapine). Measurement of fibre components (ADF, NDF and total fibre) is of greater importance in considering the nutritional value of animal feed. Methods of compositional analysis were based on internationally recognised procedures (e.g. those of the Association of Official Analytical Chemists).

5.2.1 Statistical methods

A total of 99 different analytical components were measured. Components were statistically analysed by the Applicant using a mixed-model analysis of variance with the Statistical

Analysis Software⁷ (SAS) Version 9.2. Data from the five remaining sites were analysed separately and in combination (across-sites analysis). In order to complete the statistical analysis for any component in this study, the following inclusion/exclusion criteria were applied to the data. It was deemed that the across sites analysis was conducted if less than 80% of the values were below the LLOQ (lowest limit of quantitation). At a given site, if at least two samples were above the LLOQ for each analyte, then the individual site analysis was conducted. In addition, data were not included in the analysis if greater than 80% of samples for a single entry within the study were below the LLOQ.

Evaluating large numbers of analytes typically leads to a rate of false positive results, indicated by apparent statistically significant differences in the calculated means between groups. The number of false positives increases, as the number of analytes increases. To account for this, the Applicant applied the false discovery rate (FDR) method of Benjamini and Hochberg to their data (Benjamini and Hochberg 1995; Westfall *et al.* 1999). Accordingly, P-values were adjusted to hold the false positive rate to 5%. Both adjusted and non-adjusted P-values have been provided in their statistical analysis. In the discussion of results, a significant difference between the mean in canola line 73496 and that of the control line was established only when indicated by the FDR-adjusted P-value ($P \leq 0.05$).

For the statistical analysis, data from non-GM commercial varieties provided useful background information on the extent of natural variability for each component occurring in canola varieties that have a history of safe consumption. To generate the reference range, separate studies were conducted in which respectively three and four non-GM commercial canola lines were grown at five sites in North America. Plant material was harvested, processed, and analysed using methods similar to those used in the compositional studies on canola line 73496 and its comparator. Any analytes that showed a statistically significant difference between canola line 73496 and the isogenic control were also compared to the appropriate reference range, to determine whether further consideration of the potential impact on food safety would be warranted.

5.3 Analyses of key components

The results and discussion of the compositional analyses are presented below for key constituents.

5.3.1 Proximates and fibre

Results for the proximates and fibre analyses, on a dry weight basis, are shown in Table 6. Considering the adjusted P-values (see Section 5.2.1), statistically significant differences between canola line 73496 and the comparator were not observed for any of the analytes. The unadjusted P-value indicated an apparent significant difference in the mean crude protein level, however the magnitude of the difference was notably small. In addition, the lowest and highest values for crude protein from herbicide-treated canola line 73496 were entirely within a naturally broad range for commercial non-GM canola.

Table 6: Statistical Summary of Proximates and Fibre Content in Canola line 73496 and Control Seed – Across-site Analysis (ADF=acid detergent fibre; NDF=neutral detergent fibre)

⁷ SAS website - <http://www.sas.com/technologies/analytics/statistics/stat/index.html>

Analyte	Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval ^a	
Proximates and Fiber Composition (% Dry Weight)				
Crude Protein	Mean ^b	26.6	25.9	11.3 - 54.4
	Range ^c	23.4 - 31.9	22.4 - 31.6	
	CI ^d	23.4 - 30.3	22.8 - 29.5	
	P-Value ^e		0.0275	
	Adjusted P-Value ^f		0.0974	
Crude Fat	Mean	43.6	43.9	30.6 - 65.8
	Range	37.4 - 51.5	37.0 - 49.6	
	CI	40.1 - 47.3	40.3 - 47.7	
	P-Value		0.608	
	Adjusted P-Value		0.737	
ADF	Mean	32.4	32.3	14.5 - 63.4
	Range	27.1 - 38.6	26.6 - 39.5	
	CI	28.8 - 36.5	28.7 - 36.4	
	P-Value		0.929	
	Adjusted P-Value		0.957	
Crude Fiber	Mean	28.4	28.5	12.9 - 55.5
	Range	24.3 - 35.1	22.8 - 35.5	
	CI	25.1 - 32.1	25.2 - 32.2	
	P-Value		0.889	
	Adjusted P-Value		0.935	
NDF	Mean	33.4	33.3	15.5 - 67.7
	Range	28.8 - 40.3	28.0 - 38.9	
	CI	30.4 - 36.8	30.2 - 36.7	
	P-Value		0.844	
	Adjusted P-Value		0.911	
Ash	Mean	4.08	3.92	2.01 - 7.63
	Range	2.98 - 5.10	3.00 - 5.06	
	CI	3.29 - 5.07	3.16 - 4.87	
	P-Value		0.0549	
	Adjusted P-Value		0.141	
Carbohydrates	Mean	25.3	26.0	19.6 - 33.5
	Range	16.9 - 29.1	23.7 - 28.8	
	CI	24.0 - 26.7	24.6 - 27.4	
	P-Value		0.320	
	Adjusted P-Value		0.468	

^a The statistical tolerance interval is referred to as the reference range because it uses data from separate field trials. The statistical tolerance interval was calculated from commercial canola varieties, calculated to contain with 95% confidence, 99% of the population of canola; negative limits set to zero.

^b Least squares mean were generated from a mixed model analysis.

^c Range denotes the lowest and highest individual value across sites.

^d Confidence Interval

^e Non-adjusted P-value

^f False Discovery Rate (FDR) adjusted P-value

5.3.2 Fatty Acids

The major food from canola plants is the edible seed oil, therefore fatty acid constituents are particularly relevant for safety assessment.

The levels of 30 fatty acids (C8 – C24) were measured in this study, however 12 fatty acids were excluded from the statistical analysis because more than 80% of the samples (control and/or glyphosate-treated canola line 73496) were below the LLOQ and could not be reliably measured. The levels of fatty acids in seed from canola line 73496 and the comparator are

reported in Table 7. Measurements of heptadecanoic acid (17:0), γ -linolenic acid (18:3) and erucic acid (22:1) were excluded from the Table because they occurred at levels less than 0.05% of total fatty acids, and were deemed 'non-detectable'. The levels of erucic acid were expected to be non-detectable in any case as per the definition of canola oil as 'low erucic acid rapeseed oil', and this was confirmed.

Table 7: Statistical Summary of Fatty Acids in Canola line 73496 and Control Seed – Across-site Analysis

Analyte		Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
Fatty Acids Composition (% Total Fatty Acids)				
Lauric Acid (C12:0)	Mean	0.0816	0.0811	0 - 0.216
	Range	0.0568 - 0.105	0.0563 - 0.113	
	CI	0.0639 - 0.104	0.0634 - 0.104	
	P-Value		0.863	
	Adjusted P-Value		0.919	
Myristic Acid (C14:0)	Mean	0.0530	0.0546	0.0239 - 0.112
	Range	0.0470 - 0.0749	0.0484 - 0.0670	
	CI	0.0466 - 0.0603	0.0480 - 0.0621	
	P-Value		0.212	
	Adjusted P-Value		0.355	
Palmitic Acid (C16:0)	Mean	4.08	4.19	3.10 - 5.60
	Range	3.72 - 4.40	3.85 - 4.45	
	CI	3.87 - 4.31	3.97 - 4.42	
	P-Value		0.0111	
	Adjusted P-Value		0.0605	
Palmitoleic Acid (C16:1)	Mean	0.262	0.272	0.154 - 0.378
	Range	0.225 - 0.306	0.224 - 0.317	
	CI	0.232 - 0.295	0.241 - 0.307	
	P-Value		0.0173	
	Adjusted P-Value		0.0790	
Heptadecenoic Acid (C17:1)	Mean	0.108	0.105	0 - 0.264
	Range	0.0675 - 0.140	0.0666 - 0.137	
	CI	0.0802 - 0.146	0.0776 - 0.141	
	P-Value		0.0580	
	Adjusted P-Value		0.144	
Stearic Acid (C18:0)	Mean	1.71	1.61	1.17 - 3.47
	Range	1.47 - 1.89	1.44 - 1.93	
	CI	1.57 - 1.86	1.48 - 1.76	
	P-Value		0.00577	
	Adjusted P-Value		0.0507	
Oleic Acid (C18:1)	Mean	61.6	60.1	53.1 - 74.4
	Range	58.6 - 64.1	56.8 - 63.3	
	CI	58.9 - 64.4	57.5 - 62.9	
	P-Value		<0.0001	
	Adjusted P-Value		0.00216	
Linoleic Acid (C18:2)	Mean	20.1	21.3	13.1 - 27.1
	Range	18.5 - 22.5	19.5 - 23.4	
	CI	18.5 - 21.8	19.5 - 23.1	
	P-Value		0.000107	
	Adjusted P-Value		0.00220	

Table 7 (continued): Statistical Summary of Fatty Acids in Canola line 73496 and Control Seed – Across-site Analysis

Analyte		Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
(9,15) Isomer of Linoleic Acid (C18:2)	Mean	0.0986	0.0924	0 - 0.524 ^a
	Range	0.0626 - 1.45	0.0611 - 0.496	
	CI	0.0697 - 0.139	0.0653 - 0.131	
	P-Value		0.764	
	Adjusted P-Value		0.868	
Linolenic Acid (C18:3)	Mean	8.92	9.29	4.23 - 16.8
	Range	7.46 - 10.0	8.08 - 10.5	
	CI	8.04 - 9.89	8.38 - 10.3	
	P-Value		0.00884	
	Adjusted P-Value		0.0525	
Arachidic Acid (C20:0)	Mean	0.615	0.586	0.496 - 1.01
	Range	0.585 - 0.645	0.552 - 0.640	
	CI	0.598 - 0.632	0.570 - 0.602	
	P-Value		0.00618	
	Adjusted P-Value		0.0507	
Eicosenoic Acid (C20:1)	Mean	1.38	1.36	0.999 - 1.88
	Range	1.19 - 1.54	1.24 - 1.45	
	CI	1.27 - 1.51	1.25 - 1.48	
	P-Value		0.128	
	Adjusted P-Value		0.269	
Eicosadienoic Acid (C20:2)	Mean	0.0690	0.0713	0.0308 - 0.107
	Range	0.0565 - 0.0805	0.0563 - 0.0831	
	CI	0.0583 - 0.0817	0.0602 - 0.0844	
	P-Value		0.0256	
	Adjusted P-Value		0.0974	
Behenic Acid (C22:0)	Mean	0.328	0.319	0.241 - 0.465
	Range	0.290 - 0.365	0.286 - 0.359	
	CI	0.300 - 0.358	0.292 - 0.348	
	P-Value		0.0182	
	Adjusted P-Value		0.0790	
Lignoceric Acid (C24:0)	Mean	0.178	0.170	0.0601 - 0.657
	Range	0.160 - 0.208	0.151 - 0.206	
	CI	0.160 - 0.198	0.153 - 0.189	
	P-Value		0.0183	
	Adjusted P-Value		0.0790	
Nervonic Acid (C24:1)	Mean	0.191	0.194	0.0295 - 0.542
	Range	0.121 - 0.257	0.120 - 0.254	
	CI	0.141 - 0.259	0.143 - 0.263	
	P-Value		0.303	
	Adjusted P-Value		0.458	

^a minimum and maximum values used because insufficient values above the LLOQ to calculate a tolerance interval.

No statistically significant differences were observed in the across-sites analysis between canola line 73496 and the comparator in the following 16 fatty acid analytes: 12:0 lauric acid, 14:0 myristic acid, 16:0 palmitic acid, 16:1 palmitoleic acid, 17:0 heptadecanoic acid, 17:1 heptadecenoic acid, 18:0 stearic acid, 18:2 (9,15) isomer linoleic acid, 18:3 γ -linolenic acid, 20:0 arachidic acid, 20:1 eicosenoic acid, 20:2 eicosadienoic acid, 22:0 behenic acid, 23:0 tricosanoic acid, 24:0 lignoceric acid and 24:1 nervonic acid.

A statistically significant difference between canola line 73496 and the control was observed in the levels of two fatty acids, oleic acid (18:1) and linoleic acid (18:2), in the across-sites analysis, as identified by the adjusted P-value ≤ 0.05 . From the data provided in Table 7, it

can be seen that the difference in absolute terms was small for both analytes. In addition, considering the individual site data, significant differences between the transgenic line and its comparator were not observed at all trial sites, which is a likely indication of normal variation due to the environment rather than a discernible pattern of change attributable to the genetic modification. The mean values for these analytes were within the range established for non-GM commercial reference varieties grown under similar conditions.

Overall, the observed differences do not result in any impact on the nutritional quality of the oil from canola line 73496, when compared with commercial non-GM canola varieties.

5.3.3 Amino Acids

Lipids and protein are quantitatively the most important fractions in canola seed, accounting for more than 60% of the weight (OECD, 2011). After the oil has been extracted from the seed, the canola meal is a good source of protein for animal feed. Based on international data, the oil free meal is approximately 40% protein (by weight).

The levels of 18 amino acids were measured in seed from canola line 73496 for comparison with corresponding measurements in the control canola. As can be seen from the results presented in Table 8, there were no statistically significant differences between the transgenic and isogenic control lines for any of the amino acid analytes.

5.3.4 Minerals

Levels of nine key minerals were measured in seeds from canola line 73496 and the non-GM control, as presented in Table 9. In the across-sites analysis, the mean levels of phosphorus, calcium, manganese, copper, iron, potassium, sodium and zinc were similar in both GM and non-GM lines. The mean level of magnesium was marginally lower in herbicide-treated canola line 73496 than in the control line. Although the difference was statistically significant, in absolute terms the magnitude of the difference was very small, and was not observed at all individual sites. The mean values for all key minerals in canola line 73496 observed in the across-sites analysis were within the reported reference range established by conventional commercial varieties grown under similar conditions. These results indicate no significant nutritional differences between canola line 73496, the non-GM comparator and commercial canola in the levels of key minerals.

5.3.5 Vitamins

Eleven analytes were measured to compare the vitamin composition of canola line 73496 with the non-GM control line, as shown in Table 10. The analytes included vitamins B1, B2, B3, B5, B6, B9 and a number of tocopherol compounds (vitamin E). The across-sites data showed a very small statistically significant increase in δ -tocopherol and total tocopherols in canola line 73496 compared with the control. All values for vitamin analytes in the transgenic canola were within the reference range for convention canola. The statistically significant difference in δ -tocopherol and total tocopherols between the transgenic and control lines is not of any nutritional significance.

Table 8: Statistical Summary of Amino Acids in Canola line 73496 and Control – Across-sites Analysis

Analyte	Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
Amino Acids Composition (% Dry Weight)			
Alanine	Mean	1.14	1.12
	Range	1.01 - 1.37	0.970 - 1.32
	CI	1.02 - 1.28	0.996 - 1.25
	P-Value		0.0446
	Adjusted P-Value		0.133
Arginine	Mean	1.62	1.57
	Range	1.30 - 1.98	1.32 - 1.92
	CI	1.42 - 1.85	1.37 - 1.79
	P-Value		0.0502
	Adjusted P-Value		0.133
Aspartic Acid	Mean	2.00	2.06
	Range	1.26 - 2.54	1.77 - 2.50
	CI	1.71 - 2.33	1.77 - 2.40
	P-Value		0.152
	Adjusted P-Value		0.312
Cysteine	Mean	0.606	0.618
	Range	0.505 - 0.751	0.523 - 0.722
	CI	0.516 - 0.712	0.526 - 0.726
	P-Value		0.592
	Adjusted P-Value		0.737
Glycine	Mean	1.38	1.35
	Range	1.09 - 1.61	1.17 - 1.54
	CI	1.25 - 1.52	1.22 - 1.49
	P-Value		0.162
	Adjusted P-Value		0.317
Glutamic Acid	Mean	5.05	5.01
	Range	2.48 - 6.68	4.30 - 6.40
	CI	4.23 - 6.02	4.20 - 5.97
	P-Value		0.783
	Adjusted P-Value		0.868
Histidine	Mean	0.800	0.801
	Range	0.644 - 0.966	0.667 - 0.939
	CI	0.712 - 0.899	0.713 - 0.900
	P-Value		0.945
	Adjusted P-Value		0.957
Isoleucine	Mean	1.08	1.06
	Range	0.869 - 1.30	0.922 - 1.25
	CI	0.958 - 1.22	0.935 - 1.19
	P-Value		0.113
	Adjusted P-Value		0.250
Leucine	Mean	1.88	1.83
	Range	1.52 - 2.26	1.59 - 2.19
	CI	1.66 - 2.12	1.62 - 2.07
	P-Value		0.0498
	Adjusted P-Value		0.133

Table 8 (continued): Statistical Summary of Amino Acids in Canola line 73496 and Control – Across-sites Analysis

Analyte		Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
Lysine	Mean	1.65	1.64	0.729 - 3.16
	Range	1.25 - 2.02	1.48 - 1.97	
	CI	1.45 - 1.87	1.45 - 1.86	
	P-Value		0.669	
	Adjusted P-Value		0.793	
Methionine	Mean	0.464	0.472	0.258 - 0.857
	Range	0.383 - 0.546	0.402 - 0.546	
	CI	0.408 - 0.529	0.415 - 0.538	
	P-Value		0.596	
	Adjusted P-Value		0.737	
Phenylalanine	Mean	1.12	1.10	0.545 - 2.12
	Range	0.901 - 1.35	0.927 - 1.28	
	CI	1.01 - 1.26	0.985 - 1.23	
	P-Value		0.167	
	Adjusted P-Value		0.319	
Proline	Mean	1.63	1.59	0.717 - 3.36
	Range	1.46 - 1.98	1.38 - 1.95	
	CI	1.43 - 1.87	1.39 - 1.83	
	P-Value		0.123	
	Adjusted P-Value		0.265	
Serine	Mean	1.12	1.12	0.576 - 2.08
	Range	0.719 - 1.34	0.985 - 1.31	
	CI	1.01 - 1.25	1.00 - 1.25	
	P-Value		0.971	
	Adjusted P-Value		0.971	
Threonine	Mean	1.11	1.11	0.619 - 1.95
	Range	0.847 - 1.29	0.997 - 1.26	
	CI	1.02 - 1.22	1.01 - 1.21	
	P-Value		0.687	
	Adjusted P-Value		0.793	
Tryptophan	Mean	0.325	0.312	0.153 - 0.537
	Range	0.242 - 0.429	0.236 - 0.427	
	CI	0.256 - 0.382	0.239 - 0.371	
	P-Value		0.212	
	Adjusted P-Value		0.355	
Tyrosine	Mean	0.635	0.620	0.337 - 1.20
	Range	0.552 - 0.757	0.508 - 0.737	
	CI	0.573 - 0.704	0.560 - 0.688	
	P-Value		0.205	
	Adjusted P-Value		0.355	
Valine	Mean	1.39	1.36	0.682 - 2.49
	Range	1.04 - 1.66	1.20 - 1.59	
	CI	1.23 - 1.57	1.20 - 1.54	
	P-Value		0.159	
	Adjusted P-Value		0.317	

Table 9: Statistical Summary of key Minerals in Canola line 73496 and Control – Across-sites Analysis

Analyte	Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
Minerals Composition (% Dry Weight)			
Calcium	Mean	0.465	0.474
	Range	0.365 - 0.576	0.397 - 0.572
	CI	0.406 - 0.532	0.414 - 0.542
	P-Value		0.457
	Adjusted P-Value		0.625
Phosphorus	Mean	0.788	0.756
	Range	0.543 - 1.13	0.567 - 1.14
	CI	0.608 - 1.02	0.583 - 0.980
	P-Value		0.0285
	Adjusted P-Value		0.0974
Magnesium	Mean	0.357	0.325
	Range	0.308 - 0.410	0.277 - 0.385
	CI	0.327 - 0.390	0.298 - 0.356
	P-Value		<0.0001
	Adjusted P-Value		0.00216 ^a
Manganese	Mean	0.00381	0.00351
	Range	0.00275 - 0.00463	0.00278 - 0.00418
	CI	0.00318 - 0.00458	0.00292 - 0.00422
	P-Value		0.0334
	Adjusted P-Value		0.110
Copper	Mean	0.000175	0.000185
	Range	<0.000125 ^b - 0.000251	<0.000125 ^b - 0.000276
	CI	0.000129 - 0.000236	0.000137 - 0.000250
	P-Value		0.0972
	Adjusted P-Value		0.221
Iron	Mean	0.00677	0.00634
	Range	0.00452 - 0.0200	0.00412 - 0.0116
	CI	0.00427 - 0.0107	0.00400 - 0.0100
	P-Value		0.288
	Adjusted P-Value		0.454
Potassium	Mean	0.649	0.659
	Range	0.489 - 0.866	0.487 - 0.898
	CI	0.525 - 0.803	0.533 - 0.816
	P-Value		0.401
	Adjusted P-Value		0.567
Sodium	Mean	0.00168	0.00179
	Range	0.000568 - 0.00599	0.000849 - 0.00505
	CI	0.000790 - 0.00358	0.000840 - 0.00381
	P-Value		0.507
	Adjusted P-Value		0.671
Zinc	Mean	0.00398	0.00384
	Range	0.00277 - 0.00548	0.00233 - 0.00594
	CI	0.00317 - 0.00498	0.00306 - 0.00481
	P-Value		0.288
	Adjusted P-Value		0.454

^a Statistically significant difference, FDR adjusted P-Value <0.05.

^b Analyte values were below the assay LLOQ. Sample results below the LLOQ were assigned a value equal to the LLOQ.

Table 10: Statistical Summary of Vitamin B and tocopherol levels in Canola line 73496 and Control – Across-sites Analysis

Analyte		Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
Vitamins Composition (mg/kg Dry Weight)				
Vitamin B1 (Thiamine)	Mean	14.2	14.8	4.92 - 33.1
	Range	9.65 - 21.1	10.6 - 25.8	
	CI	11.8 - 17.2	12.2 - 17.8	
	P-Value		0.611	
	Adjusted P-Value		0.737	
Vitamin B2 (Riboflavin)	Mean	2.70	3.14	1.45 - 7.33
	Range	2.04 - 3.98	2.27 - 5.88	
	CI	2.29 - 3.19	2.66 - 3.71	
	P-Value		0.00531	
	Adjusted P-Value		0.0507	
Vitamin B3 (Niacin)	Mean	188	180	61.0 - 444
	Range	156 - 245	154 - 208	
	CI	179 - 197	171 - 189	
	P-Value		0.188	
	Adjusted P-Value		0.338	
Vitamin B5 (Pantothenic Acid)	Mean	4.74	4.56	3.52 - 9.88
	Range	4.08 - 5.77	3.94 - 5.05	
	CI	4.31 - 5.21	4.15 - 5.01	
	P-Value		0.0273	
	Adjusted P-Value		0.0974	
Vitamin B6 (Pyridoxine)	Mean	3.69	3.39	1.17 - 20.9
	Range	2.58 - 6.01	2.54 - 5.25	
	CI	2.94 - 4.64	2.70 - 4.26	
	P-Value		0.0670	
	Adjusted P-Value		0.162	
Vitamin B9 (Folic Acid)	Mean	2.89	2.63	0.769 - 5.27
	Range	1.57 - 7.88	1.12 - 5.68	
	CI	2.05 - 4.06	1.87 - 3.69	
	P-Value		0.307	
	Adjusted P-Value		0.458	
α -Tocopherol	Mean	109	114	54.7 - 189
	Range	96.5 - 130	101 - 129	
	CI	98.8 - 120	103 - 125	
	P-Value		0.0179	
	Adjusted P-Value		0.0790	
β -Tocopherol	Mean	<1.25 ^a	<1.25 ^a	0 - 1.67 ^b
	Range	<1.25 ^a	<1.25 ^a	
	CI	NA	NA	
	P-Value		NA	
	Adjusted P-Value		NA	
δ -Tocopherol	Mean	3.23	3.65	1.93 - 11.9
	Range	2.80 - 4.04	3.25 - 3.99	
	CI	3.11 - 3.35	3.51 - 3.78	
	P-Value		<0.0001	
	Adjusted P-Value		0.00216 ^c	
γ -Tocopherol	Mean	170	177	106 - 378
	Range	143 - 196	146 - 205	
	CI	155 - 187	161 - 195	
	P-Value		0.00770	
	Adjusted P-Value		0.0525	
Total Tocopherols	Mean	283	295	191 - 499
	Range	243 - 313	250 - 332	
	CI	262 - 305	273 - 318	
	P-Value		0.00397	
	Adjusted P-Value		0.0465 ^c	

^a Analyte values were below the assay LLOQ. Sample results below the LLOQ were assigned a value equal to the LLOQ.

^b Tolerance interval was not available due to insufficient sample values being detected above the assay LLOQ. Minimum and maximum values are used in place of a tolerance interval.

^c Statistically significant difference, FDR adjusted P-Value <0.05.

5.3.6 Natural toxicants

Glucosinolates are key natural toxicants of rapeseed (OECD, 2011). The major compounds are 3-butenyl glucosinolate (gluconapin), 4-pentenyl glucosinolate (glucobrassicinapin), 2-hydroxy-3-butenyl glucosinolate (progoitrin) and 2-hydroxy-4-pentenyl glucosinolate (napoleiferin), however a number of other glucosinolates occur at measurable levels in canola seed.

While glucosinolates themselves are considered to be innocuous, the hydrolysis products have negative effects on animal production, particularly on thyroid activity. Because of these effects combined with low palatability in stockfeed, varieties of rapeseed were developed with consistently low concentrations of both glucosinolates (in seed meal) and erucic acid (in the oil fraction). These so-called low erucic acid-low glucosinolate varieties now define the tradename product known as *canola*. The total glucosinolate concentration in canola line 73496 seed was 5.7 μ moles/g dry weight, which is well within the accepted specifications for canola seed, which is a maximum of 30 μ moles glucosinolates (any one compound or any mixture) per gram of seed on a dry weight basis (OECD, 2011).

The levels of total glucosinolates in herbicide-treated canola line 73496 and the control line were not significantly different, as shown below in Table 11. These results support the conclusion that neither the presence of the introduced gene, nor the application spraying of glyphosate herbicide, affected the levels of glucosinolates in canola line 73496.

Table 11: Statistical Summary of Total Glucosinolates in Canola line 73496 and Control Seeds – Across-sites Analysis

Analyte		Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
Total Glucosinolates	Mean	5.77	5.66	2.17 - 30.0
	Range	2.61 - 10.5	2.77 - 13.5	
	CI	3.72 - 8.94	3.65 - 8.77	
	P-Value		0.686	
	Adjusted P-Value		0.793	

5.3.7 Anti-nutrients

The key antinutrients in canola include tannins, sinapine and phytic acid. Sinapine, a major phenolic compound in canola, imparts a bitter taste to canola meal. Phytic acid (phytate in the salt form) is a major form of storage of phosphorus in plants. Monogastric animals however lack the digestive enzyme phytase which is required to use phytate as a nutrient source. Phytic acid also has strong binding affinity to important minerals such as calcium, magnesium, iron and zinc, thus reducing the absorption of these minerals (OECD, 2011).

The results of analyses for tannins, phytic acid and sinapine in canola line 73496 and the comparator line are presented in Table 12. Based on the adjusted P-value, no significant differences were observed in the levels of these compounds in the two lines. In addition, the levels were within the range established from conventional canola varieties grown under similar conditions.

Table 12: Statistical Summary of Anti-Nutrients in Canola line 73496 and the Non-GM Control – Across-sites Analysis

Analyte		Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval
Secondary Metabolites and Anti-Nutrients Composition (% Dry Weight)				
Tannins-Soluble	Mean	0.113	0.109	0.0521 - 0.297
	Range	0.0768 - 0.180	0.0777 - 0.176	
	CI	0.0840 - 0.152	0.0812 - 0.147	
	P-Value		0.415	
	Adjusted P-Value		0.577	
Tannins-Insoluble	Mean	0.330	0.237	0.0731 - 2.32
	Range	0.155 - 0.824	0.142 - 0.492	
	CI	0.218 - 0.499	0.157 - 0.359	
	P-Value		0.00874	
	Adjusted P-Value		0.0525	
Phytic Acid	Mean	1.84	1.73	0.684 - 6.06
	Range	1.04 - 2.86	1.01 - 3.00	
	CI	1.27 - 2.68	1.19 - 2.52	
	P-Value		0.0401	
	Adjusted P-Value		0.126	
Sinapine	Mean	0.928	0.937	0.538 - 1.33
	Range	0.763 - 1.12	0.804 - 1.09	
	CI	0.852 - 1.01	0.861 - 1.02	
	P-Value		0.558	
	Adjusted P-Value		0.727	

5.3.8 Other constituents

Studies submitted:

PHI-2010-018/020: *N*-Acetylaspartate and *N*-Acetylglutamate Concentrations in Seed of an Herbicide-Treated Canola Line Containing Event DP 073496 4: U.S. and Canada Test Sites (2010)

PHI-2010-019/020: *N*-Acetylglycine, *N*-Acetylserine and *N*-Acetylthreonine Concentrations in Seed of an Herbicide-Treated Canola Line Containing Event DP 073496 4: U.S. and Canada Test Sites.

PHI-2010-145/020: Concentration of *N*-Acetylaspartate and *N*-Acetylglutamate in Whole Plant Tissues Derived from a Canola Line Containing Event DP 073496-4

PHI-2010-155/020: Concentration of *N*-Acetylglycine, *N*-Acetylserine and *N*-Acetylthreonine in Whole Plant Tissues Derived from a Canola Line Containing Event DP-073496-4

PHI-2010-107/020: Concentration of *N*-Acetylaspartate and *N*-Acetylglutamate in Processed Products from Seed of a Canola Line Containing Event DP-073496-4: U.S. and Canada Test Sites (2010)

PHI-2010-108/020: Concentration of *N*-Acetylglycine, *N*-Acetylserine and *N*-Acetylthreonine in Processed Products from Seed of a Canola Line Containing Event DP-073496-4: U.S. and Canada Test Sites.

The activity of the GAT4621 enzyme has been optimised to catalyse the acetylation of glyphosate, which produces the herbicidally inactive compound *N*-acetylglyphosate. As described in the assessments of GAT corn line DP-098140-6 (A1021) and GAT soybean line DP-356043-5 (A1006), FSANZ has previously considered whether the enzyme catalyses the acetylation of substrates other than glyphosate in the plant, such as the amino acids L-aspartate and L-glutamate.

5.3.8.1 Previous assessments

A previous assessment (A1021) summarised a study conducted to determine the substrate specificity of the GAT4621 enzyme, using a range of amino acids available in plant cells. The

enzyme assay identified catalysis of only five amino acid substrates (above the limit of quantification of the assay): L-aspartate, L-glutamate, L-serine, L-threonine, and glycine. However, the catalytic efficiency (k_{cat}/K_M) of GAT4621 on L-aspartate, L-glutamate, L-serine, and L-threonine was only about 1%, 0.8%, 0.05% and 0.06%, respectively, of the activity on glyphosate. The affinity of the GAT4621 enzyme for glycine was too low for estimation of K_M , and the catalytic efficiency could therefore not be calculated in this study.

The levels of the acetylated amino acids N-acetylglutamate (NAGlu), N-acetylaspartate (NAAsp), N-acetylthreonine (NAThr), N-acetylserine (NASer) and N-acetylglycine (NAGly) were subsequently measured in corn grain from herbicide-tolerant and non-GM control plants. Importantly, NAAsp, NAGlu, NAThr, NASer, and NAGly were found in the non-GM control corn grain, indicating that these metabolites are not novel in the GM plants.

Despite the low catalytic efficiency overall, the assessment of corn line DP-098140-6 found mean values for NAAsp, NAGlu, NAThr, NASer, and NAGly in the GM grain were significantly higher ($p < 0.05$) than those of control corn grain. However, the total level of the acetylated amino acids was low, at only 0.05% on a dry weight basis. As individual compounds, the concentrations of NAThr, NASer, and NAGly were less than 0.0003% each on a mean dry weight basis. These measurements correlate with the higher catalytic efficiency of GAT4621 on L-aspartate and L-glutamate, as determined in the substrate specificity study.

5.3.8.2 Acetylated amino acids in canola

The levels of the five acetylated amino acids were measured in seed and whole plant samples from canola line 73496 and control canola for consideration of any consequential impact on food safety. The canola was grown at five sites located in commercial canola growing regions of North America, which included three sites in Canada. Four seed samples from each site were analysed, and the results are presented in Table 13.

The two most abundant acetylated amino acids in seed from canola line 73496 were NAAsp (1480 $\mu\text{g/g}$ dry weight) and NAGlu (33 $\mu\text{g/g}$ dry weight), which were both many fold higher than in the control canola seed, and also well outside a tolerance interval set for commercial non-GM canola varieties. The mean levels of NASer and NAThr were also statistically significantly higher in canola line 73496 compared with the control, however the levels were within the tolerance interval for commercial non-GM canola. There was no significant difference in the level of NAGly in canola line 73496 and the non-GM control.

The results of the acetylated amino acid analyses in whole plant samples were similar to the seed results. That is, NAAsp and NAGlu showed large increases in the amounts present in herbicide-treated canola line 73496 compared with the non-GM control. The data provided on the whole plant samples was not considered to be as relevant for food safety assessment as the seed data and therefore the results have not been included in this report.

Table 13: Mean Concentrations of Acetylated Amino Acids in Canola Seed Grown in Canada and the U.S. – Statistical Summary

Analyte		Control Canola n=20	Herbicide-Treated 73496 Canola n=20	Tolerance Interval ^a
µg/g Dry Weight				
NAAsp	Mean ^b	1.24	1480	0.00861 - 4.43
	Range ^c	0.377 - 5.39	1200 – 1770	
	CI ^d	0 ^d - 9.38	1340 – 1640	
	P-Value ^e		<0.0001	
	Adjusted P-Value ^f		<0.0001 ^g	
NAGlu	Mean	0.628	32.8	0.0968 - 5.37
	Range	0.428 - 1.46	20.3 - 61.1	
	CI	0.00000752 - 2.50	24.4 - 42.5	
	P-Value		<0.0001	
	Adjusted P-Value		<0.0001 ^h	
NAGly	Mean	0.0751	0.0825	0.0240 - 0.338
	Range	0.0481 - 0.125	0.0424 - 0.182	
	CI	0.0540 - 0.105	0.0592 - 0.115	
	P-Value		0.454	
	Adjusted P-Value		0.454	
NASer	Mean	0.843	1.04	0.0524 - 27.2
	Range	0.389 - 3.05	0.491 - 3.55	
	CI	0.437 - 1.63	0.542 - 2.01	
	P-Value		0.0035	
	Adjusted P-Value		0.00528 ^h	
NAThre	Mean	0.110	0.546	0.0140 - 1.74
	Range	0.0531 - 0.212	0.260 - 1.64	
	CI	0.0665 - 0.181	0.331 - 0.902	
	P-Value		<0.0001	
	Adjusted P-Value		<0.0001 ^h	

^a The statistical tolerance interval was calculated from commercial canola varieties and calculated to contain with 95 percent confidence, 99 percent of the population of canola, negative limits set to zero

^b Least squares mean were generated from a mixed model analysis

^c Range denotes the lowest and highest individual value across sites

^d Confidence Interval

^e Non-adjusted P-value

^f False Discovery Rate (FDR) adjusted P-value

^g The lower limit of the confidence interval was negative on the transformed scale and was set to zero prior to back-transformation

^h Statistically significant difference, FDR adjusted P-Value <0.05.

5.3.8.3 Dietary exposure to acetylated amino acids

Whole canola seed is not commonly consumed. Typical processing of seed produces canola meal, mostly for livestock feed, and oil, which is widely consumed as food. A study was conducted to compare levels of the same five acetylated amino acids in processed canola fractions derived from the herbicide-treated GM and control lines.

As expected from the seed analyses, NAAsp and NAGlu were the most abundant acetylated amino acids, and the levels in processed canola meal from line 73496 were three orders of magnitude higher than in the processed meal fraction from control canola. As indicated in Table 14, no statistical analysis was performed but the difference for these two compounds was unequivocal.

The analysis of refined, bleached and deodorised (RBD) oil confirmed that the oil fraction does not contain plant proteins; all acetylated amino acids were not quantifiable. Dietary exposure to acetylated amino acids from the consumption of oil derived from canola line 73496 therefore would be effectively zero.

Table 14: Concentrations of Acetylated Amino Acids in Processed Fractions Produced from Canola Line 73496 and Control Canola

Fraction	Canola Line	Concentration (Range) n=2 ^a				
		NAA ^b	NAG	NAT	NAS	NAGly
µg/g ^c						
Toasted Meal (with hulls)	73496	2872 (2822 - 2921)	62.3 (57.7 - 66.8)	0.585 (0.558 - 0.612)	1.20 (1.11 - 1.29)	0.151 (0.145 - 0.157)
	Control	2.31 (2.15 - 2.47)	1.36 (1.31 - 1.41)	0.161 (0.146 - 0.175)	1.46 (1.39 - 1.53)	0.129 (0.114 - 0.144)
Toasted Meal (without hulls)	73496	3013 (3002 - 3024)	66.0 (62.6 - 69.3)	0.788 (0.535 - 1.04)	2.27 (1.98 - 2.55)	0.224 (0.184 - 0.265)
	Control	2.58 (2.00 - 3.16)	1.71 (1.63 - 1.80)	0.226 (0.195 - 0.258)	1.76 (1.51 - 2.01)	0.180 (0.163 - 0.197)
RBD Oil (with hulls)	73496	ND ^d	ND - <LLOQ	ND - <LLOQ	ND	ND - <LLOQ
	Control	ND	ND	<LLOQ	ND	ND - <LLOQ
RBD Oil (without hulls)	73496	ND - <LLOQ ^e	<LLOQ	ND	ND	ND - <LLOQ
	Control	ND	ND	ND	ND	ND

^a Statistics are not provided due to small sample size.

^b NAA = N-acetylaspartate; NAG = N-acetylglutamate; NAT = N-acetylthreonine; NAS = N-acetyliserine; NAGly = N-acetylglycine

^c Values are not corrected for moisture content.

^d ND = Not detected

^e LLOQ = Lower limit of quantification, which was 0.01 µg/g for a 100 mg sample

5.3.8.4 Safety of dietary acetylated amino acids

Acetylated amino acids are ubiquitous in nature and can be used in animal feed and in industrial applications. Acetylation of proteins is commonly employed in the food industry to alter the solubility, water absorption capacity and emulsifying properties of protein concentrates (eg. El-Adawy, 2000). The safety of acetylated amino acids has been considered previously in other assessments (Applications A1006⁸ and A1021⁹), and no food safety concerns were identified.

Biochemical studies have shown both NAAsp and NAGlu are produced enzymatically in mammals and occur in a range of tissues and organs. NASer has been identified on the N-terminus of proteins as a commonly occurring post-translational modification of cytosolic proteins in eukaryotes. It has been estimated that approximately 90% of proteins with N-terminal serine residues are acetylated (Driessen *et al.*, 1985). NAThr is also associated with acetylated proteins. N-terminal acetylation of proteins has been reported to protect them from degradation, assist in the export of soluble proteins from the cell and block activation of certain signalling pathways. A number of naturally occurring eukaryotic proteins contain NAGly, including cytochrome c, haemoglobin and ovalbumin.

The Applicant has conducted a number of *in vitro* and *in vivo* toxicology studies on NAAsp, NAGlu, NAThr, NASer and NAGly, including a two generation reproductive toxicity study with NAAsp. None of the studies showed evidence of adverse effects.

⁸ Safety Assessment of herbicide-tolerant soybean line DP-356043-5, Section 5.6.1 (see <http://www.foodstandards.gov.au/code/applications/documents/A1006%20GM%20Soybean%202nd%20AR%20FINAL%20AMENDED1.pdf>)

⁹ Supporting Document 1, Safety Assessment of herbicide-tolerant maize line DP-098140-6, Section 5.6.1 (see <http://www.foodstandards.gov.au/code/applications/documents/A1021%20GM%20Maize%20SD1%20Safety%20Assessment.pdf>)

It is evident from the studies described above that acetylated amino acids are normal constituents of foods. A body of data, including nutritional, toxicological, biochemical and metabolism studies in different mammalian species, provides evidence that acetylated amino acids have a history of safe use and safe dietary consumption. Given their presence in numerous biochemical contexts in plants and animals consumed as food, the presence of acetylated amino acids in canola line 73496 is not a food safety concern, even at significantly increased levels compared with conventional canola.

5.4 Conclusion from compositional analysis

Detailed compositional analyses of seed from herbicide-treated canola line 73496 and the near isogenic control were conducted on plants grown under normal agricultural conditions at five trial sites in canola growing regions of Canada and the U.S.A. The analyses included proximates (protein, fat, ash and carbohydrates), fibre components (ADF, NDF, CF), fatty acids, amino acids, micronutrients (minerals and vitamins) and anti-nutrients (glucosinolates, phytic acid, sinapine and tannins). The levels of these key constituents in canola line 73496 were compared to those in the near isogenic control, as well as to the ranges reported for commercial canola varieties grown under similar agricultural conditions.

Statistically significant differences were found in a number of individual analytes in the across-sites analyses between canola line 73496 and the control. However, in each case the differences occurred at only some trial sites and were small in magnitude. In addition, all nutrient mean values were within the range of natural variability for commercial canola. The composition of canola can vary significantly with the site and agricultural conditions in which it is grown, and the differences reported here are consistent with normal biological variability. Overall, the compositional data support the conclusion that there are no biologically significant differences in the levels of key constituents in seed from canola line 73496 compared with conventional canola already available in agricultural markets.

Due to the enzyme activity of GAT4621, five acetylated amino acids are produced in measurably higher amounts in canola line 73496 than in the non-GM control. In particular, NAA_{sp} and NAG_{lu} were significantly increased, however absolute levels as a percentage of total seed protein remained low at approximately 0.6% and 0.01% respectively. Acetylated amino acids are ubiquitous in the food supply and the large increase in their abundance in canola line 73496 is not a safety concern. Moreover, extraction and processing of canola oil typically removes proteins, effectively reducing potential dietary exposure to acetylated amino acids to zero.

6. Nutritional impact

In assessing the safety of a GM food, a key factor is the need to establish that the food is nutritionally adequate and will support typical growth and well-being. In most cases, this can be achieved through a detailed understanding of the genetic modification and its consequences, together with an extensive compositional analysis of the food, such as that presented in Section 5 of this report.

Where a GM food has been shown to be compositionally equivalent to conventional varieties, the evidence to date indicates that feeding studies using target livestock or other animal species will add little to the safety assessment (EFSA, 2008; OECD, 2003). If the compositional analysis indicates biologically significant changes, either intended or unintended, to the levels of certain nutrients in the GM food, additional nutritional studies should be undertaken to assess the potential impact of the changes on the whole diet.

Canola line 73496 is the result of a simple genetic modification to confer tolerance to glyphosate, with no intention to significantly alter nutritional parameters in the food. The

compositional analyses have demonstrated that the genetic modification has not altered the nutritional adequacy of canola line 73496 as a source of food when compared with that of conventional canola varieties. The introduction of foods derived from canola line 73496 into the food supply is therefore expected to have negligible nutritional impact.

References

- An, G., Mitra, A., Choi, H.K., Costa, M.A., An, K., Thornburg, R.W. and Ryan, C.A. (1989) Functional analysis of the 3' control region of the potato wound-inducible proteinase inhibitor II gene. *Plant Cell*, **1**(1):115-122.
- Astwood, J.D. and Fuchs, R.L. (1996) Allergenicity of foods derived from transgenic plants. *Highlights in food allergy. Monographs in Allergy*, **32**. pp 105-120.
- AOF (2007) Delivering market choice with GM canola. Australian Oilseeds Federation, see http://www.australianoilseeds.com/Technical_Info/industry_reports/gm_canola
- 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.
- Baylis, A.D. (2000) Why glyphosate is a global herbicide: strengths, weaknesses and prospects. *Pest Management Science* **56**: 299-308.
- Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society B* **57**:289-300.
- Bradshaw, L.D., Padgett, S.R., Kimball, S.L. and Wells, B.H. (1997) Perspectives on glyphosate resistance. *Weed Technology* **11**:189-198.
- Castle, L.A., Siehl, D.L., Gorton, R., Patten, P.A., Chen, Y.H., Bertain, S., Cho, H.J., Duck, N., Wong, J., Liu, D. and Lassner, M.W. (2004) Discovery and directed evolution of a glyphosate tolerance gene. *Science*, **304**(5674):1151-1154.
- Chen, W. and Tulsieram L., inventors (2007) Microprojectile Bombardment Transformation of Brassica. US Patent Application No. 11/270,996.
- Codex (2004) *Guideline for the Conduct of Food Safety Assessment of Foods Derived from Recombinant-DNA Plants*. Report No. CAC/GL 45-2003, Codex Alimentarius Commission, Rome.
- Codex Alimentarius (2009). Foods derived from modern biotechnology. Codex Alimentarius Commission, Joint FAO/WHO Foods Standard Programme, Food and Agriculture Organisation of the United Nations, Rome, Italy.
- Delaney, B., Astwood, J.D., Cunny, H., Eichen Conn, R., Herouet-Guicheney, C., MacIntosh, S., Meyer, L.S., Privalle, L., 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.
- Driessen, H.P.C. de Jong, W.W., Tesser, G.I. and Bloemendal, H. (1985) The mechanism of n-terminal acetylation of proteins. *CRC Critical Reviews in Biochemistry* **18**:281-325.
- Dyda, F., Klein, D.C. and Hickman, A.B. (2000) GCN5-Related N-Acetyltransferases: A structural overview. *Ann Rev Biophys and Biomol Structure* **29**:81-103.
- 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. http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1211902590265.htm. Accessed 5 April 2010.

El-Adawy, T.A. (2000) Functional properties and nutritional quality of acetylated and succinylated mung bean protein isolate. *Food Chemistry* **70**:83-91.

EPA (1996) *Bacillus licheniformis* TSCA Section 5(h)(4) Exemption: Final Decision Document. United States Environmental Protection Agency. http://www.epa.gov/biotech_rule/pubs/ra/fd005.htm.

FARRP (2011). Allergen database. University of Nebraska, Food Allergy Research and Resource Program, Lincoln, Nebraska. <http://www.allergenonline.org/databasebrowse.shtml> [accessed 14 February 2011].

FAO (2005) Pesticide residues in food - 2005. Food and Agriculture Organisation of the United Nations, Rome. <http://www.fao.org/docrep/009/a0209e/a0209e00.htm>

FSANZ 2010. Approval Report and Safety Assessment for Application A1021 Food derived from herbicide tolerant corn line DP-098140-6
<http://www.foodstandards.gov.au/code/applications/pages/applicationa1021food4210.aspx>

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.

Keil, M., Sanchez-Serrano, J., Schell, J. and Willmitzer, L. (1986) Primary structure of a proteinase inhibitor II gene from potato (*Solanum tuberosum*). *Nucleic Acids Res.*, **14**(14):5641-5650.

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. *Toxicol.Sci* **48**(2):157-162.

Ladics, G.S., Bannon, G.A., Silvanovich, A. and Cressman, R.F. (2007). Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens. *Molecular Nutrition and Food Research* **51**:985-998.

Lancashire, P.D., Bleiholder, H., Van Den Boom, T. Langeluddeke P., Stauss, R. Weber, E. and Witzemberger A. (1991) A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology* **119**:561-601.

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

Norris, S.R., Meyer, S.E. and Callis, J. (1993) The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. *Plant Mol Biol* **21**:895-906.

OECD (2011) Consensus Document on Key Nutrients and Key Toxicants in Low Erucic Acid Rapeseed ENV/JM/MONO(2011)55

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

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.

Rey, M.W., Ramaiya, P., Nelson, B.A., Brody-Karpin, S.D., Zaretsky, E.J., Tang, M., Lopez de Leon, A., Xiang, H., Gusti, V., Clausen, I.G., Olsen, P.B., Rasmussen, M.D., Anderson, J.T., Jorgensen, P.L., Larsen, T.s., Sorokin, A., Bolotin, A., Lapidus, A., Galleron, N., Ehrlich, S.D., and Berka, R.M. (2004). Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. *Genome Biol.* **5**(10):R77.

Salkinoja-Salonen, M.S., Vuorio, R., Andersson, M.A., Kämpfer, P., Andersson, M.C., Honkanen-Buzalski, T. and Scoging, A.C. (1999) Toxigenic strains of *Bacillus licheniformis* related to food poisoning. *Applied and Environmental Microbiology*, **65**:4637-4654.

Siehl, D.L., Castle, L.A. Gorton, R. and Keenan, R.J. (2007). The molecular basis of glyphosate resistance by an optimized microbial acetyltransferase. *J. Biol. Chem.* 282(15):11446-11455.

Silvanovich, A., Nemeth, M.A., Song, P., Herman, R., Tagliani, L. and Bannon, G.A. (2006). The value of short amino acid sequence matches for prediction of protein allergenicity. *Toxicological Sciences* 90:252-258.

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.

Vetting, M.W., de Carvalho, L.P., Yu, M., Hegde, S.S., Magnet, S., Roderick, S.L., Blanchard, J.S. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.*, **433**(1):212-226.

Westfall, P.H., Tobias, R.D., Rom, D., Wolfinger, R.D. and Hochberg, Y. (1999) Concepts and Basic Methods for Multiple Comparisons and Tests. In Multiple Comparisons and Multiple Tests: Using SAS. SAS Institute Inc., Cary, pp 13-40.