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

Supporting Document 1

SAFETY ASSESSMENT REPORT (UPDATED)

APPLICATION A1028 – FOOD DERIVED FROM INSECT PROTECTED AND HERBICIDE TOLERANT COTTON LINE T304-40

SUMMARY AND CONCLUSIONS

Background

A new genetically modified (GM) cotton line, T304-40, has been developed that is protected against feeding damage by Lepidopteran insect larvae, and which is also tolerant to herbicides containing glufosinate ammonium. Insect protection is conferred by expression of a modified Cry1Ab protein from *Bacillus thuringiensis* and herbicide tolerance is conferred by expression of phosphinothricin acetyltransferase (PAT) from *Streptomyces hygroscopicus*.

T304-40 cotton lines are presently only intended for cultivation in the United States of America and Canada. Therefore foods derived from cotton line T304-40, if approved by FSANZ, would only enter the food supply in Australia and New Zealand through imported products.

History of Use

The host organism is cultivated cotton (*Gossypium hirsutum* L.). Cotton is one of the oldest cultivated crops and is grown worldwide primarily as a fibre crop but also as a source of food products derived from the seed. Such products need to be highly processed because of the presence of natural toxicants (gossypol) and anti-nutrients (cyclopropenoid fatty acids) in the unprocessed seed. The main food products from cotton line T304-40 would, like cotton products currently available on the market, consist of oil and linters.

Molecular Characterisation

Cotton line T304-40 contains two novel gene cassettes. One contains a *cry1Ab* gene that encodes an insecticidal crystal protein and the other contains a *bar* gene that encodes a protein conferring tolerance to herbicides containing glufosinate ammonium (phosphinothricin). There are no antibiotic resistance markers present in line T304-40.

Comprehensive molecular analyses of cotton line T304-40 indicate that there is a single insertion site containing two, almost complete copies of the *cry1Ab* cassette, an almost complete copy of the *bar* cassette and an isolated partial terminator sequence from the *cry1Ab* cassette. The introduced genetic elements are stably inherited from one generation to the next. Plasmid backbone analysis shows that no plasmid backbone has been incorporated into the transgenic locus.

Two unexpected ORFs are present at junctions associated with the insertion site but lack the necessary regulatory sequences to express a protein. No known endogenous genes have been interrupted by insertion of the new genetic material.

Characterisation of Novel Protein

Cotton line T304-40 express two novel proteins, Cry1Ab and PAT. Expression analyses of the two proteins showed that PAT was expressed in all plant parts tested but was highest in young leaves (61.4 µg/g fresh weight). Cry1Ab was not detectable in any plant parts except the seed (3.7 µg/g fresh weight).

A number of studies were done to confirm the identity and physicochemical and functional properties of the expressed plant-derived Cry1Ab and PAT proteins, as well as to determine their potential toxicity and allergenicity. These studies have demonstrated that the Cry1Ab and PAT proteins conform in size and amino acid sequence to that expected, do not exhibit any post-translational modification including glycosylation and, in the case of PAT, exhibit the expected enzymatic activity. The bioactivity of the expressed Cry1Ab protein was unable to be tested.

In relation to potential toxicity and allergenicity, it is worth noting that the Cry1Ab and PAT proteins are inherently non-toxic to mammals and do not exhibit any potential to be allergenic to humans. In addition, bioinformatic studies have confirmed their lack of any significant amino acid sequence similarity to known protein toxins or allergens and digestibility studies have demonstrated that both proteins would be rapidly degraded in the stomach following ingestion. Acute oral toxicity studies in mice have also confirmed their absence of toxicity in animals. Both proteins exhibit a degree of heat stability however given their digestive lability, this does not raise any safety concerns. Taken together, the evidence indicates that Cry1Ab and PAT are unlikely to be toxic or allergenic to humans.

Compositional Analyses

Detailed compositional analyses were done of fuzzy seed-derived from T304-40 plants. Analyses were done of proximates (crude protein, crude fat, ash and total carbohydrates), ADF, NDF, fatty acids, amino acids, micronutrients (minerals and α-tocopherol) and anti-nutrients (gossypol, phytic acid and cyclopropenoid fatty acids). The levels were compared to levels in the non-GM parent as well as to the ranges found in commercial cotton cultivars reported in the literature. Additionally, levels of analytes were measured in processed commodities derived from control and GM cottonseed, although the results from these commodities were not analysed statistically.

For fuzzy cottonseed, across most of the categories several significant differences in analytes between the GM and non-GM seeds were found. The composition of cotton can vary significantly with the site, agricultural conditions and season of production, and differences reported here most likely reflect normal biological variability. Additionally, with the exception of palmitoleic acid, the mean analyte levels found in seeds from T304-40 fell within the range of natural variation in commercial cotton cultivars. The lower level of palmitoleic acid in seeds from T304-40 does not raise a concern because of the low contribution (<0.5%) of palmitoleic acid to total fatty acids.

For processed products derived from cottonseed there were no large discrepancies between the control and the GM line for the means of any analyte.

Taken overall, the compositional data are consistent with the conclusion that there are no biologically significant differences in the levels of key components in seed from cotton line T304-40 when compared with conventional cotton cultivars currently on the market.

Nutritional Impact

Based on the results from a broiler feeding study, it was concluded that cottonseed meal from cotton T304-40 was nutritionally adequate, and equivalent to that derived from a non-GM control cotton and a commercial non-GM cultivar, in its ability to support typical growth and well being.

Conclusion

No potential public health and safety concerns have been identified in the assessment of cotton line T304-40. On the basis of the data provided in the present Application, and other available information, food derived from cotton line T304-40 is considered as safe for human consumption as food derived from conventional cotton cultivars

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ADF	acid detergent fibre
bar	*bialaphos resistance* gene
Bt	*Bacillus thuringiensis*
BLAST	Basic Local Alignment Search Tool
bp	base pairs
Cry	Crystal protein from *Bacillus thuringiensis*
Cyt	Cytolytic protein from *Bacillus thuringiensis*
DNA	deoxyribonucleic acid
T-DNA	transferred DNA
dw	dry weight
ELISA	enzyme linked immunosorbent assay
FAO	Food and Agriculture Organization of the United Nations
FSANZ	Food Standards Australia New Zealand
GM	genetically modified
HRP	horseradish peroxidase
ILSI	International Life Sciences Institute
kDa	kilo Dalton
LB	left border
LC/MS	high performance liquid chromatography/electrospray mass spectrometry
NDF	neutral detergent fibre
OECD	Organisation for Economic Co-operation and Development
OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
OVA	ovalbumin
PAT	phosphinothricin acetyltransferase
PCR	polymerase chain reaction
PFT	pore forming toxin
L-PPT	L-phosphinothricin
RAC	raw agricultural commodity
RB	right border
RBS	ribosome binding site
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid

Ti	tumour-inducing
U.S.	United States of America
WHO	World Health Organization

[PAGE INTENTIONALLY LEFT BLANK]

1. INTRODUCTION

A genetically modified (GM) cotton line, T304-40, has been developed which is protected against feeding damage by Lepidopteran insects, particularly cotton bollworm (*Helicoverpa* spp.), and tolerant to herbicides containing glufosinate ammonium.

Protection against feeding damage by Lepidopteran insects is achieved through expression in the plant of an insecticidal Cry protein, Cry1Ab, encoded by a modified *cry1Ab* gene derived from the soil bacterium *Bacillus thuringiensis,* and tolerance to glufosinate ammonium is achieved through expression of phosphinothricin acetyltransferase (PAT) encoded by the *bar* gene derived from another soil bacterium *Streptomyces hygroscopicus*. Both of these genes have been widely used for genetic modification of a number of crop species.

The main food products derived from cotton line T304-40 would be oil and linters.

2. HISTORY OF USE

2.1 Host Organism

The host organism is cultivated cotton (*Gossypium hirsutum* L.). Cotton is one of the oldest cultivated crops and is grown primarily as a fibre crop, providing over 40% of the total fibre used in the world (OECD, 2004). Only the cotton boll, which develops from the plant ovary, is used for either textile fibre or food/feed. The cotton boll, once harvested, is processed ('ginned') to separate the cottonseed from the cotton fibre.

Cottonseed is processed into four major by-products: oil, meal, hulls and linters, of which the oil and linters are typically used as human food. Oil is the main derived product used for human consumption and, for example, in the first half of the 20th century, cottonseed oil was the major vegetable oil consumed in the United States (O'Brien, 2008). Food products from cottonseed are limited to highly processed products because of the presence of natural toxicants (gossypol) and anti-nutrients (cyclopropenoid fatty acids) in the unprocessed seed.

Worldwide, cottonseed oil ranks 6th in vegetable oil consumption. The fatty acid profile of the oil comprises mainly oleic and linoleic acids. The natural oil has a strong and unpleasant flavour and requires a process known as deodorisation to render it palatable (O'Brien, 2008). Cottonseed oil has a variety of food uses including frying oil, salad and cooking oil, and inclusion in mayonnaise, salad dressing, shortening, and margarine. In the course of processing to food grade quality oil, all proteins are destroyed by high temperatures and pressure, or are separated out by extraction with a non-polar solvent and destroyed by the temperature of solvent recovery. Any last traces of protein in the oil are removed by the alkali treatment and deodorisation steps of the oil refining.

Cotton linters are short fibres removed from the cottonseed during processing and are a major source of cellulose for both chemical and food uses. They are used as a cellulose base in products such as high fibre dietary products as well as a viscosity enhancer (thickener) in ice cream, salad dressings and toothpaste.

The other major by-products – meal and hulls – are used as stock feed. Cottonseed meal is not used for human consumption in Australia or New Zealand. It has permission to be used for human food (after processing) in the U.S. and other countries, but is primarily sold for stock feed. While cottonseed contains up to 34% protein (OECD, 2004) and has

potential to be used as a dietary source of protein, the level of gossypol in the seed is toxic to humans and other monogastrics. The gossypol is stored in pigment glands and plays a role in deterring predators. A number of options for removing gossypol in unprocessed seed have been tried over the years, including the development of glandless lines but none has so far proved viable (Lusas *et al*., 1989). Production of cottonseed flour has been reported in a number of developed countries (Lusas *et al*., 1989) and cottonseed flour has been used as a component in special products in Central America to help ease malnutrition (Scrimshaw, 1980). Cottonseed flour is also permitted for human consumption in the U.S., provided it meets certain specifications for gossypol content, although no products are currently being produced.

Cotton is not grown in New Zealand. Australia has significant plantings of the crop although the area varies annually due largely to prevailing environmental factors. In the 2006 – 07 season, 92% of the commercial cotton planted in Australia was genetically modified (Molony and Hassall, 2008).and the traits all concerned insect protection and/or tolerance to a herbicide (OGTR, 2008b). Although the main product of the cotton plant is seen as fibre, cotton is also Australia's major oilseed crop. Most cottonseed is exported as fuzzy seed destined for animal feedlots but a proportion of the seed is retained to produce oil, mainly for domestic use. In 2006, some 10,900,000 tonnes of oil was produced in Australia (FAOSTAT – available at http://faostat.fao.org/default.aspx). Cottonseed oil makes up around 15% of the total Australian domestic fat and oil supply and is used primarily in some margarines, blended vegetable cooking oils and oil for commercial deepfryers (Molony and Hassall, 2008).

The cotton cultivar 'Coker 315' has been used as the parental variety for the genetic modification described in this application, and thus is regarded as the near-isogenic line for the purposes of comparative assessment. It is not grown commercially in Australia. 'Coker' cultivars are U.S. cultivars that are widely used in producing GM cotton lines because they can be readily cultured and regenerated in the laboratory (OGTR, 2008a). Traits introduced into 'Coker' cultivars are transferred to commercial cultivars by backcrossing.

2.2 Donor Organisms

2.2.1 Bacillus thuringiensis *(Bt)*

The Cry protein expressed in cotton line T304-40 is derived from *B. thuringiensis* subspecies *berliner* 1715 (Höfte *et al*., 1986). *B. thuringiensis* is a species that, while predominantly occurring in soil, is found ubiquitously throughout the natural environment. *B. thuringiensis* is a facultative anaerobic, gram-positive spore-forming bacterium that produces characteristic insecticidal proteins, as parasporal crystals, during the sporulation phase. These crystals predominantly comprise one or more Crystal (Cry) and Cytolytic (Cyt) toxins, also called δ-endotoxins. These toxins are highly specific to their target insect species, are innocuous to humans, vertebrates and plants, and are completely biodegradable (Bravo *et al*., 2007).

Over 60 subspecies of *B. thuringiensis* have been described. *B. thuringiensis* subspecies can synthesise several types of Cry protein, which are specifically toxic to the insect orders Coleoptera, Diptera, Hymenoptera, and Lepidoptera, and also to nematodes. The Cyt toxins are mostly found in *B. thuringiensis* strains that are active against Diptera.

The effect of *B. thuringiensis* products on human health and the environment was the subject of a critical review by the WHO International Programme on Chemical Safety (WHO, 1999). The review concluded that '*B. thuringiensis* products are unlikely to pose any hazard

to humans or other vertebrates or the great majority of non-target invertebrates provided that they are free from non-*B. thuringiensis* microorganisms and biologically active products other than the insecticidal proteins'.

A number of different commercial *B. thuringiensis* formulations have been registered worldwide for use as an insecticide to be applied to foliage, soil, and water or food storage facilities. While the *B. thuringiensis* spores or vegetative cells may persist in the environment for weeks, months or years, the Cry proteins become inactive within hours or days (OECD, 2007).

Studies on mammals, particularly laboratory animals, demonstrate that *B. thuringiensis* is mostly non-pathogenic and non-toxic. *B. thuringiensis* has been demonstrated to be highly specific in its insecticidal activity and has demonstrated little, if any, direct toxicity to non-target insects (see NPTN, 2000; OECD, 2007 and references therein)

The use of *B. thuringiensis* products in the field can result in considerable aerosol and dermal exposure in humans. With the exception of case reports on ocular irritation (Samples and Buettner, 1983) and inflammation after a needle stick injury (Warren *et al*., 1984), no adverse health effects have been documented after occupational exposure to *B. thuringiensis* products. *B. thuringiensis* present in drinking water or food has not been reported to cause adverse effects on human health (NPTN, 2000; OECD, 2007; WHO, 1999).

2.2.2 Streptomyces hygroscopicus

The source of the *bar* (*bialaphos resistance*) gene is the bacterial species *Streptomyces hygroscopicus*, strain ATCC21705 (Murakami *et al*., 1986). The *Streptomycetae* bacteria were first described in the early 1900's. These organisms are generally soil-borne, although they may also be isolated from water. They are not typically pathogenic to animals including humans, and few species have been shown to be phytopathogenic (Bradbury, 1986; Kützner, 1981).

Although these organisms are not used in the food industry, the *bar* gene from *S. hygroscopicus*, has been used to confer glufosinate ammonium-tolerance in food producing crops (see Section 4.5.1.3). The *pat* gene from the closely related *S. viridochromogenes* produces a protein that is structurally and functionally equivalent to the protein encoded by the *bar* gene (Wehrmann *et al*., 1996).

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; and the genetic stability of the inserted DNA and any accompanying expressed traits.

Studies submitted:

Criel, I. (2008). Description of the T304-40 transformation methodology. Report ID: BIO2-005_TransfMethod_150. Bayer CropScience (unpublished).

Lecleir, M. (2007). Description of vector pTDL008. Report ID: BIO2-005_VectValid_076, Bayer CropScience (unpublished).

Moens, S.; Criel, I. (2008). Detailed insert characterization of *Gossypium hirsutum* transformation event T304-40. Report No. BBS07-014. Bayer CropScience (unpublished).

Habex, V.; van Herck, H. (2007). Confirmation of the absence/presence of vector backbone sequences in *Gossypium hirsutum* transformation event T304-40. Report No. BBS05-003. Bayer CropScience (unpublished).

Moens, S.; De Pestel, K. (2008). Full DNA sequence of event insert and integration site of *Gossypium hirsutum* transformation event T3-4-40. Report No. BBS08-002. Bayer CropScience (unpublished study).

De Pestel, K. (2008). Bioinformatics analysis of newly created ORFs from *Gossypium hirsutum* event T304-40. Report No. BIO2-005_Bioinfo_136. Bayer CropScience (unpublished).

De Pestel, K. (2008). Bioinformatics analysis of the pre-insertion locus of cotton transformation event T304-40. Report No. BIO2_FullSeq_137. Bayer CropScience (unpublished).

Mertens, K.;Moens, S. (2008). Structural stability analysis of *Gossypium hirsutum* transformation event T304-40 over different generations, in different backgrounds and grown in different environments. Report BIO2-004_structStab_167, Bayer CropScience (unpublished).

3.1 Method used in the genetic modification

Cotton cultivar 'Coker 315' was transformed with the T-DNA vector pTDL008 (Figure 1) via *Agrobacterium*-mediated transformation essentially as described by Cousins et al. (1991) with modifications described by Murray et al. (1999). The genes of interest were inserted into the plasmid between DNA sequences known as the Left and Right Borders (LB and RB). These border sequences were isolated from the Ti plasmid of *Agrobacterium tumefaciens* and normally delimit the DNA sequence (T-DNA) transferred into the plant (Zambryski, 1988).

Basically, cotyledon explants from *in vitro* germinated seedlings were co-cultivated with *Agrobacterium tumefaciens* strain EHA 101 containing standard binary vectors (Deblaere *et al*., 1987). Following callus induction on a medium containing glufosinate ammonium as the selection agent, callus showing tolerance to the herbicide was transferred to an embryogenesis medium for regeneration of plantlets. These plantlets (T0) were then transferred to a glasshouse for further characterisation and selection.

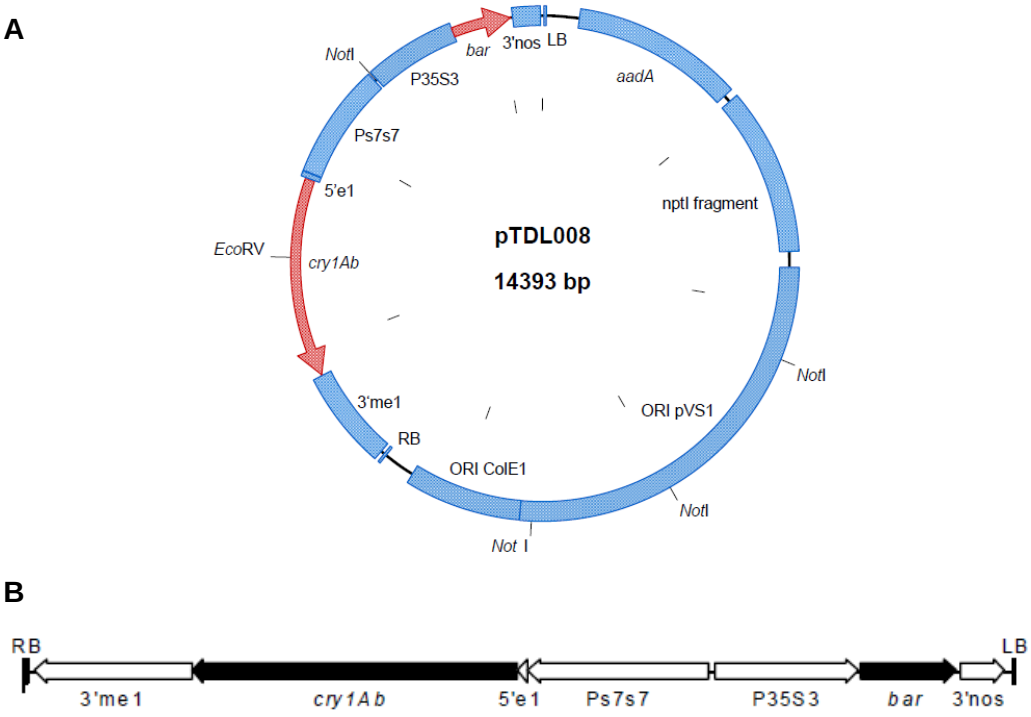

Figure 1: Vector map of plasmid pTDL008 (A) and the T-DNA region of the plasmid (B)

3.2 Function and regulation of introduced genes

Information on the genetic elements in the pTDL008 plasmid is summarised in Table 1.

Table 1: Description of the genetic elements contained in plasmid pTDL008

Genetic element	bp location on plasmid pTDL008	Size (kb)	Source	Orientation	Description & Function	References
T-DNA						
3'me1	8792-9728	0.94	*Flaveria bidentis*	Anti-clockwise	• 3' UTR of the NADP-malic enzyme gene • Terminates *cry1Ab* gene expression & directs polyadenylation	Marshall et al. (1997)
cry1Ab	9729-11582	1.85	*Bacillus thuringiensis*	Anti-clockwise	• Modified coding sequence of the *cry1Ab5* crystal protein gene	Jansens et al (2006)
5'e1	11586-11643	0.06	*Oryza sativa*	Anti-clockwise	• Sequence including leader sequence of the tapetum-specific GE1 gene • Promotes *cry1Ab* expression	Michiels et al. (1992)
Ps7s7	11644-12685	1.04	Subterranean clover stunt virus	Anti-clockwise	• Sequence including the duplicated promoter region derived from genome segment 7 • Promotes *cry1Ab* expression	Boevink et al. (1995)
P35S3	12686-13543	0.86	Cauliflower Mosaic Virus	Clockwise	• Sequence including the promoter region of the 35S transcript • Promotes *bar* expression	Odell et al. (1985)
Bar	13544-14095	0.55	*Streptomyces hygroscopicus*	Clockwise	• Coding sequence of the phosphinothricin acetyltransferase gene	Thompson et al. (1987)
3'nos	14096-	0.31	*Agrobacterium*	Clockwise	• Sequence including the 3'UTR	Depicker et al.

Genetic element	bp location on plasmid pTDL008	Size (kb)	Source	Orientation	Description & Function	References
	14393, 1-12		*tumefaciens*		of the nopaline synthase gene • Terminates *bar* gene expression and directs polyadenylation	(1982)
Plasmid Backbone						
LB	13-37		*Agrobacterium tumefaciens*		• Left border repeat • Required for the transfer of the T-DNA into the plant cell	Zambryski (1988)
	38-342				• Residual plasmid sequences from pTiAch5 flanking the left border	
aadA	343-342	1.62	*E. coli*	Anti-clockwise	• Fragment including the aminoglycoside adenyltransferase gene • Confers resistance to erythromycin, streptomycin, and spectinomycin • Allows selection of *E.coli* during vector construction.	
nptI fragment	1966-3486	1.52	Transposon Tn903	Anti-clockwise	• Fragment of the neomycin phosphotransferase 1 coding sequence • Confers resistance to kanamycin • Allows selection of *E.coli* during vector construction.	
	3487-3632	0.15			• Fragment including residual sequences upstream of the *aadA* gene	
ORI pVS1	3633-7403	3.77	*Pseudomonas aeruginosa*.		• Fragment including the origin of replication from plasmid pVS1 • Allows replication in *A. tumefaciens*	
ORI ColE1	7404-8576		*E. coli*		• Fragment including the origin of replication from plasmid pBR322 • Permits replication of plasmid in *E. coli*.	
	8577-8766				• Residual plasmid sequences from pTiAch5 flanking the right border	
RB	8767-8791	0.19	*Agrobacterium tumefaciens*		• Right border repeat • Required for the transfer of the T-DNA into the plant cell	Zambryski (1988)

3.2.1 cry1Ab gene

The modified *cry1Ab* gene[1] (Jansens *et al.*, 2006) has been derived from a gene (Genbank accession No. X04698 - first cloned and characterised by Höfte et al (1986)) which, under the latest nomenclature system, is now known as *cry1Ab5 (Bacillus thuringiensis* toxin nomenclature,database available online at http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/toxins2.html). The gene sequence has been modified for expression in plants and the gene has been truncated. The *cry1Ab* coding region in plasmid pTDL008 is 1853 bp in length and is driven by a chimeric promoter region comprising elements from the rice tapetum-specific E1 gene and the subterranean clover stunt virus genome segment 7. At the 3' untranslated region of the coding region is an mRNA termination region derived from the NADP-malic enzyme gene of *Flaveria bidentis*, a C_4 photosynthetic dicotyledonous plant.

[1] Hereafter the term 'modified' will be dropped and the gene and protein will be described just as *cry1Ab* and Cry1Ab respectively.

3.2.2 bar gene

The *bar* gene from *Streptomyces hygroscopicus* and the *pat* gene from *S. viridochromogenes* confer tolerance to herbicides containing glufosinate ammonium (phosphinothricin) – see Section 4.1.2. Both genes code for polypeptides of 183 amino acids and share 87% homology at the nucleotide sequence level (Wehrmann *et al*., 1996). Both genes have been widely used for genetic modification of food species.

The bar gene coding region (Thompson *et al*., 1987) used in plasmid pTDL008 is 551 bp in length. It is driven by the promoter of the Cauliflower Mosaic Virus 35S RNA and terminated by a TaqI fragment from the 3'untranslated region of the nopaline- synthase gene originating from the T-DNA of plasmid pTiT37 from *Agrobacterium tumefaciens.* A similar cassette has been described for use in genetic modification of herbicide tolerant cotton (Trolinder *et al.*, 2008).

3.3 Breeding of cotton plants containing transformation event T304-40

A breeding programme was undertaken for the purposes of:

- obtaining generations suitable for analysing the molecular and genetic characteristics of the line containing the T304-40 event, and
- ensuring that the T304-40 event is incorporated into a conventional breeding line for commercialisation of insect-protected, herbicide-tolerant cotton.

A number of lines (plants) with potential were chosen at the T0 stage. Basically, a series of selfing and then selection of resulting seedlings under glufosinate ammonium application was done to reduce the number of lines. This series proceeded up to generation T7. The line containing event T304-40 was eventually selected for commercial development based on both its agronomic performance and expression of the two introduced genes. The breeding programme for this line is summarised in Figure 2.

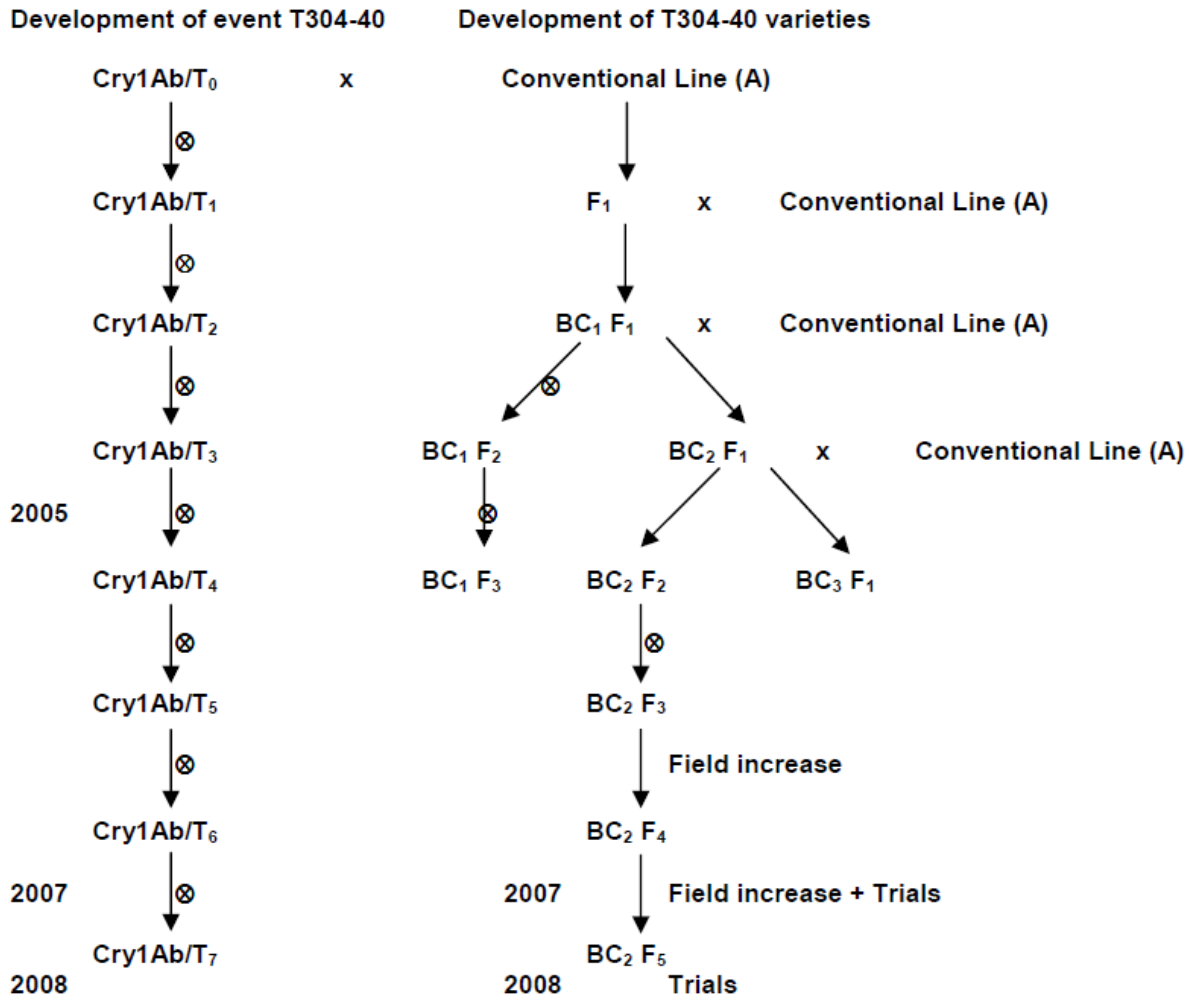

Figure 2: Breeding strategy for plants containing the T304-40 event

3.4 Characterisation of the genes in the plant

A range of analyses were undertaken in order to characterise the genetic modification in cotton line T304-40. These analyses focussed on the nature of the insertion of the introduced genetic elements and whether any unintended genetic re-arrangements may have occurred as a consequence of the transformation procedure.

3.4.1 Transgene copy number and insertion integrity

Total genomic DNA from leaf tissue of individual seedlings of the BC1 F3 generation (Figure 2) and a negative control (non-GM wild type cultivar 'FM966') was used for Southern blot analyses. A positive control (DNA from 'FM966' spiked with T-DNA from plasmid pTDL008) was also included in the Southern blot analyses. The DNA from BC1 F3 was digested with 10 different restriction enzymes (Figure 3) while control DNA was digested with one restriction enzyme (*Xba1*). The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with seven different radiolabelled probes that represented various sections of the T-DNA, including the complete T-DNA. The sizes of all hybridisation fragments were determined using a commercially available software package. Based on the Southern blot analysis, it was determined that there is a single insert in event T304-40 comprising (Figure 3):

- a single, almost complete copy of the T-DNA (A),
- an inverted incomplete copy of the *cry1Ab* gene cassette (B),
- an additional 3'me1 terminator (C).

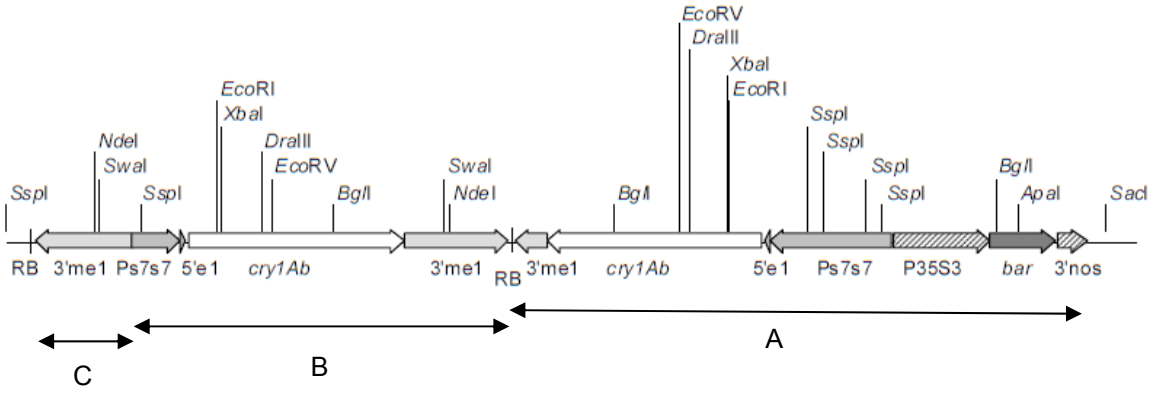

Figure 3: A. Schematic drawing of the insert in cotton line T304-40 [compare with Figure 1B which represents the arrangement of genetic elements in the T-DNA of the original plasmid] and the restriction enzymes used in the Southern blot analysis

3.4.2 Full DNA sequence of insert

Genomic DNA was obtained from leaf tissue of BC1 F3 plants and a negative control (non-GM wild type cultivar 'Coker 315'). These samples were used to determine the sequence of the integration site before transformation and the sequence of the transgenic locus. The sequence of the transgenic locus was obtained through PCR amplification of five overlapping fragments and determination of the consensus sequence. This was compared with that of the T-DNA in plasmid pTDL008.

The DNA sequence of the transgenic locus was found to consist of 479 bp of 5' flanking sequence, 9056 bp of inserted transgenic sequence and 320 bp of 3' flanking sequence (Figure 4). The 5' and 3' flanking sequences were verified as identical to sequences in the integration site DNA from cultivar 'Coker 315' and hence are of *Gossypium hirsutum* origin. A 32 bp sequence (known as the target site deletion) was found to have been deleted during the transformation process and is not present in the transgenic locus.

Analysis of the 9056 bp inserted transgenic sequence confirmed the results from the Southern blot analyses ie there is

- a partial 3'me1 terminator,
- two partial copies of the *cry1Ab* gene cassette in a tail-to-tail orientation. One of these copies has a truncated Ps7s7 promoter; the other copy has a truncated 3'me1 terminator,
- a partial copy of the *bar* gene cassette in which the 3'me1 terminator is truncated.

As a result of the transformation event, four new junctions have been created (Figure 4), two being located at the 5' and 3' ends of the insert and two being located within the insert as a result of the rearrangement.

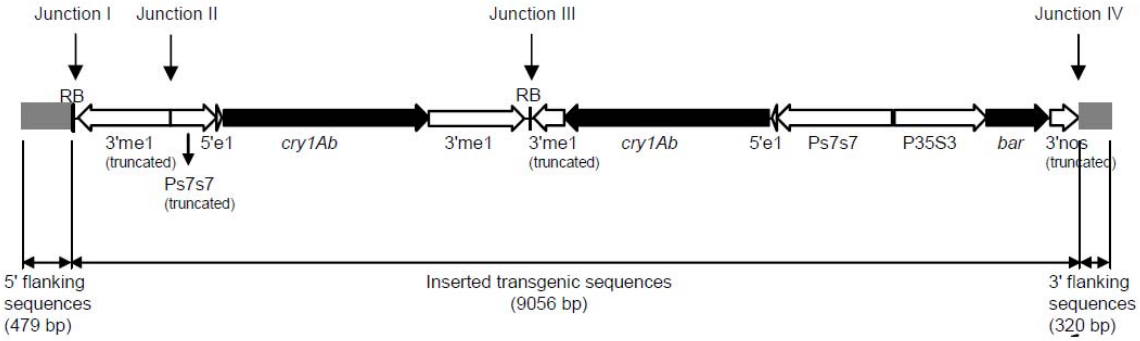

Figure 4. Location of junction regions in the inserted DNA in T304-40

3.4.3 Plasmid backbone DNA analysis

Southern blot analysis was done to determine whether any plasmid backbone had been included along with the inserted DNA in T304-40. Leaf tissue from seedlings of generation BC3 F1 (Figure 2) and a negative control (non-GM wild type cultivar 'FM966') was used for this analysis. A positive control (DNA from 'FM966 spiked with vector DNA from plasmid pTDL008) was also included in the Southern blot analyses. The DNA from BC3 F1 was digested with *Not1* (see Figure 1A for sites) and *EcoRV* (see Figure 3 for sites) restriction enzymes while control DNA was digested with *Not1* enzyme. The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with seven different overlapping radiolabelled probes covering the complete vector backbone sequence of pTDL008 and one T-DNA probe.

No hybridisation fragments were observed with either T304-40 genomic samples or 'FM966' samples while the positive control samples showed the expected hybridisation fragments after hybridisations with the vector backbone probes. These results indicate that there are no vector backbone sequences present in cotton line T304-40.

3.4.4 Open Reading Frame (ORF) analysis

Using sequence data encompassing the 5' and 3' flanking sequences as well as the inserted transgenic sequences, a bioinformatics analysis was performed to determine whether any new ORFs had been created at any of the four new junction regions (see Figure 4). The *in silico* analysis was done through the search programme, GetORF from the European Molecular Biology Open Software Suite. ORFs were defined as regions between start (ATG) and stop (TAA, TAG, TGA) codons with a minimum size coding for eight amino acids. Another programme, FGENESH (Softberry Inc.), was used to identify any potentially expressed genes in the junction regions.

In addition to the above, a homology search was done using the bioinformatics tool TSSP (Softberry Inc.) to identify any genetic elements in the junctions that play a role in the regulation of gene expression. A comparison was also made of the sequence surrounding the first ATG codon of any putative ORFs with a consensus sequence for the ribosome binding site (RBS) of plant genomes (Joshi *et al*., 1997) in order to determine if it may be a putative start of translation.

The GetORF analysis identified two ORFs at Junctions II and IV (Figure 4). FGENESH did not identify any newly created putative genes. TSSP identified 11 putative promoters or transcription sites (including the known promoter regions present in the insert), of which four were enhancer sequences. However, none of these promoter sequences was located within

100 – 200 bp upstream of a start codon and none was positioned in the correct orientation to be able to initiate transcription of the predicted ORFs. In addition, there was low overall similarity between the RBS sequence and the ATG region of the two predicted ORFs, with very few of the essential nucleotides of the RBS sequence being present. This indicated that translation of the putative ORFs would be unlikely.

Taken together, the analyses indicate that it is unlikely either of the two predicted ORFs identified in cotton line T304-40 will be transcribed and then translated into a protein.

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

An analysis, similar to the ORF analysis, of the pre-insertion locus was done in order to ascertain whether endogenous genes may have been disrupted by the insertion of the transgenic sequences. In addition to GetORF, FGENESH, and TSSP searches using the nucleotide sequence of the pre-insertion locus, a BLAST (Basic Local Alignment Search Tool) search (Altschul *et al*., 1997) was done to search for similarities between the pre-insertion locus and known cotton genomic nucleotide sequences in the European Molecular Biology Laboratory Cotton Database.

No ORFs or putative genes were predicted to span the target site deletion region and the BLAST search did not identify any functional genes in the pre-insertion locus. The TSSP search identified a putative promoter in the 5' flanking sequence but since no genes were found downstream, it was concluded that the promoter is unlikely to be biologically active.

Taken together, the analyses indicate that no known genes were interrupted by the insertion of transgenic DNA in the pre-insertion locus of T304-40.

3.5 Stability of the genetic changes

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

3.5.1 Genetic stability

The genetic stability of the modification in T304-40 was evaluated in individual plants of:

- four different generations: 12 – 18 plants from each of F1, BC1 F1, BC2 F1, BC2 F2 (Figure 2),
- four different genetic backgrounds: 20 hybrid plants crossed with each of '983050 – 99MS88', 'Coker 315', '9740' and '97201-16',
- three different harvest locations in Spain: 11 – 18 plants from each of Vinyols, Santa Oliva and Camarles.

The non-GM cultivar 'FM966' was used as a negative control and cultivar 'FM966' spiked with DNA from plasmid pTDL008 was used as a positive control.

Genomic DNA isolated from leaf tissue and the plasmid DNA was digested with *EcoRV* restriction enzyme (see Figure 3 for sites in the transgenic locus). The resulting DNA fragments were separated and transferred to a membrane for hybridisation with a radiolabelled *cry1Ab* probe (1822 bp).

The Southern blot analysis confirmed the presence of the expected hybridisation fragments in all tested transgenic DNA samples and therefore confirmed the genetic stability of the modification in T304-40 over different generations, in different genetic backgrounds and across different environments.

3.5.2 Phenotypic stability

The Applicant did not undertake any formal inheritance studies. However, at each step of the breeding programme (Figure 2), only those plants expressing the PAT protein (ie tolerance to glufosinate ammonium) were selected.

3.6 Antibiotic resistance marker genes

No antibiotic marker genes are present in T304-40. Plasmid backbone analysis shows that no plasmid backbone has been integrated into the cotton genome during transformation ie the *aadA* gene and *npt1* fragment, which were used as bacterial selectable marker genes, are not present in T304-40.

3.7 Conclusion

Cotton line T304-40 contains two novel gene cassettes. One contains a *cry1Ab* gene that encodes an insecticidal crystal protein and the other contains a *bar* gene that encodes a protein conferring tolerance to herbicides containing glufosinate ammonium (phosphinothricin). There are no antibiotic resistance markers present in line T304-40.

Comprehensive molecular analyses of cotton line T304-40 indicate that there is a single insertion site containing two, almost complete copies of the *cry1Ab* cassette, an almost complete copy of the *bar* cassette and an isolated partial terminator sequence from the *cry1Ab* cassette. The introduced genetic elements are stably inherited from one generation to the next. Plasmid backbone analysis shows that no plasmid backbone has been incorporated into the transgenic locus. Two unexpected ORFs are present at junctions associated with the insertion site but lack the necessary regulatory sequences to express a protein. No known endogenous genes have been interrupted by insertion of the new genetic material.

4. CHARACTERISATION OF NOVEL PROTEINS

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

expressed as expected, including whether any post-translational modifications have occurred.

Cotton line T304-40 expresses two novel proteins, a Cry1Ab protein and the PAT protein. A number of different analyses were done to determine the identity, physiochemical properties, *in planta* expression, bioactivity and potential toxicity and allergenicity of the two proteins. Because the expression of proteins *in planta* is usually too low to allow purification of sufficient quantities for safety assessment studies, a bacterial expression system was used to generate large quantities of both proteins. The equivalence of the bacterial-produced proteins to the plant-produced proteins was determined as part of the protein characterisation.

4.1 Function and phenotypic effects

4.1.1 Cry1Ab protein

The general mechanism of insecticidal activity of Cry proteins is well understood (see eg Bravo *et al*., 2007; Gill *et al*., 1992; OECD, 2007; Schnepf *et al*., 1998), with the mode of action being characterised principally in lepidopteran insects. The Cry proteins belong to a class of bacterial toxins known as pore-forming toxins (PFT) that are secreted as water-soluble proteins, which after undergoing conformational change, are able to insert into, or translocate across, the cell membranes of their host. There are two main groups of PFT: (i) the α-helical toxins in which the α-helix regions form the trans-membrane pore; and (ii) the β-barrel toxins, that insert into the membrane by forming a β-barrel composed of β-sheet hairpins from each monomer (Parker and Feil, 2005). The Cry proteins belong to the α-helical group of PFT, along with other toxins such as exotoxin A (from *Pseudomonas aeruginosa*) and diphtheria toxin.

The primary action of Cry toxins is to lyse midgut epithelial cells in the target insect by forming pores in the apical microvilli membrane of the cells, which subsequently leads to ion leakage and cell lysis. The crystal inclusions ingested by susceptible larvae dissolve in the alkaline environment of the gut, and the solubilised inactive protoxins are cleaved by midgut proteases yielding 60-70 kDa protease resistant proteins (Bravo *et al*., 2007). Toxin activation involves the proteolytic removal of an N-terminal peptide (25-30 amino acids for Cry1 toxins, 58 residues for the Cry3A and 49 for Cry2Aa) and approximately half of the remaining protein from the C-terminus in the case of the long Cry protoxins. The activated toxin then binds to specific receptors on the brush border membrane of the midgut epithelium columnar cells (Aronson and Shai, 2001; Hofmann *et al*., 1988) before inserting into the membrane.

Toxin insertion leads to formation of lytic pores in microvilli apical membranes (Aronson and Shai, 2001; de Maagd *et al*., 2001) and eventually to cell lysis and disruption of the gut epithelium. The septicaemia that inevitably follows may be mediated by an influx of enteric bacteria into the haemocoel (Broderick *et al*., 2006).

The majority of Cry1A proteins comprise two functional segments (Höfte and Whiteley, 1989), the N-terminal portion containing the insecticidal toxin and the C-terminal portion. The N-terminal region of the Cry1Ab protein is responsible for conferring resistance to Lepidopteran pests of cotton including the cotton bollworm (*Helicoverpa zea*), tobacco budworm (*Heliothis virescens*), pink bollworm (*Pectinophora gossypiella*), fall armyworm (*Spodoptera frugiperda*), and beet armyworm (*Spodoptera exigua*). Cry1Ab is one of the most studied of the proteins produced by *B. thuringiensis* (Vazquez-Padron *et al*., 2004).

4.1.2 PAT protein

The *bar* gene from *Streptomyces hygroscopicus* confers tolerance to the antibiotic called bialaphos (Murakami *et al*., 1986) that is also produced by *S. hygroscopicus* ie the bacterium has evolved a mechanism to avoid the toxicity of its own product. Bialaphos, now also used as a non-selective herbicide, is a tripeptide composed of two L-alanine residues and an analogue of glutamate known as L-phosphinothricin (L-PPT) (see Thompson *et al*., 1987) more recently known also as glufosinate ammonium. Free L-PPT released from bialaphos by peptidases (or applied directly as a synthetic herbicide) inhibits glutamine synthetase which in turn leads to rapid accumulation of ammonia and subsequent cell death.

The homologous polypeptide produced by the *bar* and *pat* genes (see Section 3.3.2) is known as phosphinothricin acetyltransferase (PAT); it is an acetyl transferase with enzyme specificity for both L-PPT and demethylphosphinothricin (DMPT) in the acetylation reaction (Thompson *et al*., 1987). In the presence of acetyl-CoA, PAT catalyses the acetylation of the free amino group of L-PPT to N-acetyl-L-PPT, a herbicidally-inactive compound. The kinetics and substrate specificity of the PAT enzyme are well characterised; it has a high specificity for L-PPT and has been shown to have a very low affinity to related compounds and amino acids; even excess glutamate is unable to block the L-PPT-acetyltransferase reaction (Thompson *et al*., 1987).

The acetyltransferase activity is heat- and pH-dependent (Wehrmann *et al*., 1996). PAT is active between temperatures of 25-55^{o}C, with maximum activity occurring between 40 and 45°C. Complete thermoinactivation occurs after 10 minutes at 60^{o}C and above. The optimum pH for PAT activity is 8.5, but it is active over a broad pH range of 6 to 11.

4.2 Protein expression analysis

Study submitted:

Currier, T.;Massengill, J. (2008). Protein expression analysis of cotton event T304-40, expressing Cry1Ab and PAT/*bar* proteins, USA, 2007. Report No. CP07B005. Bayer CropScience (unpublished).

The Cry1Ab and PAT proteins are expected to be expressed in all plant tissues. Plants were grown in a single greenhouse trial in the U.S. The Cry1Ab and PAT protein levels were determined by enzyme linked immunosorbent assay (ELISA) using commercial ELISA kits specific for each protein. The parts (matrices) sampled from 'Coker 315' control plants and glufosinate ammonium-sprayed BC_2F_4 plants (see Figure 2) are given in Table 2. For roots, stems and leaves, samples were taken at several growth stages. Samples were taken from leaf tissue of each plant used for analysis to confirm (by lateral flow strips) whether the plants were transgenic or control.

Table 2: Description of the plant parts sampled for protein quantification

Matrix	Definition
Roots	All parts of the plant below the soil line.
Stems	The central stem of the plant without leaves or petioles
Leaves	The petiole and fully expanded leaf of the plant.
Squares	The developing part of the plant that will give rise to a flower. These will be harvested at a size of about 1-3 cm measured from the base to the tip of the square.
Apex	Approximately the top 3-4 cm of the plant, which connects directly to the stem.
Bolls	The developing fruit of the plant harvested at a size of about 1-2 cm.
Whole plant	All above ground parts of the plant. Plants must contain at least 1 boll of any size and may contain flowers.
Pollen	Material collected from anthers.
Nectar	Material collected from nectaries of the flower.
Flowers	The entire flower containing all reproductive parts of the plant. Flowers were harvested on the day the flowers open.
Grain	Acid delinted seeds pulled from the cotton fibers of fully expanded bolls.

A commercially available software programme (SoftMax Pro™ - Molecular Devices, version 4.0) was used to calculate the concentrations of immunoreactive Cry1Ab and PAT proteins from optical density values. No Cry1Ab or PAT protein was detected in samples taken from 'Coker 315' plants. For T304-40 plant samples, the average Cry1Ab protein content was below the limit of detection for all parts except the seed (Table 3) while PAT protein was found in all parts and was highest in young leaves.

Table 3: Average concentration of Cry1Ab and PAT proteins in various plant parts from T304-40

Sample	Growth stage	Average protein content in µg/g fresh weight ± SD	
		Cry1Ab	PAT
Roots	1	0[A]	21.30 ± 3.90
	2	0[A]	17.60 ± 3.10
Stem	1	0[A]	38.10 ± 6.50
	3	0[A]	17.00 ± 3.70
Leaves	1	0[A]	61.40 ± 6.40
	2	0[A]	42.60 ± 7.50
	3	0[A]	18.40 ± 2.50
Squares	3	0[A]	44.50 ± 16.10
Apex	3	0[A]	26.10 ± 6.40
Bolls	3	0[A]	11.40 ± 2.10
Whole plant	3	0[A]	47.30 ± 10.70
Pollen	3	0[A]	0.129 ± 0.06
Nectar	3	0[A]	0.01 ± 0.01
Flowers	3	0[A]	40.20 ± 8.00
Grain	4	3.70 ± 0.002	35.60 ± 27.10

[A] Value below the limit of detection

4.3 Protein characterisation and equivalence

The *cry1Ab* gene sequence used in plasmid pTDL008 would be expected to encode a protein with a molecular weight of 69 kDa and containing 617 amino acids. The deduced amino acid sequence of the Cry1Ab protein in T304-40 is identical to the native protein (Höfte *et al*., 1986), now known as Cry1Ab5 (see Section 3.2.1), except that it is truncated at the C-terminal end and an alanine has been inserted at position 2 of the N-terminal end.

The PAT protein would be expected to comprise 183 amino acids and have a molecular weight of approximately 21 kDa.
The Cry1Ab and PAT proteins are not produced in sufficient quantity in cotton T304-40 to isolate enough for the toxicological and biochemical studies required for a Safety Assessment. A standard procedure to overcome this type of problem is to produce the protein in a bacterial system and, if this protein shows equivalence to the *in planta*-produced protein, to then use the bacterially-produced protein for the toxicological and biochemical studies. The Cry1Ab and PAT proteins were therefore expressed in recombinant *E. coli* and characterisation tests were done to confirm the identity and equivalency of these bacterially-produced proteins to those produced in cotton line T304-40.

Studies submitted:

Martone, A. (2008a). Structural and functional equivalence of Cry1Ab, Cry2Ae and PAT/*bar* proteins produced in *Bacillus thuringiensis* (*Bt*) and *Escherichia coli* to Cry1Ab, Cry2Ae and PAT/*bar* in TwinLink™ cotton, *Gossypium hirsutum*, USA 2007. Study CY07B006 , Bayer CropScience (unpublished).

Martone, A. (2008b). Structural and functional equivalence of Cry1Ab, Cry2Ae and PAT/*bar* proteins produced in *Escherichia coli* and *Bacillus thuringiensis* to the Cry1Ab, Cry2Ae and PAT/*bar* proteins from events T304-40, GHB119 and TwinLink™ cotton seed, *Gossypium hirsutum*, USA 2008. Study CP08B011 , Bayer CropScience (unpublished).

Martone,A. (2008c). Structural and Functional Equivalence of Cry1Ab and PAT/*bar* Proteins Produced in *Escherichia coli* to Cry1Ab and PAT/*bar* Proteins from Event T304-40 and TwinLink™ Cotton, *Gossypium hirsutum*. USA, 2008. Study CP08B003, Bayer CropScience (unpublished).

Rouquie, D. (2007). Cry1Ab protein: epitope homology and N-glycosylation searches. Study SA07187, Bayer CopScience (Unpublished).

Herouet-Guicheney, C. (2006). Phosphinothricin acetyl transferase (PAT) BAR gene product: epitope homology and N-glycosylation searches. Study SA 06000, Bayer CropScience (unpublished).

Habex, V. (2004). Description of the amino acid sequence of the Cry1Ab protein. Report No. Cry1Abaas-02, Bayer CropScience (unpublished).

De Beuckeleer, M. (2003). Description of the Amino Acid Sequence of the PAT Protein encoded from the *bar* Gene. Report No. Pat/bar aas/01, Bayer CropScience (unpublished).

In three separate studies (see Martone 2008a,b,c above), the bacterially-derived proteins were compared to Cry1Ab and PAT proteins isolated from seed and leaf tissue of both line T304-40 and a Bayer CropScience cultivar known as 'TwinLink'™ which was produced by the conventional crossing of line T304-40 with another GM line (GHB119) containing the *cry2Ae* gene (see eg OGTR, 2008a). The presence of event T304-40 in material derived from 'TwinLink'™ and line T304-40 was confirmed by PCR using appropriate primers.

4.3.1 Microbially expressed proteins

Microbial Cry1Ab protein and PAT protein were obtained from a bacterial expression system (*Escherichia coli*) using standard methods. The concentrations of the proteins thus obtained were 1.08 mg/ml for Cry1Ab and 1.2 mg/ml for PAT. In order to characterise the bacterially-produced proteins, a number of parameters were measured:

- molecular weight (SDS-PAGE)
- immunoreactivity (western blotting)
- protein activity (bioassay for Cry1Ab and enzymatic assay for PAT)
- peptide sequencing (analysis of tryptic digest by HPLC/electrospray mass spectrometry (LC/MS)
- glycosylation status (*in silico* analysis and glycoprotein detection)

The molecular weights of the proteins were calculated using a regression analysis derived from plotting migration of marker proteins in the SDS-PAGE gel against molecular weight. An approximate molecular weight of 72 kDa was obtained for the Cry1Ab protein and of 20 kDa for the PAT protein. This is considered to be good agreement, within the limitation of analysis, with the actual molecular weights of 69 kDa for Cry1Ab protein and 21 kDa for the PAT protein.

Immunoreactivity was tested by incubating blotted polyvinylidene fluoride membranes with rabbit polyclonal antibodies raised against the appropriate (ie Cry1Ab or PAT) protein followed by incubation with an enzyme linked goat-anti-rabbit secondary antibody. For both proteins there was staining of one major band on the membrane and this demonstrated the immunological relationship of the proteins isolated from *E. coli* with the equivalent native proteins.

The bioactivity of the Cry1Ab protein derived from *E. coli* was confirmed by an insect (*Helicoverpa zea* larvae) feeding assay. The activity of the PAT protein from *E. coli* was determined by a spectrophotochemical method based on the ability of the PAT enzyme to generate free Coenzyme A sulphydryl groups during the transfer of the acetyl group of Acetyl Coenzyme A to L-PPT (D'Halluin *et al*., 2009). Detection of an absorbance 10% above background at 412 nm indicated that the PAT protein was active.

Taken together, the above three analyses confirmed the identity of the proteins obtained from the bacterial expression system as Cry1Ab and PAT. The further analyses described below were designed to provide some characteristics of the proteins upon which comparison with proteins isolated from plants containing event T304-40 could be done.

4.3.1.1 HPLC/Electrospray mass spectrometry (LC/MS)

Peptide maps were produced from the selected ion chromatograms obtained for tryptic digests of the Cry1Ab and PAT proteins isolated from *E. coli*. Based on the theoretical peptide sequences of the Cry1Ab and PAT proteins, it was estimated that the Cry1Ab protein from *E. coli* provided 80% coverage and the PAT protein covered 96%.

4.3.1.2 Glycosylation analysis

N-glycosylated proteins are glycosylated on an asparagine residue and commonly contain an asparagine-X-serine/threonine sequence (N-X-S/T), where X is any amino acid except proline (Orlando and Yang, 1998). Although rare, the sequence asparagine-X-Cysteine (N-X-C) can also be an N-glycosylation site (Miletich and Broze Jr., 1990). The occurrence of these motifs does not, however, indicate that the protein will necessarily be glycosylated and since *E. coli*, like most prokaryotes, lacks the capacity for protein glycosylation (Wacker *et al*., 2002) it would be exceptional to find that N-glycosylation had occurred in the bacterially-derived proteins.

An *in silico* approach was used to search the Cry1Ab and PAT proteins for the occurrence of the two motifs given above. The search of the Cry1Ab protein found six potential sites on the basis of the N-X-S/T consensus sequence and one site based on the N-X-C consensus sequence. No potential N-glycosylation sites were identified for the PAT protein.

Only an experimental approach could confirm whether any of the potential sites identified in the Cry1Ab protein were actual N-glycosylation sites. To assess whether post-translational glycosylation had occurred, glycosylation analysis of a purified Cry1Ab protein sample from *E. coli* was undertaken using a commercially available glycoprotein detection kit. No glycoprotein staining was detected. Although not required, glycosylation analysis of a purified PAT protein sample from *E. coli* was also undertaken and, as expected, no glycostaining was detected.

4.3.2 Protein equivalence

Having established the authenticity and characteristics of the bacterially-derived Cry1Ab and PAT proteins, it was then necessary to confirm whether the plant-derived proteins were equivalent. Initially, this was done by comparing the bacterially-derived proteins with purified Cry1Ab and PAT proteins obtained from ground leaves of 'TwinLink'™ plants. With regard to migration on an SDS-PAGE gel and immunoreactivity the plant and bacterial proteins were indistinguishable. With regard to protein activity, an enzyme assay indicated that the PAT proteins from both *E. coli* and 'TwinLink'™ leaves had similar activity. However, the Cry1Ab protein could not be isolated from 'TwinLink'™ leaves in enough quantity to perform an insect assay.

LC/MS analysis of the Cry1Ab protein from 'TwinLink'™ leaves covered 74% of the theoretical peptide sequences, and of those peptides analysed by selective ion monitoring, 93% were identical to those in the *E. coli*-derived protein. The PAT protein isolated from 'TwinLink'™ leaves covered 87% of the theoretical sequences and 90% of these were identical to those in the *E. coli*-derived protein.

As with the *E.coli*-derived Cry1Ab protein, glycoprotein staining of plant-derived Cry1Ab indicated that none of the potential glycosylation sites was actually glycosylated. The absence of any glycosylation sites in the PAT protein obtained from 'TwinLink'™ leaves was similarly confirmed by the results of glycoprotein staining.

In addition to the above studies, the Applicant also attempted an N-terminal sequencing analysis (Edman degradation) of the Cry1Ab and PAT proteins produced in *E. coli* and 'TwinLink'™ leaves. The analysis did not prove useful because of either blocking of the N-terminal sequence (a common and often unavoidable problem encountered in protein sequencing) or lack of sufficient protein concentration to permit Edman degradation.

To confirm whether the Cry1Ab and PAT proteins produced in 'TwinLink'™ leaves were equivalent to those produced in seeds and leaves of line T304-40 as well as to the proteins produced in *E. coli*, two further minor studies were undertaken. Purified plant-derived Cry1Ab and PAT proteins were obtained from ground leaves and seeds from 'TwinLink'™ and line T304-40 and compared with the bacterially-derived proteins. With regard to migration on an SDS-PAGE gel and immunoreactivity the plant-derived proteins from both tissue types and both GM lines were indistinguishable.

4.3.3 Conclusion

The studies described above allowed the determination of the identity and physicochemical and functional properties of the Cry1Ab and PAT proteins produced by the same event in two tissue types from two GM cotton lines, and allowed comparison of these with the bacterially-produced proteins.

A range of characterisation methods confirmed the identity and non-glycosylated status of Cry1Ab and PAT proteins produced in both a bacterial expression system and in leaves from 'TwinLink'™. From the equivalence shown between the proteins produced in 'TwinLink'™ leaves and in line T304-40, the identity and non-glycosylated status of the proteins produced in line T304-40 is demonstrated.

With regard to the equivalence of proteins produced *in planta* and in *E. coli:*

- The Cry1Ab proteins isolated from both sources were found to be equivalent in all parameters analysed but it was not possible to compare the activity of the Cry1Ab proteins from plant and bacterial sources.

- The PAT proteins isolated from both sources were found to be equivalent in all parameters analysed.

- Based on weight of evidence, it is concluded that microbially-derived Cry1Ab and PAT proteins are suitable surrogates for use in safety assessment studies.

4.4 Potential toxicity of the introduced proteins

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

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

4.4.1 History of human consumption

4.4.1.2 Cry1Ab

The Cry1Ab protein expressed in T304-40 is ubiquitous in the environment (OECD, 2007; Schnepf et al., 1998) and is therefore a natural contaminant of many human foods. No food safety issues have been raised with the use of products from Bt crops (Delaney *et al*., 2008; Mendelsohn *et al*., 2003; OECD, 2007; Shelton *et al*., 2009). In addition, the Cry1Ab protein has been used in microbial formulations for over 30 years. There is no evidence, from this long history of use as a plant pesticide, of any associated toxicity to humans (Kough, 2003).

4.4.1.3 PAT

Streptomyces hygroscopicus and *S. viridochromogenes* are common soil bacteria, therefore humans have a long history of exposure to the PAT protein through the consumption of roots and vegetables.

Since 1995, humans have also been directly exposed to the PAT protein through the consumption of foods derived from glufosinate ammonium tolerant canola, soybean, cotton and corn, without any evidence of toxicity (Delaney *et al*., 2008; Hérouet *et al*., 2005).

4.4.2 Similarities with known protein toxins

Studies submitted:

Rouquie, D. (2007). Cry1Ab Protein: Overall Amino Acid Sequence Homology Search with Known Toxins and Allergens. Study Report No. SA 07186, Bayer CropScience (unpublished).
Herouet-Guicheney, C. (2006). Phosphinothricin Acetyltransferase (PAT) Protein *bar* Gene Product: Overall Amino Acid Sequence Homology Search with Known Toxins and Allergens. Study Report SA 06001, Bayer CropScience (unpublished).

Bioinformatic analyses are useful for assessing whether introduced proteins share any amino acid sequence similarity with known protein toxins. The Cry1Ab (617 amino acids) and PAT (183 amino acids) sequences were compared with protein sequences present in a number of large public reference databases: Uniprot_Swissprot, Uniprot_TrEMBL, PIR (Protein Informatin Resource), PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept.

The similarity searches used the BLASTP (Basic Local Alignment Search Tool Protein) algorithm (Altschul *et al*., 1997) and the BLOSUM62 scoring matrix. BLASTP is now frequently applied for searching for similarities in protein sequences by performing local alignments. This detects more similarities that would be found using the entire query sequence length. A parameter known as the *E* value (see eg Baxevanis, 2005) represents the probability that a particular alignment is due to random chance. Comparisons between highly homologous proteins yield *E*-values approaching zero, indicating the very low probability that such matches would occur by chance. A larger *E*-value indicates a lower degree of similarity. All database sequences with an E-value of 1 or lower were identified by default by the BLASTP program. Although a statistically significant sequence similarity generally requires a match with an *E*-value of less than 0.01 (Pearson, 2000), setting a threshold *E*-value of 1.0 ensures that proteins with even limited similarity will not be excluded. Commonly, for protein-based searches, hits with *E*-values of 10^{-3} or less and sequence identity of 25% or more are considered significant although any conclusions reached need to be tempered by an investigation of the biology behind the putative homology (Baxevanis, 2005).

None of the proteins returned from the BLASTP search with the Cry1Ab protein sequence were associated with known toxic or anti-nutritional properties. The only significant similarities were with other Bt proteins (which are non-toxic). This suggests that the Cry1Ab protein itself is unlikely to be a toxin or anti-nutrient. Similarly, the results of the overall homology search with the PAT protein showed no similarity with known toxins but similarity only with other acetyltransferase proteins (which are non-toxic). This suggests that it is unlikely the *bar* gene encodes known toxic proteins. These results are entirely expected given that there have not been previous toxicity concerns with these two proteins (see eg Delaney *et al*., 2008).

4.5.2 In vitro *digestibility*

See Section 4.6.3.

4.5.3 Thermolability

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

Studies submitted:

Rouquie, D. (2007). Cry1Ab Protein: Heat Stability Study. Study SA07112, Bayer CropScience (unpublished)
Esdaile, D.J. (2002). Phosphinothricin Acetyltransferase (PAT) Bar Gene Product Heat Stability Study. Study SA 02175, Bayer CropScience (unpublished).

4.5.3.1 Cry1Ab

Cry1Ab protein obtained from a bacterial expression system (see Section 4.3) was incubated for 10, 30 or 60 minutes at 60°, 75° or 90° C. The integrity of the protein was detected by gel electrophoresis (SDS-PAGE) coupled with a Coomassie blue staining, or by a Western blot analysis. No significant changes to structure were indicated by SDS-PAGE after heat treatment at 60° for any of the times, at 75° for up to 30 min or at 90° for up to 10 min. Marked but not complete degradation of the protein was noted after incubation at 90°. These results indicate that the Cry1Ab protein is partially heat stable at 90° for 60 min.

The Western blot analysis using an anti-Cry1Ab protein polyclonal primary antibody and an alkaline phosphatase-coupled secondary antibody indicated that immunoreactivity was detectable in all treatments ie conformational changes that may be associated with heat denaturation do not affect epitope accessibility.

4.5.3.1 PAT

PAT protein obtained from a bacterial expression system (see Section 4.3) was incubated for 10, 30 or 60 minutes at 60°, 75° or 90° C. The integrity of the protein was detected by gel electrophoresis (SDS-PAGE) coupled with a Coomassie blue staining. No changes to protein migration were noted in any of the treatments ie the PAT protein is heat stable at 90° for 60 min.

4.5.4 Acute oral toxicity study

An acute oral toxicity study using mice was conducted in order to examine the potential toxicity of the Cry1Ab protein obtained from a bacterial expression system (see Section 4.3).

Study submitted:

Rouquie, D. (2007). Cry1Ab Protein. Acute Oral Toxicity by Oral Gavage in Mice. Study SA 07109, Bayer CropScience (unpublished).

Test material	Cry1Ab preparation from *E. coli*
Vehicle	50 mM Na_2CO_3 (Cry1Ab has limited solubility in H_2O)
Test Species	Crl:OF1 mice (five females) – 7 weeks old on day of treatment
Dose	2 x separate doses of 1000 mg/kg[2] body weight Cry1Ab test substance by oral gavage, within 4 h. Actual total dose 1971.8 mg/kg Cry1Ab since the purity of the protein is 98.59% in the test substance – see Section 4.3)
Control	None

Mice were observed for mortality, body weight gain and clinical signs over 14 days. At the end of the study all animals were killed and examined for organ or tissue damage or dysfunction. All mice survived for the duration of the study. No clinical signs of systemic toxicity were observed. No macroscopic abnormalities were present in the mice at necropsy on day 14. Under the conditions of this study, administration of Cry1Ab protein to female mice at a dose of 1,971 mg /kg bw produced no test substance-related clinical signs of

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

toxicity, body weight losses, macroscopic abnormalities or mortality. These results support the conclusion that the Cry1Ab protein is not acutely toxic.

The Applicant did not supply an acute oral toxicity study for the PAT protein, as expressed in T304-40. However, FSANZ has previously assessed a number of acute oral toxicity studies of the PAT protein and these studies indicate that the PAT protein is unlikely to be toxic to humans.

4.6 Potential allergenicity of the introduced proteins

The potential allergenicity of novel proteins is evaluated using an integrated, step-wise, case-by-case approach relying on various criteria used in combination, since no single criterion is sufficiently predictive of either allergenicity or non-allergenicity (see eg Thomas *et al*., 2009). The assessment focuses on: the source of the novel protein; any significant amino acid sequence similarity between the novel protein and known allergens; the structural properties of the novel protein, including susceptibility to digestion, heat stability and/or enzymatic treatment; and specific serum screening if the novel protein is derived from a source known to be allergenic or has amino acid sequence similarity with a known allergen. In some cases, such as where the novel protein has sequence similarity to a known allergen, additional *in vitro* and *in vivo* immunological testing may be warranted. Applying this approach systematically provides reasonable evidence about the potential of the novel protein to act as an allergen.

The allergenic potential of Cry1Ab and PAT proteins was assessed by:

- consideration of the source of the gene encoding each protein and history of use or exposure
- bioinformatic comparison of the amino acid sequence of the Cry1Ab and PAT proteins with known protein allergen sequences
- evaluation of the lability of the microbially produced and purified Cry1Ab and PAT proteins from *E. coli* using *in vitro* gastric and intestinal digestion models; and thermolability

4.6.1 Source of each protein

As described in Section 4.4.1, both the Cry1Ab protein and PAT protein are derived from common soil bacteria to which humans have been naturally exposed and which may have been inadvertently ingested on fresh produce without eliciting adverse effects. Neither *Bacillus. thuringiensis* nor *Streptomyces hygroscopicus* is considered to be a source of allergenic proteins (see eg EFSA, 2007; OECD, 2007).

4.6.2 Similarity to known allergens

Studies submitted:

Rouquie, D. (2007). Cry1Ab Protein: Overall Amino Acid Sequence Homology Search with Known Toxins and Allergens. Study Report No. SA 07186, Bayer CropScience (unpublished).

Rouquie, D. (2007). Cry1Ab protein: epitope homology and N-glycosylation searches. Study SA07187, Bayer CopScience (Unpublished).

Herouet-Guicheney, C. (2006). Phosphinothricin Acetyltransferase (PAT) Protein *bar* Gene Product: Overall Amino Acid Sequence Homology Search with Known Toxins and Allergens. Study Report SA 06001, Bayer CropScience (unpublished).

Herouet-Guicheney, C. (2006). Phosphinothricin acetyl transferase (PAT) BAR gene product: epitope homology and N-glycosylation searches. Study SA 06000, Bayer CropScience (unpublished).

Bioinformatic analysis provides part of a 'weight of evidence' approach for assessing potential allergenicity of novel proteins introduced to GM plants (Goodman, 2006; Thomas *et al*., 2005). It is a method for comparing the amino acid sequence of the introduced protein with sequences of known allergens in order to indicate potential cross-reactivity between allergenic proteins and the introduced protein. As with the bioinformatic analysis that looked at similarities of Cry1Ab and PAT with known protein toxins (see Section 4.4.2), the generation of an *E* value provides an important indicator of significance of matches (Baxevanis, 2005; Pearson, 2000).

The Cry1Ab (617 amino acids) and PAT (183 amino acids) sequences were compared with protein sequences present in a number of large public reference databases: Uniprot_Swissprot, Uniprot_TrEMBL, PIR (Protein Informatin Resource), PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept (see Section 4.4.2). The criterion used to indicate potential allergenicity was a minimum of 35% identity on a window of 80 amino acids within the sequence of an allergenic protein. In addition, in order to identify any short homologous amino acid sequences that may represent allergenic epitopes, the Cry1Ab and PAT proteins were subdivided into eight linearly contiguous amino acid blocks that were compared with all known allergen sequences contained in a reference database.

For the Cry1Ab protein, homology was only observed with sequences of other Cry proteins (which are non-allergenic). No matches were detected with known allergenic epitopes. It was concluded that it is unlikely the Cry1Ab protein is allergenic.

Similarly, for the PAT protein, homology was only observed with sequences of other acetyltransferases (which are non-allergenic) from various sources. No matches were detected with known allergenic epitopes. It was concluded that it is unlikely the PAT protein is allergenic.

4.6.3 In vitro *digestibility*

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

A pepsin digestibility assay was conducted to determine the digestive stability of the Cry1Ab and PAT proteins. In addition to the pepsin protocol using simulated gastric fluid (SGF), a second digestibility study was done using simulated intestinal fluid (SIF) containing pancreatin, which is a mixture of enzymes including amylase, trypsin, lipase, ribonuclease and protease. The relevance of the SIF study however is limited because ordinarily an ingested protein would first be exposed to pepsin-mediated hydrolysis in the acidic environment of the stomach before being subject to further digestion in the small intestine.

4.6.3.1 Simulated gastric fluid (SGF) studies

Studies submitted:

Rouquie, D. (2007). Cry1Ab Protein *In Vitro* Digestibility Study in Human Simulated Gastric Fluid. Study SA 07110, Bayer CropScience (unpublished).
Rouquie, D. (2002). Phosphinothricin acetyltransferase (PAT) *bar* Gene Product In Vitro Digestibility Study in Simulated Gastric Fluid. Study SA 02173, Bayer CropScience (unpublished).

The *in vitro* digestibility of the *E. coli*-derived Cry1Ab and PAT proteins in SGF (U.S.Pharmacopeia, 1990) containing pepsin was evaluated by SDS-PAGE. Essentially the methods used for each protein were the same although the method for the Cry1Ab protein, performed some 5 years after that for the PAT protein, closely followed that recommended by Thomas et al (Thomas *et al*., 2004) who noted the desirability of using a protocol that is standard across different laboratories. Digestibility of the proteins in SGF was measured by incubating samples at 37º for selected times (0, 0.5, 1, 2, 5, 10, 20, 30 and 60 minutes) and subjecting these to SDS-PAGE. Proteins were visualised by Coomassie staining of the resulting gel.

Two control proteins were treated in parallel: horseradish peroxidase (HRP) is known to hydrolyse readily in pepsin and served as a positive control; ovalbumin (OVA) is known to persist in pepsin and was used as a negative control.

The Cry1Ab protein was rapidly hydrolysed in SGF, with complete digestion after two minutes exposure. In the method used for the Cry1Ab analysis, the HRP positive control was hydrolysed within five minutes while the OVA negative control persisted for 60 minutes.

The PAT protein was rapidly hydrolysed in SGF, with complete digestion after 30 seconds exposure. In the method used for the PAT analysis, the HRP positive control was also rapidly hydrolysed (< 30 seconds) while the OVA negative control persisted for 60 minutes.

4.6.3.2 Simulated intestinal fluid (SIF) studies

Studies submitted:

Rouquie, D. (2007). Cry1Ab Protein *In Vitro* Digestibility Study in Simulated Intestinal Fluid. Study SA 07111, Bayer CropScience (unpublished).
Esdaile, D.J. (2002). Phosphinothricin acetyltransferase (PAT) bar Gene Product In Vitro Digestibility Study in Simulated Intestinal Fluid. Study SA 02174, Bayer CropScience (unpublished).

The digestibility of *E. coli*-derived Cry1Ab and PAT proteins in SIF containing pancreatin (U.S.Pharmacopeia, 1990) was assessed using SDS-PAGE. Digestibility of the protein in SIF was measured by incubating samples with SIF containing porcine pancreatin, for specified time intervals (0, 0.5, 2, 5, 10, 20, 30 and 60 minutes), and analysing by SDS-PAGE with Coomassie staining, and also Western blot analysis.

No visible Cry1Ab band at 69 kDa was observed following SDS-PAGE or Western blot analysis at incubation times of 30 seconds or greater. However a number of bands with a range of molecular weights were observed across all the incubation times, indicating the formation of stable degradation products. It is therefore concluded that there is a rapid (< 30 seconds) but incomplete degradation of Cry1Ab by pancreatin.

The PAT protein was digested within seconds after contact with SIF in the presence of pancreatin with all residual fragments being completely digested within five minutes.

4.6.4 Thermolability

See Section 4.5.3

4.7 Conclusion

Cotton line T304-40 express two novel proteins, Cry1Ab and PAT. Expression analyses of the two proteins showed that PAT was expressed in all plant parts tested but was highest in young leaves (average of 61.4 µg/g fresh weight). Cry1Ab was not detectable in any plant parts except the seed (average of 3.7 µg/g fresh weight).

A number of studies were done to confirm the identity and physicochemical and functional properties of the expressed plant-derived Cry1Ab and PAT proteins, as well as to determine their potential toxicity and allergenicity. These studies have demonstrated that the Cry1Ab and PAT proteins conform in size and amino acid sequence to that expected, do not exhibit any post-translational modification including glycosylation and, in the case of PAT, exhibit the expected enzymatic activity. The bioactivity of the expressed Cry1Ab protein was unable to be tested.

In relation to potential toxicity and allergenicity, it is worth noting that the Cry1Ab and PAT proteins are inherently non-toxic to mammals and do not exhibit any potential to be allergenic to humans. In addition, bioinformatic studies have confirmed their lack of any significant amino acid sequence similarity to known protein toxins or allergens and digestibility studies have demonstrated that both proteins would be rapidly degraded in the stomach following ingestion. Acute oral toxicity studies in mice have also confirmed their absence of toxicity in animals. Both proteins exhibit a degree of heat stability however given their digestive lability, this does not raise any safety concerns. Taken together, the evidence indicates that Cry1Ab and PAT are unlikely to be toxic or allergenic to humans.

5. COMPOSITIONAL ANALYSES

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

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

5.1 Key components

Fuzzy or whole cottonseed (the linted cottonseed remaining after the ginning process which removes fibres for textile production) is the most usual source of food products derived from cotton. Cottonseed oil is the primary cotton product used for human consumption. For cotton, the key components that should be considered in the comparison include proximates

(cottonseed only), fatty acids, tocopherol, gossypol and the cyclopropenoid fatty acids (OECD, 2004). The Applicant also undertook analyses for amino acid, minerals and phytic acid content, which are of greater relevance for animal feed.

5.2 Study design and conduct for key components

Studies submitted:

Oberdörfer, R. (2008). General Information for the Nutritional Impact Assessment of GM Cotton Traits. Report No. 08 B 013, Bayer CropScience (unpublished).

Oberdörfer, R. (2009). Nutritional Impact Assessment Report on Insect-Resistant Cotton (Event T304-40). Report No. 08 B 011, Bayer CropScience (unpublished).

Martone, A. (2008). Analyses of Raw Agricultural Commodity (Fuzzy Seed) of Cry1Ab Cotton Event T304-40 for PAT/*bar* and Cry1Ab and its Non-Transgenic Counterpart for PAT/*bar* and Cry1Ab Proteins. Report No. CP07B011, Bayer CropScience (unpublished).

Mackie, S.J.W. (2008). Production of Processed Commodities of Transgenic Cry1Ab Cotton and the Non-Transgenic Counterpart, USA 2007. Study CP07B008, Bayer CropScience (unpublished).

Ideally, the comparator in compositional analyses should be the near isogenic parental line grown under identical conditions. In the case of cotton line T304-40, the generation selected for testing was BC2F4 (see Figure 2) and the control was the non-GM parental line 'Coker 315'. Both lines were grown in 2007 at eight sites across three regions in North America, representative of the range of growing regions for cotton[3]. Both the seed that was planted and the seed that was subsequently harvested, was analysed to verify seed identity. The non-transgenic 'Coker 315' seed planted in the field contained less than 0.1% of T304-40 seed with 95% confidence as well as containing less than 0.1% of other traits/lines (Bollgard, Bollgard II, Liberty and Roundup Ready). T304-40 seed was free of other traits/lines (MON531, Bollgard II and Roundup Ready) at a level of 0.1%. Extracts of raw agricultural commodity (RAC) of 'Coker 315' harvested from the trials were analysed by discriminating PCR for the presence of T304-40; only endogenous PCR product was found thereby confirming that the samples were harvested from non-GM plots. Similarly, all extracts of RAC from plants notionally containing T304-40 were shown to contain DNA sequences from T304-40. In addition to the above testing, all plant material was handled and stored so as to minimise the likelihood of cross-contamination.

Plants were grown under conditions typical of production practices. At each site, three plots of 'Coker 315' and six plots of T304-40 were planted in a randomised block design; three of the GM plots were sprayed two times during the growth phase with glufosinate ammonium herbicide. In total this meant that for each 'treatment' there were 3 (plots) x 8 (sites) = 24 readings. Plots were harvested by hand, or with a cotton picker, from the interior rows of each plot and the cotton was ginned at the trial site. Methods of composition analysis were based on internationally recognised procedures (e.g. those of the Association of Official Analytical Chemists), methods specified by the manufacturer of the equipment used for analysis, or other published methods.

A total of 72 fuzzy cottonseed samples provided results for 42 analytes. These results were statistically analysed using analysis of variance. The results for generation BC2F4 and 'Coker 315' were compared to a combined literature range for each analyte, compiled by the Applicant from published literature for commercially available cottonseed[4]. Any statistically significant differences between T304-40 and the 'Coker 315' control were compared to the reference range to assess whether the differences were likely to be biologically meaningful.

[3] Sites were located in St. Landry - 2 sites(Louisiana), Tate - 2 sites(Mississippi), Uvalde, (Texas), Wharton (Texas) and Hockley – 2 sites (Texas)

[4] Published literature for cotton included OECD (2004), ILSI (2007) and Codex (Codex, 2001).

5.3 Analyses of key components

Compositional analyses of the fuzzy cottonseed included proximates (crude protein, crude fat, ash and total carbohydrates), acid detergent fibre (ADF), neutral detergent fibre (NDF), fatty acids, amino acids, micronutrients (minerals and α-tocopherol) and anti-nutrients (gossypol, phytic acid and cyclopropenoid fatty acids). For a number of analytes, 'over all sites' analysis was not valid because there were significant site x treatment interactions. For these analytes, analysis was therefore done on a 'by site' basis. The results presented in the following tables show means pooled from all sites. The range, given as a calculated variable, denotes the lowest and highest individual value obtained for any plot within the sites.

5.3.1 Proximates and fibre

Results of the proximate and fibre analysis of the cottonseed are shown in Table 4. Statistically significant differences between 'Coker 315' and T304-40 (both sprayed and unsprayed) were noted for ash but the means were all within the range reported in the literature. No statistically significant differences were observed between T304-40 and the control mean values for other proximates and fibre.

Table 4 Percentage dry weight (dw) of proximates and fibre in fuzzy cottonseed from 'Coker 315' and T304-40.

Analyte	Calculated variable	'Coker 315' (% dw)	T304-40 not glufosinate sprayed (%dw)	T303-40 - glufosinate sprayed (%dw)	Combined literature range (%dw)
Protein (%dw)	Mean	24.06	24.68	24.94	11.7 – 34.2
	Range	21.0 – 26.9	21.6 – 29.4	22.4 – 27.6	
Fat (%dw)	Mean	18.25	18.33	17.95	11.8 – 36.3
	Range	16.1 – 20 .9	15.6 – 20.5	15.6 – 20.4	
Carbohydrate (%dw)[1]	Mean	54.01	53.06	53.25	36.4 – 74.4
	Range	50.7 – 59.4	48.0 – 57.9	49.0 – 57.6	
ADF (%dw)	Mean	40.79	40.67	41.20	29.0 – 66.9
	Range	32.8 – 51.1	33.3 – 47.1	30.8 – 46.8	
NDF (%dw)	Mean	47.81	45.31	46.41	38.1 – 71.4
	Range	43.0 – 55.7	40.8 – 50.1	40.7 – 54.7	
Ash (%dw)	Mean	3.68	3.95	3.87	3.2 – 5.0
	Range	3.0 – 4.3	3.1 – 4.9	3.3 – 4.7	

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

5.3.2 Fatty Acids

The levels of 10 fatty acids in fuzzy cottonseed from T304-40 and in 'Coker 315' control seed were measured. Results of the analysis are given in Table 5 and can be summarised as follows:

- For arachidic and behenic acids, across the majority of sites, there was no significant difference between the means for T304-40 seeds and the means for 'Coker 315' seeds. Note, however, that for behenic acid, three readings (of a possible 24) were below the limit of quantification in non-sprayed T304-40 plants.
- For stearic acid the means for T304-40 seeds were significantly higher at all sites than means for 'Coker 315' seeds but fell within the range found in the published literature.

- For myristic, palmitic and palmitoleic acids the means for T304-40 seeds were significantly lower at the majority of sites than the mean for 'Coker 315' seeds. The palmitoleic acid means for T304-40 while outside the range found in the published literature, do not raise a concern because of the low contribution (<0.5%) of palmitoleic acid to total fatty acids.
- For oleic, linoleic and linolenic acids the results across all sites were inconsistent ie at approximately half of the sites results were significant, while at the other half results were not significant. However, all values (both mean and ranges) fell within the range found in the published literature.
- For lignoceric acid most readings were below the limit of quantification for both 'Coker 315' and for T304-40 seeds.

Table 5. Percentage composition, relative to total fat, of major fatty acids in fuzzy cottonseed from 'Coker 315' and T304-40.

Fatty acid	Calculated variable	'Coker 315' (% total)	T304-40 not glufosinate sprayed (% total)	TT303-40 (glufosinate sprayed) (%total)	Combined literature range (% total)
Myristic acid (C14:0)	Mean	0.76	0.66	0.65	0.53 – 1.17
	Range	0.61 – 0.89	0.53 – 0.79	0.55 – 0.79	
Palmitic acid (C16:0)	Mean	24.12	23.4	23.4	21.1 – 29.9
	Range	23.2 – 25.3	22.7 – 24.5	22.6 – 24.5	
Stearic acid (C18:0)	Mean	2.48	2.71	2.76	2.15 – 3.4
	Range	2.34 – 2.70	2.44 – 3.17	2.44 – 3.14	
Arachidic acid (C20:0)	Mean	0.29	0.29	0.29	0.2 - 0.48
	Range	0.27 – 0.33	0.26 – 0.33	0.27 – 0.33	
Behenic acid (C22:0)	Mean	0.19	Not calculated[1]	0.17	0.1 - 0.27
	Range	0.15 – 0.21	< 0.12 – 0.20	0.14 -0.19	
Lignoceric acid (C24:0)	Mean	Not calculated[1]	Not calculated[1]	Not calculated[1]	0.11 - 0.3
	Range	<0.12 – 0.24	<0.12 – 0.13	<0.12 – 0.12	
Palmitoleic acid (C16:1)	Mean	0.48	0.44	0.43	0.46 – 0.86
	Range	0.44 – 0.54	0.39 – 0.47	0.40 – 0.46	
Oleic acid (C18:1)	Mean	15.25	15.15	14.95	13.4 – 22.0
	Range	14.1 – 16.4	14.1 – 16.3	13.8 – 16.1	
Linoleic acid (C18:2)	Mean	56.18	56.96	57.14	36.3 – 64.0
	Range	54.6 – 58.5	55.5 – 58.8	55.1 – 59.1	
Linolenic acid (C18:3)	Mean	0.21	0.21	0.22	0.08 - 0.62
	Range	0.18 – 0.32	0.20 – 0.23	0.19 – 0.23	

[1]Mean not calculated because there were some readings below the limit of quantification

5.3.3 Amino acids

Levels of 18 amino acids were measured in seed from 'Coker 315' and T304-40 seed. Since asparagine and glutamine are converted to aspartate and glutamate respectively during the analysis, levels for aspartate include both aspartate and asparagine, while glutamate levels include both glutamate and glutamine. Results of the analysis are given in Table 6. While there were significant differences at some sites, at the majority of sites, there was no significant difference in the mean level of any amino acid between seeds from T304-40 and those from 'Coker 315'. At all sites, the range of values obtained for each amino acid was within the range found in the combined literature.

Table 6. Percentage dry weight (dw), relative to total dry weight, of amino acids in fuzzy cottonseed from 'Coker 315' and T304-40.

Amino acid	Calculated variable	'Coker 315' (% dw)	T304-40 not glufosinate sprayed (% dw)	T303-40 (glufosinate sprayed) (% dw)	Combined literature range (% dw)
Alanine	Mean	0.8	0.83	0.89	0.42 – 1.51
	Range	0.65 – 0.96	0.67 – 1.00	0.69 – 1.05	
Arginine	Mean	2.44	2.58	2.77	1.05 – 4.4
	Range	1.96 – 3.02	1.95 – 3.19	2.10 – 3.45	
Aspartic acid	Mean	1.93	2.10	2.17	1.0 – 3.55
	Range	1.64 – 2.29	1.65 – 2.4	1.78 – 2.54	
Cysteine	Mean	0.38	0.39	0.39	0.16 – 0.86
	Range	0.29 – 0.43	0.26 – 0.46	0.32 – 0.44	
Glutamic acid	Mean	4.02	4.2	4.52	1.96 – 8.16
	Range	3.30 – 4.93	3.31 – 5.26	3.42 – 5.48	
Glycine	Mean	0.86	0.89	0.95	0.44 – 1.58
	Range	0.71 – 1.03	0.72 – 1.05	0.73 – 1.13	
Histidine	Mean	0.58	0.6	0.64	0.31 – 1.03
	Range	0.48 – 0.7	0.48 – 0.74	0.51 – 0.77	
Isoleucine	Mean	0.67	0.7	0.74	0.35 – 1.17
	Range	0.56 – 0.82	0.57 – 0.83	0.56 – 0.9	
Leucine	Mean	1.21	1.26	1.34	0.63 – 2.23
	Range	1.0 – 1.47	1.08 – 1.52	1.03 – 1.61	
Lysine	Mean	0.94	0.97	1.03	0.52 – 1.65
	Range	0.79 – 1.12	0.79 – 1.16	0.81 – 1.22	
Methionine	Mean	0.32	0.32	0.32	0.15 – 0.54
	Range	0.23 – 0.39	0.20 – 0.39	0.26 – 0.39	
Phenylalanine	Mean	1.13	1.18	1.27	0.54 – 2.03
	Range	0.93 – 1.38	0.94 – 1.45	0.96 – 1.54	
Proline	Mean	0.76	0.77	0.83	0.41 – 1.39
	Range	0.63 – 0.94	0.63 – 0.92	0.61 – 1.07	
Serine	Mean	0.9	0.9	1.0	0.5 – 1.63
	Range	0.72 – 1.1	0.73 – 1.17	0.72 – 1.22	
Threonine	Mean	0.66	0.69	0.73	0.34 – 1.21
	Range	0.54 – 0.8	0.56 – 0.83	0.57 – 0.85	
Trytophan	Mean	0.2	0.21	0.22	0.1 – 0.49
	Range	0.17 – 0.25	0.17 – 0.27	0.17 – 0.26	
Tyrosine	Mean	0.55	0.57	0.61	0.32 – 1.17
	Range	0.44 – 0.68	0.45 – 0.69	0.49 – 0.71	
Valine	Mean	0.94	0.98	1.04	0.45 – 1.67
	Range	0.78 – 1.16	0.79 – 1.22	0.88 – 1.29	

5.3.4 Anti-nutrients

Levels of key anti-nutrients in seeds from T304-40 and 'Coker 315' are given in Table 7. While there were significant differences at some sites, at the majority of sites, there was no significant difference in the mean level of free gossypol, total gossypol, malvalic acid and sterculic acid between seeds from T304-40 and those from 'Coker 315'. At all sites, the range of values obtained for each anti-nutrient was within the range found in the combined literature.

The level of phytic acid was significantly higher in seeds from T303-40 than from 'Coker 315' but the level was not outside that found in the combined literature range for this analyte. For dihydrosterculic acid, the mean for seeds from sprayed T304-40 plants was significantly lower than that for seeds from unsprayed T304-40 plants or from 'Coker 315'. However, the level was not outside that found in the combined literature range for this analyte.

Table 7. Levels of anti-nutrients in fuzzy cottonseed from 'Coker 315' and T304-40.

Anti-nutrient	Calculated variable	'Coker 315'	T304-40 not glufosinate sprayed	T303-40 (glufosinate sprayed)	Combined literature range
Free gossypol (% dw)	Mean	0.62	0.60	0.61	0.23 – 1.4
	Range	0.29 – 0.74	0.46 – 0.81	0.44 – 0.78	
Total gossypol (% dw)	Mean	0.69	0.69	0.69	0.46 – 1.99
	Range	0.52 – 0.85	0.52 – 0.86	0.54 – 0.89	
Phytic acid (% dw)	Mean	1.38	1.51	1.49	0.85 – 2.57
	Range	0.82 – 1.87	0.94 – 1.96	0.9 – 1.85	
Malvalic acid (%tot fat)	Mean	0.45	0.47	0.46	0.17 – 1.5
	Range	0.37 – 0.52	0.39 – 0.56	0.39 – 0.58	
Sterculic acid (% tot fat)	Mean	0.32	0.31	0.29	0.12 – 0.92
	Range	0.24 – 0.39	0.23 – 0.38	0.21 – 0.38	
Dihydrosterculic acid (% tot fat)	Mean	0.16	0.16	0.14	0.11 – 0.5
	Range	0.13 – 0.19	0.13 - .21	0.10 – 0.16	

5.3.5 Minerals and vitamins

Levels of key minerals and vitamins in seeds from T304-40 and 'Coker 315' are given in Table 8. At all sites, there was no significant difference in the mean level of potassium between seeds from T304-40 and those from 'Coker 315'. The results for the other analytes were somewhat inconsistent across the sites but in all cases, both the means and the individual ranges obtained fell within the published literature range.

Table 8. Levels of key minerals and vitamins in fuzzy cottonseed from 'Coker 315' and T304-40.

Analyte	Calculated variable	'Coker 315'	T304-40 not glufosinate sprayed	T303-40 (glufosinate sprayed)	Combined literature range
Calcium (% dw)	Mean	0.14	0.19	0.18	0.09 – 0.33
	Range	0.11 – 0.16	0.14 – 0.23	0.15 – 0.24	
Phosphorus (% dw)	Mean	0.53	0.57	0.55	0.31 – 0.86
	Range	0.35 – 0.70	0.34 – 0.75	0.35 – 0.71	
Potassium (% dw)	Mean	1.11	1.12	1.12	0.96 – 1.42
	Range	0.98 – 1.29	0.96 – 1.34	0.94 – 1.31	
Magnesium (% dw)	Mean	0.34	0.37	0.36	0.27 – 0.49
	Range	0.28 – 0.39	0.29 – 0.44	0.30 – 0.41	
Iron (mg/kg dw))	Mean	42.0	47.1	44.1	23.2 – 160.0
	Range	26.8 – 62.3	30.2 – 65.1	29.3 – 65.7	
Zinc (mg/kg dw)	Mean	31.7	36.3	35.6	17.8 – 63.0
	Range	23.8 – 42.5	28.0 – 48.9	27.8 – 45.1	
α-tocopherol (mg/kg dw)	Mean	122.0	119.0	124.0	16.0 – 245.0*
	Range	79 - 157	88 - 158	90.8 - 171	

* Reference range calculated from crude oil figure in Codex (Codex, 2001) and converted to mg/kg/dw based on a seed fat content of 11.8 – 36.3% dw.

5.3.6 Analytes in processed commodities

In addition to analysing the fuzzy cottonseed, the Applicant supplied comparative compositional data for a range of processed commodities derived from cottonseed ie lint, linters, hulls, delinted seeds, meal, toasted meal, crude oil and refined deodorised oil. The cotton that was processed to obtain these various products was grown at a single site (Uvalde, Texas) at which there was a single plot of 'Coker 315' and a single plot of T304-40 plants, sprayed twice with a glufosinate ammonium herbicide. Measures to test for, and ensure purity of, both the planted and harvested seed were followed as described in Section 5.2.

FSANZ has considered the results from these analyses but the data are not presented in this Assessment. No statistical analyses of the compositional results were done because of the low sample number. For lint and linters there was also no published reference range with which to compare the means. For some analytes, particularly the amino acids and fat, ADF and NDF in cottonseed meal and toasted cottonseed meal, the mean levels for both the 'Coker 315' control and T304-40 fell outside the published reference range. Overall, there were no large discrepancies between the control and the GM line for the means of any analyte.

5.4 Conclusion from compositional analysis

Detailed compositional analyses were done of fuzzy seed-derived from T304-40 plants. Analyses were done of proximates (crude protein, crude fat, ash and total carbohydrates), ADF, NDF, fatty acids, amino acids, micronutrients (minerals and α-tocopherol) and anti-nutrients (gossypol, phytic acid and cyclopropenoid fatty acids). The levels were compared to levels in the non-GM parent as well as to the ranges found in commercial cotton cultivars reported in the literature. Additionally, levels of analytes were measured in processed commodities derived from control and GM cottonseed, although the results from these commodities were not analysed statistically.

For fuzzy cottonseed, across most of the categories several significant differences in analytes between the GM and non-GM seeds were found. The composition of cotton can vary significantly with the site, agricultural conditions and season of production, and differences reported here most likely reflect normal biological variability. Additionally, with the exception of palmitoleic acid, the mean analyte levels found in seeds from T304-40 fell within the range of natural variation in commercial cotton cultivars. The lower level of palmitoleic acid in seeds from T304-40 does not raise a concern because of the low contribution (<0.5%) of palmitoleic acid to total fatty acids.

For processed products derived from cottonseed there were no large discrepancies between the control and the GM line for the means of any analyte.
Taken overall, the compositional data are consistent with the conclusion that there are no biologically significant differences in the levels of key components in seed from cotton line T304-40 when compared with conventional cotton cultivars currently on the market.

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 an understanding of the genetic modification and its consequences, together with an extensive compositional analysis of the food.

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

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

Cotton line T304-40 is the result of a simple genetic modification to confer insect protection and herbicide tolerance with no intention to significantly alter nutritional parameters in the food. In addition, extensive compositional analyses have been undertaken to demonstrate the nutritional adequacy of T304-40 and these indicate it is equivalent in composition to conventional cotton varieties. The introduction of cotton line T304-40 into the food supply is therefore expected to have little nutritional impact.

The Applicant submitted one animal feeding study with cotton T304-40, the results of which are included below.

6.1 Broiler feeding study

Studies submitted:

Stafford, J.M. (2009). Broiler Chicken Nutritional Equivalency Study with T304-40 Cotton. Project No. TX99X071, Bayer CropScience (unpublished).

Mackie, S.J.W. (2008). Production of Cottonseed Samples of Insect-Resistant Cotton, the Non-Transgenic Counterpart and a Non-Transgenic Commercial Cotton Variety, USA 2007. Study No. CP07B002, Bayer CropScience (unpublished).

This 42-day study compared growth, performance and carcass yield of Ross #308 broiler chickens fed diets containing approximately 10% toasted cottonseed meal from seeds of cotton T304-40 (sprayed with glufosinate ammonium herbicide) from generation BC2F4 (see Figure 2) with those fed diets containing approximately 10% cottonseed meal obtained from one near-isogenic line ('Coker 315') and one non-GM commercial cultivar ('FiberMax® 958'). Plants of the three seed types were grown in 2007 in two field trials in southern US (Texas) to produce bulk cottonseed samples of each type. DNA characterisation, using discriminating PCR analysis, was undertaken of both the raw commodity and the test diets derived from the raw commodity to validate purity.

Broilers were housed 10 broilers per pen (replicate) with 14 replicates (7 replicates per gender) per treatment to give 140 broilers in each of the 3 cotton line treatments and a total of 420 birds.

Diets were formulated to meet nutrient requirements of a typical commercial broiler diet and were fed in three phases (Starter: 0 – day 7; Grower: day 8 – day 21; and Finisher: day 22 – day 42) according to standard commercial poultry farming practice. The diets were also

designed to be isoenergetic, isoproteic and as similar as possible relative to limiting amino acids in terms of both the cottonseed meal source and the broiler growth phase. Feed and drinking water were available *ad libitum* throughout the study; the feed was weighed and refreshed at least weekly.

Birds were observed at least daily for overall health, behaviour and/or evidence of toxicity. Body weights were determined on days 7, 21, 35 and 42 and feed consumption was calculated for each pen on a weekly basis and converted to mean feed consumption per bird. At study termination, 21 birds/gender/treatment (total of 126 birds) were randomly selected and processed to collect carcass and carcass part yield data.

Analysis of variance was used to analyse two data sets, the first containing those data based on cage average (survival; feed consumption; and feed:body weight conversion ratio) and the second containing individual data (live body weight; chilled carcass weight; abdominal fat pad; breast, thigh, wing and leg weights; tissue conversion ratio).

No statistically significant effects were detected in any of the data that were attributable to the consumption of cottonseed meal derived from cotton T304-40. Broilers consuming a diet containing T304-40 cottonseed meal demonstrated health and growth characteristics comparable to broilers consuming cottonseed meal diets derived from non-GM cotton.

Based on the results from this study, it was concluded that cottonseed meal from cotton T304-40 was nutritionally adequate, and equivalent to that derived from a non-GM control cotton and a commercial non-GM cultivar, in its ability to support typical growth and well being.

REFERENCES[5]

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

Aronson, A.I. and Shai, Y. (2001) Why *Bacillus thuringiensis* insecticidal toxins are so effective: unique features of their mode of action. *FEMS Microbiology Letters* 195:1-8.

Astwood, J.D. and Fuchs, R.L. (1996) Allergenicity of foods derived from transgenic plants. *Highlights in food allergy. Monographs in Allergy*, 32. 105-120.

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.

Boevink, P., Chu, P.W.G. and Keese, P. (1995) Sequence of subterranean clover stunt DNA: affinities with the geminiviruses. *Virology* 207:354-361.

Bradbury, J.F. (1986) *Guide to Plant Pathogenic Bacteria*. Cambridge News Ltd, Aberystwyth, p190-197.

Bravo, A., Gill, S.S. and Soberón, M. (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* 49:423-435.

[5] All website references were current as at 23 November 2009

Broderick, N.A., Raffa, K.F. and Handelsman, J. (2006) Midgut bacteria required for *Bacillus thuringiensis* insecticidal activity. *Proceedings of the National Academy of Sciences* 103:15196-15199.

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

Cousins, Y.L., Lyon, B.R. and Llewellyn, D.J. (1991) Transformation of an Australian cotton cultivar: prospects for cotton improvement through genetic engineering. *Australian Journal of Plant Physiology* 18:481-494.

D'Halluin, K., De Block, M., Denecke, J., Janssens, J., Leemans, J., Reynaerts, A. and Botterman, J. (2009) The *bar* gene as selectable and screenable marker in plant engineering. *Methods in Enzymology* :415-426.

de Maagd, R.A., Bravo, A. and Crickmore, N. (2001) How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends in Genetics* 17:193-199.

Deblaere, R., Reynaerts, A., Höfte, H., Hernalsteens, J.-P., Leemans, J. and van Montagu, M. (1987) Vectors for cloning in plant cells. *Methods in Enzymology* 153:277-292.

Delaney, B., Astwood, J.D., Cunny, H., Eichen Conn, R., Herouet-Guicheney, C., MacIntosh, S., Meyer, L.S., Privalle, L., Gao, Y., Mattsson, J., Levine, M. and ILSI. (2008) Evaluation of protein safety in the context of agricultural biotechnology. *Food and Chemical Toxicology* 46:S71-S97.

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

EFSA. (2007) Opinion of the Scientific Panel on Genetically Modified Organisms on an Application (Reference EFSA-GMO-UK-2004-04) for the Placing on the Market of Glufosinate Tolerant Genetically Modified Rice LLRICE62 for Food and Feed Uses, Import and Processing under Regulation (EC) No 1829/2003 from Bayer CropScience GmbH. *The EFSA Journal* 588:1-25.

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, available online at http://www.efsa.europa.eu/cs/BlobServer/Non_Scientific_Document/gmo_report_feedingtrials.pdf?ssbinary=true.

Gill, S.S., Cowles, E.A. and Pietrantonio, P.V. (1992) The mode of action of *Bacillus thuringiensis* endotoxins. *Annual Review of Entomology* 37:615-636.

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.

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

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

Hofmann, C., Vanderbruggen, H., Höfte, H., Van Rie, J., Jansens, S. and Van Mellaert, H. (1988) Specificity of *Bacillus thuringiensis* δ-endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target inscet midguts. *Proceedings of the National Academy of Sciences* 85:7844-7848.

Höfte, H., de Greve, H., Seurinck, J., Jansens, S., Mahillon, J., Ampe, C., Vandekerckhove, J., Vanderbruggen, H., van Montagu, M., Zabeau, M. and Vaeck, M. (1986) Structural and functional analysis of a cloned delta endotoxin of *Bacillus thuringiensis berliner* 1715. *European Journal of Biochemistry* 161:273-280.

Höfte, H. and Whiteley, H.R. (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiological Reviews* 53:242-255.

ILSI (2007) *International Life Sciences Institute Crop Composition Database Version 3.0*. http://www.cropcomposition.org/cgi-perl/search_ora.cgi.

Jansens, S., Van Houdt, S. and Reynaerts, A. (2006) Plants made insect resistant by transformation with a nucleic acid encoding a modified Cry1Ab protein and method for making same, (US 7,049,491 B2): PatentStorm http://www.patentstorm.us/patents/7049491.html.

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

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.

Kough, J. (2003) The safety of *Bacillus thuringiensis* for human consumption. In: Metz, M. eds. *Bacillus thuringiensis: A Cornerstone of Modern Agriculture*. Chapter 1. Food Products Press, pp. 1-10.

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

Lusas, E.W., Rhee, K.C. and Koseoglu, S.S. (1989) Status of Vegetable Food Proteins from Lesser-Used Sources. In: Applewhite, T.H. eds. *Proceedings of the World Congress on Vegetable Protein Utilization in Human Foods and Animal Feedstuffs*. Chapter 31. American Oil Chemists' Society, Champaign, Illinois, pp. 175-203.

Marshall, J.S., Stubbs, J.D., Chitty, J.A., Surin, B. and Taylor, W.C. (1997) Expression of the C_4 Me1 gene from Flaveria bidentis requires an interaction between 5' and 3' sequences. *The Plant Cell* 9:1515-1525.

Mendelsohn, M., Kough, J., Vaituzis, Z. and Matthews, K. (2003) Are *Bt* crops safe? *Nature Biotechnology* 21:1003-1009.

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.

Michiels, F., Morioka, S., Scheirlinck, T. and Komari, T. (1992) Stamen-specific promoters from rice. In: Plant Genetic Systems N.V.(Brussels, B. eds. PCT/EP1992/000274(WO/1992/013956): http://www.wipo.int/pctdb/en/wo.jsp?wo=1992013956.

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

Molony, L.-A. and Hassall, G.H.D. (2008) *Signposts for Australian Agriculture - The Australian Cotton Industry*. National Land and Water Resources Audit, Australian Government. http://nlwra.gov.au/files/products/national-land-and-water-resources-audit/pn21908/pn21908.pdf.

Murakami, T., Anzai, H., Imai, S., Satoh, A., Nagaoka, K. and hompson, C.J. (1986) The bialaphos biosynthetic genes of *Streptomyces hygroscopicus*: Molecular cloning and characterization of the gene cluster. *Molecular and General Genetics* 205:42-50.

Murray, F., Llewellyn, D., McFadden, H., Last, D., Dennis, E.S. and Peacock, W.J. (1999) Expression of the *Talaromyces flavus* glucose oxidase gene in cotton and tobacco reduces fungal infection, but is also phytotoxic. *Molecular Breeding* 5:219-232.

NPTN. (2000) *Bacillus thuringiensis* (Technical Fact Sheet). National Pesticide Telecommunications Network, Oregon State University & U.S. Environmental Protection Agency, http://npic.orst.edu/factsheets/BTtech.pdf.

O'Brien, R.D. (2008) *Fats and Oils: Formulating and Processing for Applications*. 3rd ed, CRC Press.

Odell, J.T., Nagy, F. and Chua, N.H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* 313(6005):810-812.

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

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

OECD (2004) *Consensus Document on Compositional Considerations for New Varieties of Cotton (Gossypium hirsutum and Gossypium barbadense: Key Food and Feed Nutrients and Anti-Nutrients*. Report No. ENV/JM/MONO(2004)16, Organisation for Economic Co-operation and Development. http://www.olis.oecd.org/olis/2004doc.nsf/LinkTo/NT0000480E/$FILE/JT00168142.PDF.

OECD (2007) *Consensus Document on Safety Information on Transgenic Plants Expressing Bacillus thuringiensis - Derived Insect Control Proteins*. Report No. ENV/JM/MONO(2007)14, Organisation for Economic Co-operation and

Development.
http://www.olis.oecd.org/olis/2007doc.nsf/LinkTo/NT00002DF6/$FILE/JT03230592.PDF.

OGTR (2008a) *Risk Assessment and Risk Management Plan for DIR 087: Limited and Controlled Release of Cotton Genetically Modified for Insect Resistance and Herbicide Tolerance*. Document prepared by the Office of the Gene Technology Regulator, Australia. http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/dir087-2008.

OGTR (2008b) *The Biology of Gossypium hirsutum L. and Gossypium barbadense L. (cotton)*. Document prepared by the Office of the Gene Technology Regulator, Australia. http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/cotton-3/$FILE/biologycotton08.pdf

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.

Parker, M.W. and Feil, S.C. (2005) Pore-forming protein toxins: from structure to function. *Progress in Biophysics and Molecular Biology* 88:91-142.

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.

Samples, J.R. and Buettner, H. (1983) Ocular infection caused by a biological insecticide. *The Journal of Infectious Diseases* 148:614.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D.R. and Dean, D.H. (1998) Bacillus thuringiensis and its pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews* 62(3):775-806.

Scrimshaw, N.S. (1980) The background and history of Incaparina. *The United Nations University Press Food and Nutrition Bulletin* 2(2): http://www.unu.edu/Unupress/food/8F022e/8F022E01.htm.

Shelton, A.M., Zhao, J.-Z. and Roush, R.T. (2009) Economic, ecological, food safety and social consequences of the deployment of Bt transgenic plants. *Annual Review of Entomology* 47:845-881.

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

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

Thomas, K., MacIntosh, S., Bannon, G., Herouet-Guicheney, C., Holsapple, M., Ladics, G., McClain, S., Vieths, S., Woolhiser, M. and Privalle, L. (2009) Scientific advancement of novel

protein allergenicity evaluation: An overview of work from the HESI Protein Allergenicity Technical Committee (2000 - 2008). *Food and Chemical Toxicology* 47:1041-1050.

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

Trolinder, L., Jefferson, G. and Debeuckeleer, M. (2008) Herbicide tolerant cotton plants and methods for producing and identifying same. (US7442504), PatentStorm http://www.patents.com/Herbicide-tolerant-cotton-plants-methods-producing-identifying-same/US7442504/en-US/.

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

Vazquez-Padron, R.I., de la Riva, G., Agüero, G., Silva, Y., Pham, Si.M., Soberón, M., Bravo, A. and Aitouche, A. (2004) Cryptic endotoxic nature of *Bacillus thuringiensis* Cry1Ab insecticidal crystal protein. *FEBS Letters* 570:30-36.

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

Warren, R.E., Rubenstein, D., Ellar, D.J., Kramer, J.M. and Gilbert, R.J. (1984) *Bacillus thuringiensis* var *israelensis*: protoxin activation and safety. *The Lancet* 323:678-679.

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

WHO. (1999) Microbial pest control agent *Bacillus thuringiensis*. Environmental Health Criteria. Vol 217, World Health Organization, Geneva.

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