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

# Supporting Document 1

# SAFETY ASSESSMENT REPORT (APPROVAL)

# APPLICATION A1040 – FOOD DERIVED FROM INSECT-PROTECTED AND HERBICIDE-TOLERANT COTTON LINE GHB119

## SUMMARY AND CONCLUSIONS

### Background

A new genetically modified (GM) cotton line, GHB119, 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 Cry2Ae protein from *Bacillus thuringiensis* and herbicide tolerance is conferred by expression of phosphinothricin acetyltransferase (PAT) from *Streptomyces hygroscopicus*.

The Applicant anticipates that cotton lines containing event GHB119 will be commercially cultivated in major cotton-producing countries, including Australia. Food products containing event GHB119 would therefore be expected to enter the Australian and New Zealand food supply via local production and imports.

### 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 GHB119 would, like cotton products currently available on the market, consist of oil and linters.

### Molecular Characterisation

Cotton line GHB119 contains two novel gene cassettes. One contains a modified *cry2Ae* 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 marker genes present in line GHB119. Comprehensive molecular analyses of cotton line GHB119 indicate there is a single insertion site containing one complete copy of the two gene cassettes. 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. Three unexpected ORFs are present at the 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 GHB119 expresses two novel proteins, Cry2Ae and PAT. Expression analyses of the two proteins showed that the Cry2Ae protein is detectable in all parts of the plant but is not present in nectar; it is lowest in pollen and highest in leaves during the early stages of growth (average of 9.33 µg/g fresh weight). PAT is probably expressed in all plant parts tested but is often at levels below the limit of detection. It is likely to be highest in young leaves (average of 27.4 µg/g fresh weight). Both Cry2Ae and PAT are detectable in fuzzy cottonseed and a range of processed products derived from fuzzy cottonseed but not in the oil.

A number of studies were done to confirm the identity and physicochemical and functional properties of the expressed plant-derived Cry2Ae and PAT proteins, as well as to determine their potential toxicity and allergenicity. These studies have demonstrated that the Cry2Ae and PAT proteins conform in size and amino acid sequence to that expected, do not exhibit any post-translational modification including glycosylation and exhibit the expected activity. In relation to potential toxicity and allergenicity, it is worth noting that Cry2Ae 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 Cry2Ae and PAT are unlikely to be toxic or allergenic to humans.

## Compositional Analyses

Detailed compositional analyses were done of fuzzy seed derived from GHB119 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, data were obtained from two further studies using the GM cultivar 'TwinLink'™, which is the product of a conventional cross between line GHB119 and another GM cotton line. These studies measured various constituents in fuzzy seed as well as in processed products.

For fuzzy cottonseed, across most of the categories but most notably in the amino acids, some significant differences were found in individual analytes between seeds from GHB119 and those of the non-GM control. 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. The mean analyte levels found in seeds from GHB119 fell within the range of natural variation in commercial cotton cultivars. The compositional analysis of seeds from 'TwinLink'™ indicated few significant differences from seeds of the non-GM control.
For processed products derived from 'TwinLink'™ 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 containing event GHB119 when compared with conventional cotton cultivars currently on the market.

**Nutritional Impact**

Based on the results from a broiler feeding study using meal from TwinLink™ cotton, it was concluded that cottonseed meal containing event GHB119 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 GHB119. On the basis of the data provided in the present Application, and other available information, food derived from cotton line GHB119 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 |
| CCI | Confidential Commercial Information |
| Cry | Crystal protein from *Bacillus thuringiensis* |
| Cyt | Cytolytic protein from *Bacillus thuringiensis* |
| DNA | deoxyribonucleic acid |
| T-DNA | transferred DNA |
| dw | dry weight |
| EFSA | European Food Safety Authority |
| ELISA | enzyme linked immunosorbent assay |
| FAO | Food and Agriculture Organization of the United Nations |
| FSANZ | Food Standards Australia New Zealand |
| GEM2 | Same as Cry2Ae protein |
| GM | genetically modified |
| HPLC | high performance liquid chromatography |
| 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 |
| SD | standard deviation |
| 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 Organisation |

# 1. Introduction

A genetically modified (GM) cotton line, GHB119, has been developed which is protected against feeding damage by Lepidopteran insects, particularly cotton bollworm (*Helicoverpa* spp.), a major cotton pest in Australia, and tolerant to herbicides containing glufosinate ammonium.

Protection against feeding damage by Lepidopteran insects is achieved through expression in the plant of an insecticidal crystal protein, Cry2Ae, encoded by a modified *cry2Ae* gene derived from the soil bacterium *Bacillus thuringiensis.* 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 GHB119 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 (Figure 1), 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 20$^{th}$ 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 6$^{th}$ 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, 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. Subsequent alkali treatment and deodorisation steps of the oil refining process are likely to remove any last detectable traces of protein in the oil.

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.

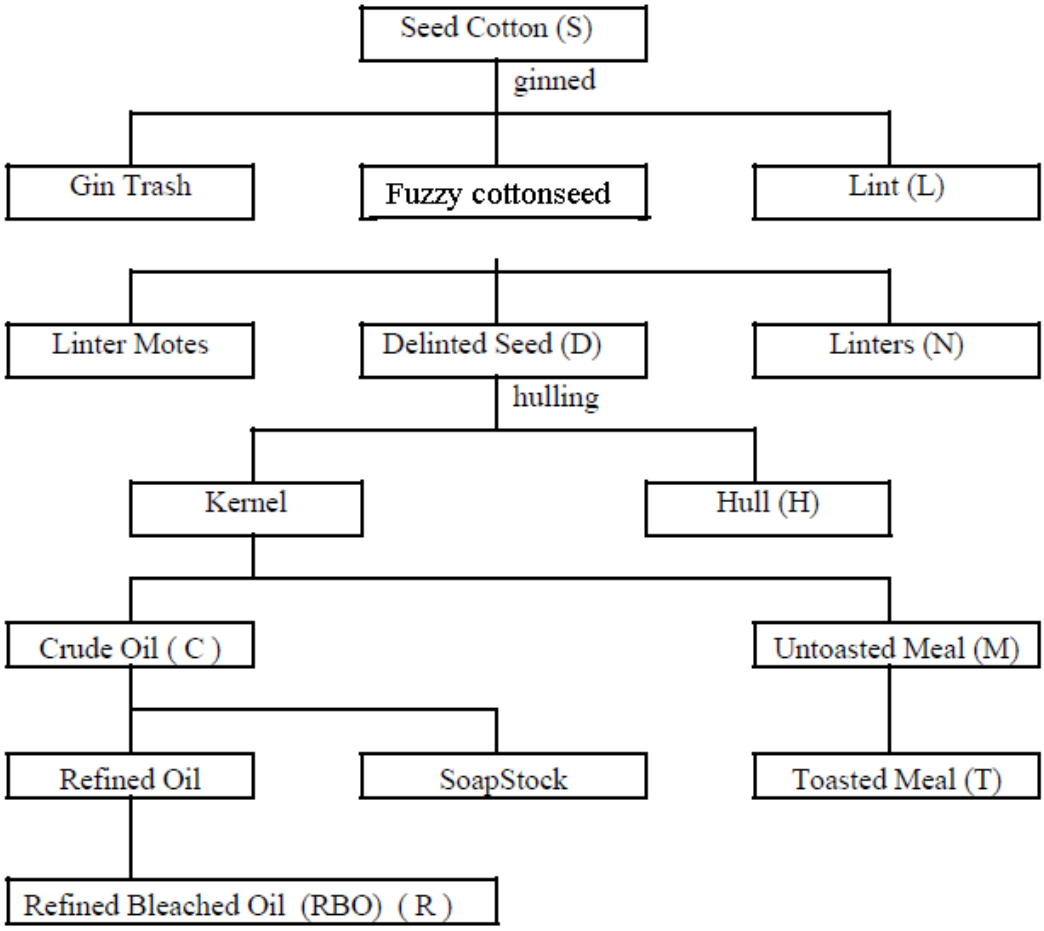


*Figure 1: The major processed fractions obtained from cottonseed*

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 – 2007 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[1] 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).

[1] Fuzzy (or whole) cottonseed is the linted cottonseed remaining after the ginning process which removes fibres for textile production (refer to Figure 1).

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 deep fryers (Molony and Hassall, 2008).

Another possible food product that can be derived from the cotton plant is bee pollen. This substance is produced by bees during foraging and is taken back to the hive to be fed to larvae and young adult bees (Krell, 1996). It comprises pollen grains that are pelleted by the bee in the corbiculae ('pollen baskets') located on the posterior pair of legs. Beekeepers can collect the pellets by placing a screen at the entrance to a hive; as the bees pass through the screen, the pellets are dislodged and fall into a collection tray. The pellets are frozen or dried for storage and are then packaged for sale as bee pollen, which is generally consumed as the raw product without any further processing. The size of this market in Australia is so small that it has not been included in the most recent survey of the Australian Honeybee industry (Crooks, 2008).

The cotton cultivar 'Coker 312' was used as the parental variety for the genetic modification described in this application, and thus is regarded as the near-isogenic line for the purposes of comparative assessment with GHB119. 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 (Krell, 1996; OGTR, 2008a). Traits introduced into 'Coker' cultivars are transferred to commercial cultivars by backcrossing.

Additional comparators have also been used for various studies. These include the conventional cotton cultivars 'FM 966' and 'FM 958', that are FiberMax® cultivars developed in Australia by the Commonwealth Scientific and Industrial Research Organisation and selected by a joint venture breeding programme of Aventis and Cotton Seed International for production conditions in the U.S.

An additional GM cultivar, known as 'TwinLink'™ (see eg OGTR, 2008a), has also been used in various studies evaluated as part of this assessment. This cultivar[2] was a source of Cry2Ae and PAT proteins for the protein equivalence studies described in Section 4, a source of analytes for the compositional analyses of fuzzy seed and processed products described in Sections 5.3.6 and 5.3.7, and as the source of feed for a broiler study (Section 6.1). 'TwinLink'™ was produced (refer to Figure 3) by the conventional crossing of line GHB119 with another GM line (T304-40[3]) containing the *cry1Ab* gene. In studies involving 'TwinLink'™, the Applicant used the cultivar 'Coker 315' as the non-GM comparator since this was the near-isogenic line from which T304-40 was derived.

## 2.2 Donor organisms

### 2.2.1 Bacillus thuringiensis (Bt)

The Cry protein expressed in cotton line GHB119 is derived from *B. thuringiensis* subspecies *dakota*. *B. thuringiensis* is a facultative anaerobic, gram-positive spore-forming bacterium that, while typically referred to as a soil bacterium, probably has its main ecological niche in insects (Federici, 1999).

[2] Currently 'TwinLink'™ is undergoing regulatory approval processes in a number of countries. As a product of conventional breeding, food derived from 'TwinLink'™ does not require approval by FSANZ providing both GM parents are approved separately for food use.

[3] Food derived from cotton line T304-40 was added to the *Australia New Zealand Food Standards Code* in May 2010. (http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm)

The species (often referred to just as '*Bt*') is more appropriately regarded as a complex of over sixty subspecies that are characterised by the production of a parasporal body during the sporulation phase (Federici, 1999). The subspecies can be distinguished from one another on the basis of immunological differences in flagellar (H antigen) serotype or molecular techniques (see eg Yu *et al*., 2002). The parasporal body contains one or more crystalline protein inclusions that are toxic to insects and are categorised as either Crystal (Cry) or Cytolytic (Cyt) toxins, also called δ-endotoxins. The Cry toxins 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. Both types of toxins are innocuous to humans, vertebrates and plants, and are completely biodegradable (Bravo *et al*., 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). Infection in humans is unusual although there have been at least two clinical reports, one in the wounds of a soldier (Hernandez *et al*., 1998) and one in burn wounds (Damgaard *et al*., 1997), and in both cases impaired immunosuppression was implicated in the cause of the infection. *B. thuringiensis* has also been rarely associated with gastroenteritis (see eg Jackson *et al*., 1995) but generally, *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).

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).

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. The use of *B. thuringiensis* products in the field can result in considerable aerosol and dermal exposure in humans. Studies of human populations exposed to *Bacillus thuringiensis* subsp. *kurstaki*, aerial spraying in Oregon (USA) and British Columbia (Canada) did not indicate any association with short-term clinical illness, except possibly in people who were immunocompromised (Green *et al*., 1990; Valadares de Amorim *et al*., 2001).

### 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 GHB119 transformation methodology. Report ID: BIO2-005_TransfMethod_151. Bayer CropScience (unpublished).

Habex, V.; Verhaege, S. (2008). Detailed insert characterization of *Gossypium hirsutum* transformation event GHB119 by Southern blot analysis. Study No. BBS07-009, Bayer CropScience (unpublished).

Verhaege, S.; Habex, V. (2008). Full DNA sequence of event insert and integration site of *Gossypium hirsutum* transformation event GHB119. Report No. BBS08-001. Bayer CropScience (unpublished study).

Moens, S. (2008). Confirmation of the absence of vector backbone sequences in *Gossypium hirsutum* transformation event GHB119. Report No. BBS06-007. Bayer CropScience (unpublished).

Verhaege, S.; De Pestel, K. (2008). Bioinformatics analysis of newly created ORFs from *Gossypium hirsutum* event GHB119. Report No. BIO2-005_Bioinfo_130. Bayer CropScience (unpublished).

Verhaege, S.; De Pestel, K. (2008).. Bioinformatics analysis of the pre-insertion locus of *Gossypium hirsutum* transformation event GHB119. Report No. BIO2-005_FullSeq_131. Bayer CropScience (unpublished).

Verhaege, S.; Criel, I. (2008). Structural stability analysis of *Gossypium hirsutum* transformation event GHB119 in different generations, in different backgrounds and when grown in different environments. Report BIO2-005_StructStab_141, Bayer CropScience (unpublished).

## 3.1 Method used in the genetic modification

Cotton cultivar 'Coker 312' was transformed via agrobacterium-mediated transformation (Deblaere *et al*., 1987) basically following the method of Reynaerts & De Sonville ( 2002). 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, embryogenic callus derived from hypocotyl explants of *in vitro* germinated seedlings was co-cultivated with the disarmed *Agrobacterium tumefaciens* strain C58C1$^{R}$ containing the helper Ti plasmid pEHA101 and T-DNA vector pTEM12 (refer to Figure 2). Following plant regeneration, transformed plantlets were selected on a medium containing glufosinate ammonium as the selection agent. These plantlets (T0) were then transferred to a glasshouse for further characterisation and selection.

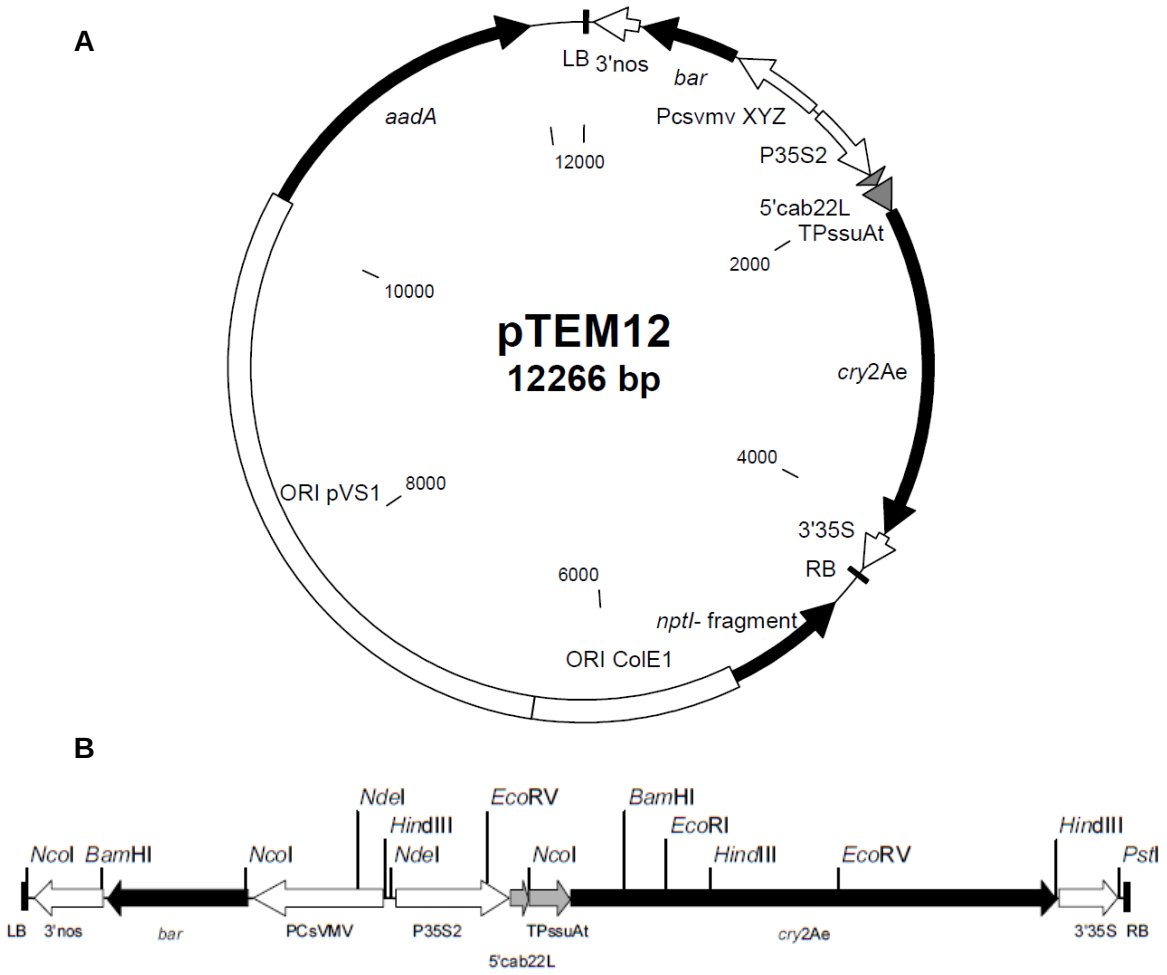


*Figure 2: (A) Vector map of plasmid pTEM12 and (B) the T-DNA region of the plasmid*

## 3.2 Function and regulation of introduced genes

Information on the genetic elements in the pTEM12 plasmid is summarised in Table 1.

**Table 1: Description of the genetic elements contained in plasmid pTEM12**

| Genetic element | bp location on plasmid pTEM12 | Size (kb) | Source | Orientation | Description & Function | References |
|---|---|---|---|---|---|---|
| ***T-DNA*** | | | | | | |
| *3'nos* | 26 - 335 | 0.31 | *Agrobacterium tumefaciens* | Anti-clockwise | • Sequence including the 3'UTR of the nopaline synthase gene<br>• Terminates bar gene expression and directs polyadenylation | *Depicker et al. (1982)* |
| *Bar* gene | 336 - 887 | 0.55 | *Streptomyces hygroscopicus* | Anti-Clockwise | • Coding sequence of the phosphinothricin acetyltransferase gene | Thompson et al. (1987) |
| *Pcsvmv XYZ* | 888 - 1423 | 0.53 | *Cassava vein mosaic virus* | Anti-clockwise | • Sequence including the promoter region<br>• Constitutive promoter<br>• Promotes bar expression | Verdaguer et al (1996) |
| *P35S2* | 1424 - 1920 | 0.49 | *Cauliflower mosaic virus* | Clockwise | • Sequence including the promoter region of the 35S transcript<br>• Constitutive promoter<br>• Promotes cry2Ae expression | Odell et al. (1985) |

| Genetic element | bp location on plasmid pTEM12 | Size (kb) | Source | Orientation | Description & Function | References |
|---|---|---|---|---|---|---|
| *5'cab22L* | 1921 - 1990 | 0.07 | *Petunia hybrida* | Clockwise | • Sequence including the untranslated leader sequence of the chlorophyll a/b binding protein gene<br>• Functions as a an efficient leader in 35S promoter constructions | Harpster et al (1988) |
| *TPssuAt* | 1991 - 2155 | 0.16 | *Arabidopsis thaliana* | Clockwise | • Coding sequence of the transit peptide of the ribulose-1,5-bisphosphate carboxylase small subunit gene ats1A<br>• Targets the Cry2Ae protein to the chloroplast to avoid degradation in the cytoplasm | De Almeida et al (1989) |
| *cry2Ae* gene | 2156 - 4051 | 1.89 | *Bacillus thuringiensis* | Clockwise | • Coding sequence of the cry2Ae gene (adapted to plant codon usage) | CCI |
| *3'35S* | 4052 - 4320 | 0.26 | *Cauliflower mosaic virus* | Clockwise | • Sequence including the 35S transcript<br>• Terminates cry2Ae gene expression and directs polyadenylation | Depicker et al. (1982) |
| ***Plasmid Backbone*** | | | | | | |
| RB | 4321 - 4345 | 0.19 | *Agrobacterium tumefaciens* | | • Right border repeat<br>• Required for the transfer of the T-DNA into the plant cell | Zambryski (1988) |
| | 4346 - 4537 | | | | • Residual plasmid sequences from pTiAch5 flanking the right border | Zhu et al (2000) |
| | 4538 - 5248 | | | | • Fragment of the neomycin phosphotransferase coding sequence of the nptI gene from transposon Tn903 | Oka et al (1981) |
| ORI ColE1 | 5249 - 6421 | 1.17 | *E. coli* | | • Fragment including the origin of replication from plasmid pBR322<br>• Permits replication of plasmid in E. coli. | Bolivar et al (1977) |
| ORI pVS1 | 6422 - 10192 | 3.77 | *Pseudomonas aeruginosa.* | | • Fragment including the origin of replication from plasmid pVS1<br>• Allows replication in A. tumefaciens | Hajdukiewicz et al (1994) |
| *aadA* | 10193 - 11961 | 1.76 | *E. coli* | Clockwise | • Fragment including the aminoglycoside adenyltransferase gene<br>• Confers resistance to erythromycin, streptomycin, and spectinomycin<br>• Allows selection of E.coli during vector construction. | Fling et al (1985) |
| | 11962 - 12266 | | | | • Residual plasmid sequences from pTiAch5 flanking the left border | |
| LB | 1-25 | | *Agrobacterium tumefaciens* | | • Left border repeat<br>• Required for the transfer of the T-DNA into the plant cell | Zambryski (1988) |

### *3.2.1* cry2Ae *gene*

The cry2Ae gene used in plasmid pTEM12 gives rise to a protein more correctly known as Cry2Ae1[4] (Bacillus thuringiensis toxin nomenclature database available online at http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/toxins2.html). The DNA sequence of the gene has been optimised for expression in plants (the Applicant has supplied the sequence but this is CCI). The structure of the Cry2Ae cassette is discussed in Arnaut et al (2007).

[4] Hereafter referred to just as Cry2Ae

The *cry2Ae* coding region in plasmid pTEM12 is 1895 bp in length and is driven by a chimeric 5'untranslated region comprising elements from the Cauliflower mosaic virus 35S promoter and an untranslated leader sequence from *Petunia hybrida* chlorophyll a/b binding protein which augments gene expression. A DNA sequence at the 5' end of the *cry2Ae* coding region encodes the *TPssuAt* transit peptide gene from *Arabidopsis thaliana* and allows chloroplast targeting of the encoded Cry2Ae protein. This in turn optimises stability of the Cry2Ae protein. At the 3' untranslated region of the coding region is a transcript termination and polyadenylation region from the 35S gene from Cauliflower mosaic virus.

### *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 4.2.2. 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 pTEM12 is 551 bp in length. It is driven by a promoter region of the Cassava vein mosaic virus and terminated by a sequence of the 3'untranslated region of the *nopaline-synthase (nos)* gene originating from the T-DNA of plasmid pTiT37 from *Agrobacterium tumefaciens*.

## 3.3 Breeding of cotton plants containing transformation event GHB119

A breeding programme was undertaken for the purposes of:

- obtaining generations suitable for analysing the molecular and genetic characteristics of line GHB119
- ensuring that the GHB119 event is incorporated into a conventional breeding line for commercialisation of insect-protected, herbicide-tolerant cotton.
- conventional crossing of plants containing event GHB119 with plants containing event T304-40 in order to obtain the GM cultivar'TwinLink'™ (see Section 2.1).

The breeding pedigree for the various generations is given in Figure 3.

For event GHB119, a number of lines (plants) with potential were chosen at the $T_0$ stage ($Cry2T_0$). 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 ($Cry2T_7$). The line containing event GHB119 was eventually selected for commercial development based on both its agronomic performance and expression of the two introduced genes and $T_0$ plants were back-crossed to a conventional breeding line. The same procedure was followed for event T304-40.

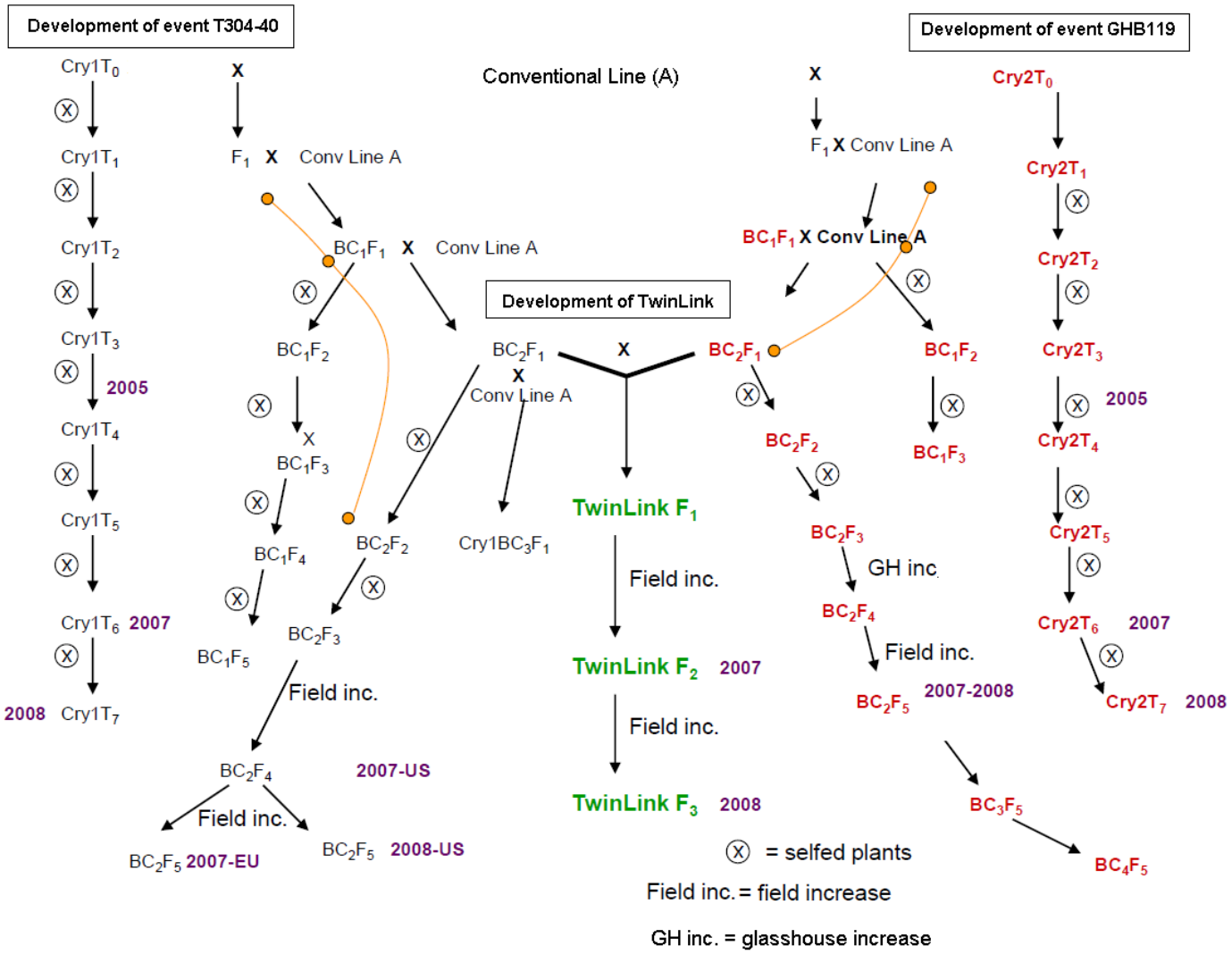


*Figure 3: Breeding strategy for plants containing event GHB119 and plants containing event T304-40*

## 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 GHB119. 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 $Cry2T_2$ generation (refer to Figure 3) and a negative control (non-GM cultivar 'FM966') were used for Southern blot analyses. A positive control (DNA from 'FM966' spiked with T-DNA from plasmid pTEM12) was also included in the Southern blot analyses. The DNA from $Cry2T_2$ was digested with one of the following enzymes or enzyme combinations: *Dra*I, *Eco*RV, *Hind*III, *Nco*I, *Bam*HI, *Eco*RI, *Pst*I, *Ava*I, *Nde*I, *Xho*I, *Eco*RI/*Pst*I and *Eco*RI/*Nde*I (see location of seven of these enzymes in Figure 2B). DNA from 'FM966' was digested with one restriction enzyme (*EcoRV*). 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 lengths of all hybridisation fragments were determined using a commercially available software package (Genetools). Based on the Southern blot analysis, it was determined that there is a single insert in event GHB119 and that the arrangement of the T-DNA genetic material is the same as that in the pTEM12 plasmid (refer to Figure 2).

### *3.4.2 Full DNA sequence of insert*

Genomic DNA was obtained from leaf tissue of $Cry2T_2$ plants (refer to Figure 3) and a negative control (non-GM wild type cultivar 'Coker 312'). 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 three overlapping fragments and determination of the consensus sequence. This was compared with that of the T-DNA in plasmid pTEM12 using commercially available software (Clone Manager). Four-fold sequence cover was obtained for each base.

The DNA sequence of the transgenic locus was found to consist of 358 bp of 5' flanking sequence, 4302 bp of inserted transgenic sequence identical to that of the T-DNA of the pTEM12 plasmid, and 320 bp of 3' flanking sequence (refer to Figure 4). The 5' and 3' flanking sequences were verified as identical to sequences in the integration site except that there were 9 less and 11 more base pairs in the 5' flanking region and 3' flanking region, respectively, of the transgenic locus. An eight 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.

The sequence analysis confirms that the flanking sequences in event GHB119 are of *Gossypium hirsutum* origin and that all inserted sequences originate from the T-DNA of the transforming pTEM12 plasmid.

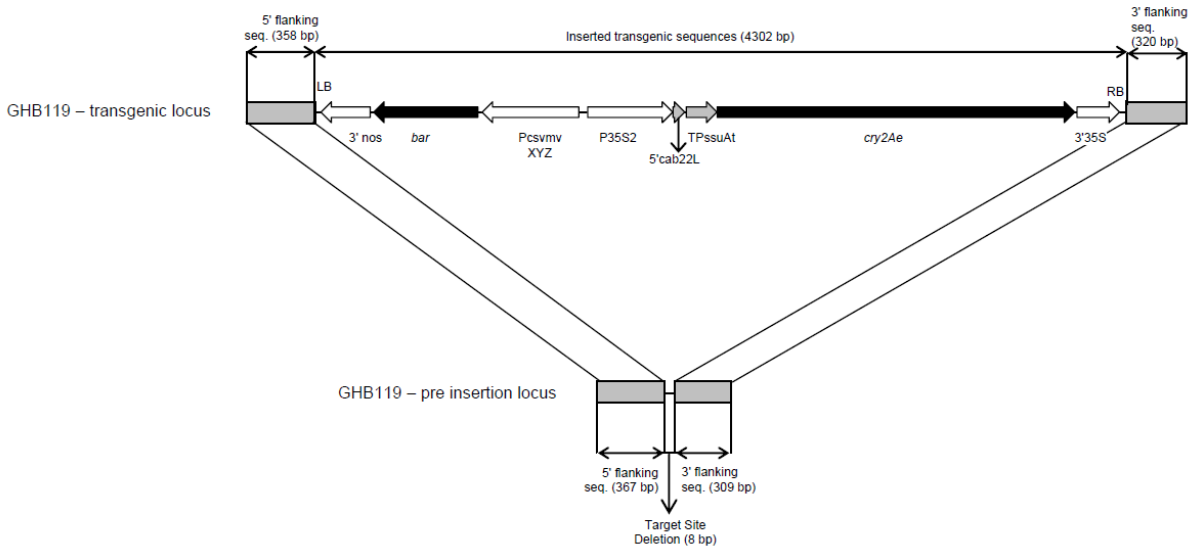


*Figure 4: Schematic drawing of the transgenic locus of event GHB119 compared with the T-DNA of plasmid pTEM12 and the pre-insertion site of 'Coker 312'*

### *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 GHB119. Leaf tissue from seedlings of generation $Cry2T_2$ (refer to Figure 3) 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 $Cry2T_2$ was digested with *EcoRV* and *Dra1* restriction enzymes while control DNA was digested with *EcoRV* enzyme. The resulting DNA fragments were separated and transferred to a membrane for sequential hybridisation with five overlapping radiolabelled probes (covering the complete vector backbone sequence of pTEM12) and one T-DNA probe.

No hybridisation fragments were observed with either GHB119 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 GHB119.

### *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 in the junctions between the T-DNA insert and the host genomic DNA. 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 standard stop (TAA, TAG, TGA) codons with a minimum size coding for eight amino acids. Another programme, FGENESH (Softberry Inc.), which predicts introns and exons by statistical sequence analysis and polyA signals by homology search with known plant consensus sequences, was used to identify any 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 one ORF spanning the 5' junction and two ORFs spanning the 3' junction (refer to Figure 5). FGENESH did not identify any newly created putative genes overlapping either the 5' or 3' junctions.

TSSP identified five putative promoters (refer to Figure 5), which included one promoter region (Promoter region 5) already present in the pre-insertion site.

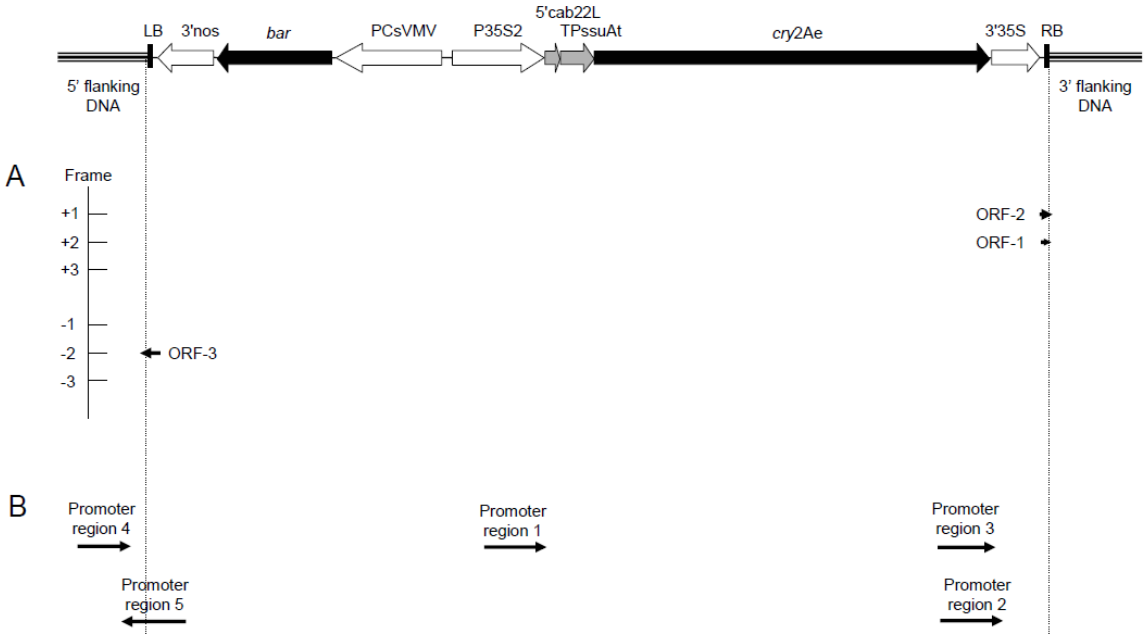


*Figure 5: Schematic overview of the newly created ORFs (A) and the prediction of regulatory elements (B) in the Gossypium hirsutum transformation event GHB119*

Since the 5' flanking regions around promoter region 5 do not show homology to known proteins, it is not likely that promoter region 5 would alter endogenous gene expression levels. The following points can be made about the remaining four promoter regions:

- All four of the newly-created promoter regions contain a transcription starting site and transcription binding sites, but only three of these (1, 4 and 2) also have a TATA box.
- Promoter region 4, while close to ORF-3 is not in the correct orientation to initiate transcription of ORF-3.
- Promoter regions 1 and 2 are in the correct orientation to initiate transcription of ORFS 1 and 2 but promoter region 1 is too far upstream.
- It may be possible for promoter region 2 to initiate transcription of ORFs 1 and 2 but the positioning of the TATA box some 200 nucleotides upstream of the start codons of ORFs 1 and 2 is not optimal.

The above evidence suggests that, of the five promoter regions, only promoter region 2 is a possible but unlikely regulatory element.

There is low overall similarity between the RBS sequence and the ATG region of the three predicted ORFs, with very few of the essential nucleotides of the RBS sequence being present. This indicates that, even if transcription occurred, translation of the putative ORFs would be unlikely.

Taken together, the analyses indicate that it is unlikely that any of the three predicted ORFs identified in cotton line GHB119 would 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.

The results of the BLAST search were consistent with the pre-insertion locus occurring in a region with repetitive elements containing no functional genes. No ORFs or putative genes were predicted to span the target site deletion region. The TSSP search identified one promoter region on the forward strand from bp 15 - 274 and one promoter on the reverse strand from bp 513 - 236 of the pre-insertion locus of the GHB119 event. However, since the homology search did not find any known endogenous genes in the pre-insertion locus and the FGENESH tool did not predict any genes, it is unlikely that the promoters predicted by TSSP are 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 GHB119.

## 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 event GHB119 was evaluated in individual plants of:

- three different generations: 4 – 18 plants from each of $F_1$, $BC_1$ $F_1$ and $BC_2$ $F_1$ (refer to Figure 3)
- two different genetic backgrounds: 22 hybrid plants obtained by crossing with each of ‘Coker 312 and ‘FM966’
- six different harvest locations in the U.S.: 10 - 14 plants from each of Chula (Georgia), Newport (Arkansas), Proctor (Arkansas), Senatobia (Mississippi), East Bernard (Texas) and Levelland (Texas).

The non-GM cultivar ‘FM966’, grown at all locations, was used as a negative control and cultivar ‘FM966’ spiked with DNA from plasmid pTEM12 was used as a positive control. Genomic DNA isolated from leaf tissue and the plasmid DNA was digested with *EcoRV* restriction enzyme (refer to Figure 2B for sites in the transgenic locus). The resulting DNA fragments were separated and transferred to a membrane for hybridisation with a radiolabelled probe to the entire T-DNA sequence in pTEM12 (4345 bp).

In one of the negative controls (*EcoRV*-digested DNA from 'FM966' grown in Levelland, Texas) there was an unexpected but weak hybridisation of the T-DNA probe to a 9 kb fragment. Further analysis revealed that this was due to aspecific hybridisation of the probe to an endogenous genomic fragment; this weak aspecific hybridisation fragment was also visible in all genomic DNA samples derived from the test substance.

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 GHB119 over different generations, in different genetic backgrounds and across different environments.

### *3.5.2 Phenotypic stability*

The Applicant did not undertake any inheritance studies on line GHB119 alone. However, the independent segregation of event GHB119 was evaluated in a conventional cross between lines GHB119 and T304-40 at the $BC_2F_1$ generation (refer to Figure 3) producing the cultivar 'TwinLink'™. Eighty progeny were analysed using a PCR-based method for determining zygosity, that utilises three oligonucleotide primers. Four genotypes were possible in the expected ratio 1:1:1:1 (refer to Table 2).

**Table 2: Segregation analysis of progeny derived from a cross between GHB119 x T304-40**

| Genotype | Expected ratio | Observed ratio |
|---|---|---|
| WT/WT* | 1 | 0.96 |
| GHB119/WT | 1 | 1.25 |
| T303-40/WT | 1 | 0.75 |
| GHB119/T303-40 | 1 | 1.05 |

* WT = wild type

Chi squared ($X^2$) analysis of the results showed that there was no significant difference ($p = 0.05$; $X^2$ less than 7.815 with 3 degrees of freedom) between the expected and observed ratios. This confirmed the independent inheritance of the GHB119 and T304-40 events.

While no formal inheritance studies of GHB119 were undertaken, confirmation of phenotype was implied by the selection, at each stage of the breeding programme, of only those plants expressing the PAT protein (i.e. tolerance to glufosinate ammonium).

## 3.6 Antibiotic resistance marker genes

No antibiotic marker genes are present in GHB119. Plasmid backbone analysis shows that no plasmid backbone has been integrated into the cotton genome during transformation i.e. the *aadA* gene, which was used as a bacterial selectable marker gene, is not present in GHB119.

## 3.7 Conclusion

Cotton line GHB119 contains two novel gene cassettes. One contains a cry2Ae 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 GHB119.

Comprehensive molecular analyses of cotton line GHB119 indicate that there is a single insertion site containing one complete copy of the two cassettes comprising the T-DNA from plasmid pTEM12.

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. Three unexpected ORFs are present at the 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.

Two types of novel proteins were considered:

- those that may be potentially generated as a result of the creation of ORFs during the introduction of the T-DNA of plasmid pTEM12 (see Section 3.4.4)
- those that were expected to be directly produced as a result of the translation of the introduced genes. Cotton line GHB119 expresses two novel proteins, a Cry2Ae 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. It should be noted that for the PAT protein, most of the relevant studies were submitted and considered in association with Application A1028 (FSANZ, 2010) and details have not been reiterated here.

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

**Study submitted:**

Capt, A. (2008).Cotton Transformation Event GHB119: *In Silico* Analysis of Putative Open Reading Frame (ORF) Sequences for Identifying Potential Homologies to Known Toxins and Allergens. Regulatory Toxicology Position Paper, Bayer CropScience (unpublished).

Open reading frame analysis identified the formation of three unexpected ORFs in event GHB119 (Section 3.4.4). The sequences corresponding to the three ORFs were compared with protein sequences present in a number of large public reference databases: Uniprot_Swissprot, Uniprot_TrEMBL, PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept. An Allergen database was also set up by compiling all amino acid sequences referenced with the keyword ‘allergen’ in similar public reference databases. The similarity searches used the BLASTP (Basic Local Alignment Search Tool Protein) (Altschul *et al*., 1997) and FindPattern (Genetic Computer Group - http://mikrobiologie.uni-graz.at/public/GCG/gcg_11/html/findpatterns.html) algorithms (for more detailed information of this type of analysis see Section 4.5.2). No significant similarities of the three ORFs to any allergens, toxins or anti-nutrient proteins were found.

It is concluded that there are very low allergen or toxin concerns relating to the three ORFs created by the transformation procedure used to generate cotton line GHB119. It is unlikely that any of the three ORFs identified in cotton line GHB119 are able to undergo transcription and express a protein (see discussion in Section 3.4.4). However, even if transcription could occur, the protein products are unlikely to be of concern.

## 4.2 Function and phenotypic effects of the Cry2Ae and PAT proteins

### *4.2.1 Cry2Ae 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. 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 *cry2* genes show a rather limited homology to other *cry* genes (Höfte and Whiteley, 1989). Cry2 protoxins (approximately 70 kDa) are smaller than the more widely studied Cry1 protoxins (approximately 140 kDa) but still require N-terminal processing (proteolytic removal of approximately 49 residues) to form an active toxin (Bravo *et al*., 2007). Cry2A proteins do, however, share a three-domain organisation with other Cry proteins therefore suggesting a common functional property (Hernández-Rodríguez *et al*., 2008). They are active predominantly against lepidopteran insects but, for example, the Cry2Aa protein may also be toxic to dipterans (Widner and Whiteley, 1990). Commercial microbial formulations, which include the Cry2A class of protein, have been used for control of lepidopteran pests for more than 40 years. In GM plants (often referred to as 'Bt-protected' plants) Cry2A proteins are used particularly to provide protection against *Helicoverpa* spp. (Hernández-Rodríguez *et al*., 2008). Cry2Ae has been incorporated into cotton plants specifically to provide protection against cotton bollworm (*Helicoverpa zea*) and tobacco budworm (*Heliothis virescens*).

### *4.2.2 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* i.e. 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 3.2.2) 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.

## 4.3 Protein expression analysis

**Studies submitted:**

Currier, T (2008). Protein Expression Analysis of Cotton Event GHB119, Expressing Cry2Ae and PAT/*bar* Proteins, USA, 2007. Report No. CY07B005. Bayer CropScience (unpublished).
Kowite, W.J. (2007). Production of RAC (Fuzzy Seed) Samples of GEM2 Cotton and the Non-Transgenic Counterpart, USA, 2006. Study No. CY06B001. Bayer CropScience (unpublished).
Martone, A. (2008). Analyses of Raw Agricultural Commodity (Fuzzy Seed) of Cry2Ae Cotton Event GHB119 for PAT/*bar* and Cry2Ae and its Non-Transgenic Counterpart for PAT/*bar* and Cry2Ae Proteins. Study No. Cy07B001. Bayer CropScience (unpublished).
Martone, A. (2008). Analyses of Processed Fractions from TwinLink™ Cotton Seed and Its Non-Transgenic Counterpart for PAT/bar, Cry1Ab and Cry2Ae Proteins, 2008, USA. Study No. CP08B010. Bayer CropScience (unpublished).

### *4.3.1 Novel protein expression in plant tissues*

The Cry2Ae and PAT proteins are expected to be expressed in all plant tissues since the *cry2Ae* and *bar* genes are driven by constitutive promoters (refer to Table 1). Plants were grown in a single greenhouse trial in the U.S. The Cry2Ae and PAT protein levels were determined by enzyme linked immunosorbent assay (ELISA) using commercial ELISA kits specific for each protein. The matrices sampled from 'Coker 312' control plants and glufosinate ammonium-sprayed $BC_2F_4$ plants (refer to Figure 3) are given in Table 3. The results for each matrix were an average from 5 replicates (plants). For roots, stems and leaves, samples were taken at several growth stages.

The purity of the seed sown for the study was confirmed by PCR or lateral flow strip detection. The seed containing event GHB119 had approximately 97% purity, while the seed of 'Coker 312' was shown to be free of the adventitious presence of a number of GM traits (Roundup Ready, Bollgard, Bollgard II, and Liberty). In addition, 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 3: Description of the plant matrices 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™) was used to calculate the concentrations of immunoreactive Cry2Ae and PAT proteins from optical density values. No Cry2Ae protein was detected in samples taken from 'Coker 312' plants. PAT protein was measured in some stem samples taken from 'Coker 312' but the levels were below the limit of quantitation. For GHB119 plants, Cry2Ae protein was detected in all parts except nectar (where there was insufficient material available to test). The protein was lowest in pollen and highest in leaves at the earlier growth stages (refer to Table 4). While detected in all GHB119 plant parts, PAT protein measurements were low and the average measurements given in Table 4 were below the limit of detection except in stems, young leaves and Stage 1 roots.

**Table 4: Average concentration of Cry2Ae and PAT proteins in various plant parts from GHB119**

| Sample | Growth stage | Average protein content in µg/g fresh weight ± SD | |
|---|---|---|---|
| | | Cry2Ae | PAT |
| Roots | 1 | 3.24 ± 0.55 | 19.20 ± 3.50 |
| | 2 | 2.15 ± 0.45 | 10.50 ± 3.40[A] |
| Stem | 1 | 2.45 ± 0.76 | 20.90 ± 4.30 |
| | 3 | 2.31 ± 0.38 | 18.00 ± 4.70 |
| Leaves | 1 | 9.33 ± 3.39 | 27.40 ± 6.90 |
| | 2 | 8.29 ± 1.64 | 28.40 ± 3.50[A] |
| | 3 | 4.44 ± 1.78 | 19.60 ± 3.70[A] |
| Squares | 3 | 1.76 ± 0.75 | 16.20 ± 3.40[A] |
| Apex | 3 | 2.38 ± 0.65 | 25.30 ± 9.20[A] |
| Bolls | 3 | 0.75 ± 0.15 | 6.57 ± 1.64[A] |
| Whole plant | 3 | 9.80 ± 4.25 | 19.00 ± 4.80[A] |
| Pollen | 3 | 0.33 ± 0.27 | 0.47 ± 0.33[A] |
| Nectar | 3 | ND[B] | 0.01 ± 0.01[A] |
| Flowers | 3 | 1.14 ± 0.15 | 10.20 ± 2.50[A] |
| Grain | 4 | 0.90 ± 0.56 | 2.37 ± 2.98[A] |

[A] Value below the limit of detection; [B] Insufficient material available for testing

### *4.3.2 Novel protein content in raw agricultural commodity (RAC)*

Fuzzy cottonseed is the RAC produced from ginned cottonseed (refer to Figure 1).

Fuzzy seed samples of sprayed and unsprayed GHB119 generation $BC_2F_5$ and the non-GM control 'Coker 312' were obtained from separate field trails in 2006 in the US States of Georgia, Arkansas, Mississippi and Texas. The identity and purity of the fuzzy seed was confirmed by PCR analysis of sub-samples. Ginning was done at each of the trial locations using 'research scale' gins, and kernel and lint coat fractions were obtained from the fuzzy seed. Cry2Ae and PAT proteins were detected in the fractions using ELISA, and protein values for the fuzzy seed were mathematically calculated based on the relative weight of these fractions.

Cry2Ae and PAT proteins were detected in all kernel and lint fractions obtained from GHB119 fuzzy seed, with the kernel fraction having the highest protein levels. The average quantities detected are given in Table 5. Statistical analysis indicated that site was the only factor that had a significant effect on the amount of Cry2Ae and PAT.

*Table 5: Average quantities of Cry2Ae and PAT proteins in fuzzy seed fractions derived from GHB119*

| Sample | Average Cry2Ae content | | Average PAT content | |
|---|---|---|---|---|
| | Unsprayed | Sprayed | Unsprayed | Sprayed |
| Kernel (µg/g ± SD) | 2.70 ± 0.85 | 2.60 ± 0.82 | 96.9 ± 6.50 | 92.80 ± 11.30 |
| Lint coat (µg/g ± SD) | 0.24 ± 0.34 | 0.19 ± 0.23 | 1.09 ± 1.92 | 0.88 ± 1.61 |
| Fuzzy seed (mean and range) * | 1.55 (1.30 – 1.81) | 1.47 (1.38 – 1.5) | 50.7 (44.5 – 54.9) | 49.9 (45.5 – 56.1) |
| Average percent of total crude protein | 0.00084 | 0.0001 | 0.027 | 0.028 |

* standard deviation was not calculated for fuzzy seed data because the value is the weighted numerical sum of the average kernel and lint coat measurements.

Traces of Cry2Ae (6.3 – 88.6 ng/g) and PAT (211 – 3330 ng/g) protein were also detected in all but 3 of the 'Coker 312' samples. The Applicant suggests that this is a consequence of inadvertent cross contamination during post-harvest handling.

### *4.3.3 Novel protein content in processed fractions*

The Applicant provided data for the levels of Cry2Ae and PAT in various processed fractions of the GM cultivar 'TwinLink'™ (refer to Section 2.1) which was grown in one field trial in Texas, U.S. under typical cotton production conditions. The seed was harvested and processed into food or feed commodities (refer to Figure 1) and the novel protein levels were quantified using ELISA (refer to Table 6).

**Table 6: Levels of Cry2Ae and PAT in processed fractions of seed derived from 'TwinLink'™**

| Processed Fraction | Cry2Ae (ng/g) mean +/-SD | PAT (ng/g) mean +/- SD |
|---|---|---|
| Seed cotton | 956 +/- 330 | 87300 +/- 31400 |
| Lint coat seed cotton | 113 +/- 16.3 | 9950 +/- 480 |
| Kernel seed cotton | 4060 +/- 326 | 368000 +/- 31400 |
| Lint seed cotton | Below limit of detection | 1310 +/- 954 |
| Lint | Below limit of detection | 1000 +/- 213 |
| Linters | 987 +/- 217 | 21000 +/- 1800 |
| Delinted seed | 2830 +/- 268 | 196000 +/- 24600 |
| Lint coat delinted seed | 271 +/- 35 | 5920 +/- 488 |
| Kernel Delinted seed | 4440 +/- 266 | 316000 +/- 24600 |
| Hulls | 528 +/- 41.2 | 17000 +/- 4050 |
| Meal; | 314 +/- 39.4 | 1550 +/- 230 |
| Toasted meal | 26 +/- 5.3 | 265 +/- 42 |

| Processed Fraction | Cry2Ae (ng/g) mean +/-SD | PAT (ng/g) mean +/- SD |
|---|---|---|
| Crude oil | Below limit of detection | Below limit of detection |
| Refined bleached oil | Below limit of detection | 38.8 +/- 9.1 |

The analyses showed that the Cry2Ae protein was detected in all processed fractions (particularly the kernel fractions) except lint seed cotton, lint, crude oil and refined bleached oil. PAT protein was detected in all processed fractions except crude oil. It is somewhat anomalous that PAT was detected in refined oil, especially as it was not detected in the crude oil. Any oil extraction process involves high temperature and pressure that would be expected to denature proteins, and refined oil has undergone further alkali and deodorization which should remove any traces of remaining protein. It is likely that the presence of PAT in refined oil was due to cross contamination during processing.

### 4.4 Protein characterisation and equivalence

The *cry2Ae* gene sequence used in plasmid pTEM12 would be expected to encode a protoxin with a molecular weight of 71 kDa and containing 631 amino acids. The deduced amino acid sequence of the Cry2Ae protein in GHB119 is identical to the native protein (Baum *et al.*, 2003) known as Cry2Ae1 (refer to Section 3.2.1), except that a methionine has been deleted from position 1. The PAT protein would be expected to comprise 183 amino acids and have a molecular weight of approximately 21 kDa.

The Cry2Ae and PAT proteins are not produced in sufficient quantity in cotton GHB119 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 Cry2Ae and PAT proteins were therefore expressed in recombinant *B. thuringiensis* and *E. coli* respectively and characterisation tests were done to confirm the identity and equivalency of these bacterially-produced proteins to those produced in cotton line GHB119.

**Studies submitted:**

Martone, A. (2008a). Structural and Functional Equivalence of Cry2Ae and PAT/*bar* Proteins Produced in *Bacillus thuringiensis* (*Bt*) and *Escherichia coli* to Cry2Ae and PAT/*bar* Proteins in GHB119 Cotton, *Gossypium hirsutum*, USA 2007. Study CY07B002 , Bayer CropScience (unpublished).

Martone, A. (2008b). 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).[5]

Martone, A. (2008c). 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).[5]

Rouquie, D. (2008a). Cry2Ae protein: epitope homology and N-glycosylation searches. Study SA 08207, Bayer CopScience (Unpublished).

Rouquie, D. (2008b). Cry2Ae protein with transit peptide: epitope homology and N-glycosylation searches. Study SA 08209, 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).[6]

Currier, T.C. (2005). Analysis to determine if the GEM2 Protein from Cotton Leaves is Glycosylated. Study No. BK05B003, Bayer CropScience (unpublished).[7]

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).[5]

[5] Study was submitted with Application A1028 but the Cry2Ae protein data are relevant to Application A1040

[6] This study was submitted and considered in Application A1028 (http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm)

[7] GEM2 is an alternative designation for Cry2Ae (Kepiro, 2008)

In three separate studies (see Martone 2008a,b,c above), the bacterially-derived proteins were compared to Cry2Ae and PAT proteins isolated from seed and leaf tissue of both line GHB119 and 'TwinLink'™.

### *4.4.1 Microbially expressed proteins*

Microbial Cry2Ae protein and PAT protein were obtained from bacterial expression systems (*Bacillus thuringiensis* and *Escherichia coli*, respectively) using standard methods. 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 Cry2Ae 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. A molecular weight of between 66 – 70 kDa was obtained for the Cry2Ae protein over a number of separate analyses 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 71 kDa for Cry2Ae 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 (i.e. Cry2Ae 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 the relevant bacterial expression system with the equivalent native proteins.

The bioactivity of Cry2Ae protein derived from *B. thuringiensis* 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*., 1992). 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 Cry2Ae 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 GHB119 could be done.

#### *4.4.1.1 HPLC/Electrospray mass spectrometry (LC/MS)*

Peptide maps were produced from the selected ion chromatograms obtained for tryptic digests of the Cry2Ae and PAT proteins isolated from bacterial expression systems. Based on the theoretical peptide sequences of the Cry2Ae and PAT proteins, it was estimated that the Cry2Ae protein from *B. thuringiensis* provided 68% coverage and the PAT protein from *E. coli* covered 96%.

#### *4.4.1.2 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* and *B. thuringiensis*, like most prokaryotes, lack the capacity for protein glycosylation (Abu-Qarn *et al*., 2008; Wacker *et al*., 2002) it would be unusual to find that N-glycosylation had occurred in the bacterially-derived proteins.

An *in silico* approach was used to search the Cry2Ae (both with and without the transit peptide sequence) and PAT proteins for the occurrence of the two motifs given above. The search of the Cry2Ae protein found ten potential sites on the basis of the N-X-S/T consensus sequence and one site based on the N-X-C consensus sequence. The presence of the transit peptide sequence did not alter this outcome. 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 Cry2Ae protein were actual N-glycosylation sites. To assess whether post-translational glycosylation had occurred, glycosylation analysis of a purified Cry2Ae protein sample from *B. thuringiensis* 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.

### *4.4.2 Protein equivalence*

Having established the authenticity and characteristics of the bacterially-derived Cry2Ae 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 Cry2Ae 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. A *Helicoverpa zea* feeding study indicated that the Cry2Ae protein isolated from 'TwinLink'™ leaves had the same biological activity as the Cry2Ae derived from *B. thuringiensis*.

LC/MS analysis of 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. The Cry2Ae protein from 'TwinLink'™ leaves covered 46% of the theoretical peptide sequences, and of those peptides analysed by selective ion monitoring, 57% were identical to those in the *B. thuringiensis*-derived protein. This low coverage of the theoretical sequence of the Cry2Ae protein is a reflection of the folding of the protein which, in turn, determines the access of the trypsin used for digestion. It indicates that, for the Cry2Ae protein, LC/MS alone is not a sufficient methodology for determining protein equivalence.

As with the *B. thuringiensis*-derived Cry2Ae protein, glycoprotein staining of plant-derived Cry2Ae indicated that none of the potential glycosylation sites were 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 Cry2Ae and PAT proteins produced in the bacterial expression systems and 'TwinLink'™ leaves.

The expected sequence would be asparagine, asparagine, valine, leucine, asparagine. The analysis did not prove useful for the PAT protein because of blocking of the N-terminal sequence (a common and often unavoidable problem encountered in protein sequencing). Partial blocking was also a problem for the Cry2Ae protein produced in 'TwinLink'™ leaves; the N-terminal sequence was, however, detected by LCMS. For the Cry2Ae protein obtained from the bacterial system, results suggested that the first five amino acids of the N-terminal sequence were methionine, asparagine, asparagine, valine, leucine. However, since the LCMS analysis of the bacterial Cry2Ae detected the N-terminal peptide, it is likely that the identification of a methionine on the N-terminal sequence by Edman degradation may have been caused by contamination during analysis.

To confirm whether the Cry2Ae and PAT proteins produced in 'TwinLink'™ leaves were equivalent to those produced in seeds and leaves of line GHB119 as well as to the proteins produced in the bacterial expression systems, two further minor studies were undertaken. Purified plant-derived Cry2Ae and PAT proteins were obtained from ground leaves and seeds from 'TwinLink'™ and line GHB119 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.

### *4.4.3 Conclusion*

The studies described above allowed the determination of the identity and physicochemical and functional properties of the Cry2Ae 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 Cry2Ae 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 GHB119, the identity and non-glycosylated status of the proteins produced in line GHB119 is demonstrated.

With regard to the equivalence of proteins produced in planta and in the bacterial expression systems:

- The Cry2Ae proteins isolated from *B. thuringiensis* and from plants were found to be equivalent in all parameters analysed.
- The PAT proteins isolated from *E. coli* and from plants were found to be equivalent in all parameters analysed.
- Based on weight of evidence, it is concluded that microbially-derived Cry2Ae and PAT proteins are suitable surrogates for use in safety assessment studies.

## 4.5 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.

### *4.5.1 History of human consumption*

#### *4.5.1.1 Cry2Ae*

The Cry2Ae protein expressed in line GHB119 is ubiquitous in the environment (OECD, 2007; Schnepf et al., 1998) and is therefore a natural contaminant of many human foods. Cry2A proteins have been used in microbial formulations sprayed on food crops for over 40 years. The HD1 isolate from *B. thuringiensis* subsp. *kurstaki* was the first Bt isolate to be developed commercially for the control of lepidopteran pests (Federici, 1999) and a Cry2A protein (formerly referred to as P2; Höfte and Whiteley, 1989) is one of its (as well as of other commercial isolates) components (Donovan *et al*., 1988; Moar *et al*., 1994). There is no evidence, from this long history of use as a plant pesticide, of any associated toxicity to humans (Kough, 2003). In addition, no safety issues have been raised with the use of food products derived from Bt crops (Delaney *et al*., 2008; Mendelsohn *et al*., 2003; OECD, 2007; Shelton *et al*., 2009), including those containing Cry2Ae (EPA, 2008). The Cry2Ae protein shows approximately 89% sequence homology with the Cry2Ab protein produced by GM Bollgard II® cotton (FSANZ, 2002) and GM YieldGuard VT® Pro corn (FSANZ, 2008) that are both grown commercially worldwide.

#### *4.5.1.2 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).

### *4.5.2 Similarities with known protein toxins*

**Studies submitted:**

Rouquie, D. (2008b). Cry2Ae protein with transit peptide: overall amino acid sequence homology search with known toxins and allergens, Study SA 08208, Bayer CropScience (unpublished).

Capt, A. (2009). Cry2Ae Protein: Amino Acid Sequence Homology Search with Known Toxins. PSI Number: TX99L096, 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).[8]

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

[8] This study was submitted and considered in Application A1028 (http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm)

The Cry2Ae (both with and without the transit peptide) and PAT sequences were compared with protein sequences present in a number of large public reference databases: eg Uniprot_Swissprot, Uniprot_TrEMBL, PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept.

The similarity searches used the BLASTP (Basic Local Alignment Search Tool Protein) algorithm (Altschul *et al*., 1997), Version 2.2.20 (February 2009) 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 Cry2Ae protein sequence (either with or without the transit peptide) 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 Cry12Ae 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).

### *4.5.3* In vitro *digestibility*

See Section 4.6.3.

### *4.5.4 Thermolability*

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

**Studies submitted:**

Rouquie, D. (2008). Cry2Ae Protein: Heat Stability Study. Study SA 08128, Bayer CropScience (unpublished)
Esdaile, D.J. (2002). Phosphinothricin Acetyltransferase (PAT) Bar Gene Product Heat Stability Study. Study SA 02175, Bayer CropScience (unpublished).[9]

#### *4.5.4.1 Cry2Ae*

Cry2Ae protein obtained from a bacterial expression system (refer to Section 4.4) was incubated for 10, 30 or 60 minutes at 60°, 75° or 90°C.

[9] This study was submitted and considered in Application A1028 (http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm)

The integrity of the protein was detected by gel electrophoresis (SDS-PAGE) coupled with a Coomassie blue staining, and 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° for 30 min.

The Western blot analysis using an anti-Cry2Ae protein polyclonal rabbit primary antibody and a peroxidase-coupled goat anti-rabbit secondary antibody indicated that immunoreactivity was detectable in all 60°C and 75°C treatments and in the 90°C treatment after 10 min. Immunoreactivity was not detected after 30 min at 90°C. These results indicate that the Cry2Ae protein is partially heat stable at 90°C for 30 min but is completely degraded at this temperature after 60 min.

*4.5.4.2 PAT*

PAT protein obtained from a bacterial expression system (refer to Section 4.4) 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.

### *4.5.5 Acute oral toxicity study*

An acute oral toxicity study using mice was conducted in order to examine the potential toxicity of the Cry2Ae protein obtained from a bacterial expression system (refer to Section 4.4).

**Study submitted:**

Rouquie, D. (2006). GEM2 Protein. Acute Oral Toxicity by Oral Gavage in Mice. Study SA 06235, Bayer CropScience (unpublished).

| | |
|---|---|
| Test material | *GEM2 (Cry2Ae) preparation from* B. thuringiensis |
| Vehicle | 50 mM $Na_2CO_3$ (Cry2Ae 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[10] body weight Cry2Ae test substance by oral gavage, within 4 h. Actual total dose was 1,187 mg/kg Cry2Ae since the purity of the protein was 93.85% in the test substance) |
| 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. Dilated uterine horns were noted in three of the mice but this was considered to reflect the stage of oestrus and not to be connected with administration of the test material. Under the conditions of this study, administration of Cry2Ae protein to female mice at a dose of 1,187 mg /kg bw produced no test substance-related clinical signs of toxicity, body weight losses, macroscopic abnormalities or mortality. These results support the conclusion that the Cry2Ae protein is not acutely toxic.

[10] 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).

The Applicant did not supply an acute oral toxicity study for the PAT protein. 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 Cry2Ae 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 Cry2Ae and PAT proteins with known protein allergen sequences
- evaluation of the lability of the microbially produced and purified Cry2Ae and PAT proteins using *in vitro* gastric and intestinal digestion models; and thermolability

### *4.6.1 Source of each protein*

As described in Section 4.2, both the Cry2Ae 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. (2008a). Cry2Ae protein: epitope homology and N-glycosylation searches. Study SA 08207, Bayer CopScience (Unpublished).

Rouquie, D. (2008b). Cry2Ae protein with transit peptide: overall amino acid sequence homology search with known toxins and allergens, Study SA 08208, Bayer CropScience (unpublished).

Rouquie, D. (2008c). Cry2Ae protein with transit peptide: epitope homology and N-glycosylation searches. Study SA 08209, 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).[11]

Herouet-Guicheney, C. (2006). Phosphinothricin acetyl transferase (PAT) BAR gene product: epitope homology and N-glycosylation searches. Study SA 06000, Bayer CropScience (unpublished).[11]

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).

[11] This study was submitted and considered in Application A1028 (http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm)

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 Cry2Ae and PAT with known protein toxins (refer to Section 4.5.2), the generation of an *E* value provides an important indicator of significance of matches (Baxevanis, 2005; Pearson, 2000).

The Cry2Ae (631 amino acids) sequence with and without its transit peptide was compared with all known allergen sequences contained in a reference allergen database (AllergenOnline, version 8.0) using the FASTA algorithm and BLOSUM62 scoring matrix. 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 Cry2Ae protein was subdivided into eight linearly contiguous amino acid blocks that were compared using the FindPatterns algorithm. Bioinformatic analysis of the PAT protein is described in Application A1028 (FSANZ, 2010).

For the Cry2Ae protein either with or without the transit peptide, 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 Cry2Ae 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 (Thomas *et al*., 2004) was conducted to determine the digestive stability of the Cry2Ae 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. (2008). Cry2Ae Protein *In Vitro* Digestibility Study in Human Simulated Gastric Fluid. Study SA 08126, Bayer CropScience (unpublished).
Rouquie, D. (2009). Cry2Ae Protein *In Vitro* Digestibility Study in Human Simulated Gastric Fluid – Complementary Study. Study SA 09100, 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).[12]

The *in vitro* digestibility of the *B. thuringiensis*-derived Cry2Ae protein in SGF (U.S.Pharmacopeia, 1990) containing pepsin was evaluated by incubating samples at 37º for selected times (0, 0.5, 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. Western blotting of the Cry2Ae SDS gels was also performed using an anti-Cry2Ae rabbit polyclonal antibody.

Both the SDS gels and Western blotting indicated that the Cry2Ae protein was rapidly hydrolysed in SGF, with 90% digestion in less than two minutes and complete digestion after five minutes exposure.

The procedure for testing the digestibility of the PAT protein in SGF was provided in Application A1028 (FSANZ, 2010). The PAT protein was rapidly hydrolysed in SGF, with complete digestion after 30 seconds exposure.

In the methods used for both the Cry2Ae and PAT analysis, the HRP positive control was rapidly hydrolysed (< 30 seconds) while the OVA negative control had not been completely digested after 60 minutes.

#### *4.6.3.2 Simulated intestinal fluid (SIF) studies*

**Studies submitted:**

Rouquie, D. (2008). Cry2Ae Protein *In Vitro* Digestibility Study in Simulated Intestinal Fluid. Study SA 08127, 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).[11]

The digestibility of *B. thuringiensis*-derived Cry2Ae protein in SIF containing pancreatin (U.S.Pharmacopeia, 1990) was assessed using SDS-PAGE. Digestibility of the protein in SIF was measured by incubating samples at 37$^{o}$ C 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 (anti-Cry2Ae rabbit polyclonal antibody).

A Cry2Ae band was observed following SDS-PAGE and Western blot analysis at all incubation times, but with diminishing intensity. It is therefore concluded that there is only a partial digestion of Cry2Ae by pancreatin after 60 min at 37$^{o}$ C.

[12] This study was submitted and considered in Application A1028 (http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm)

The procedure for testing the digestibility of the PAT protein in SIF was provided in Application A1028 (FSANZ, 2010). 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 4.5.4

### 4.7 Conclusion

Cotton line GHB119 expresses two novel proteins, Cry2Ae and PAT. Expression analyses of the two proteins showed that the Cry2Ae protein is detectable in all parts of the plant but is not present in nectar; it is lowest in pollen and highest in leaves during the early stages of growth (average of 9.33 µg/g fresh weight). PAT is probably expressed in all plant parts tested but is often at levels below the limit of detection. It is likely to be highest in young leaves (average of 27.4 µg/g fresh weight). Both Cry2Ae and PAT are detectable in fuzzy cottonseed and a range of processed products derived from fuzzy cottonseed but not in the oil.

A number of studies were done to confirm the identity and physicochemical and functional properties of the expressed plant-derived Cry2Ae and PAT proteins, as well as to determine their potential toxicity and allergenicity. These studies have demonstrated that the Cry2Ae and PAT proteins conform in size and amino acid sequence to that expected, do not exhibit any post-translational modification including glycosylation and exhibit the expected activity. In relation to potential toxicity and allergenicity, it is worth noting that Cry2Ae 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 Cry2Ae and PAT are unlikely to be toxic or allergenic to humans.

## 5. Compositional analysis

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

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

## 5.1 Key components

Fuzzy cottonseed 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 - malvalic, sterculic and dihydrosterculic 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). Composition of Raw Agricultural Commodity (Ginned Cottonseed) of the Insect-Tolerant Cotton (Event GHB119) and the Non-Transgenic Counterpart (Coker 312) Grown in Spain, Germany, 2008. . Report No. 08 B 003, Bayer CropScience (unpublished).

Oberdörfer, R. (2009). Composition of Raw Agricultural Commodity (Ginned Cottonseed) of Glyphosate-Tolerant Cotton (Event GHB614), the Non-Transgenic Counterpart (Coker 312) and Four Commercial Cotton Varieties Grown in Spain in 2007. . Report No. 08 B 002, Bayer CropScience (unpublished).

Oberdörfer, R. (2009). Composition of Raw Agricultural Commodity (Ginned Cottonseed) of the Insect-Tolerant Cotton (Event GHB119) and the Non-Transgenic Counterpart (Coker 312) Grown in Spain in 2008. . Report No. 09 B 003, Bayer CropScience (unpublished).

Oberdörfer, R. (2009). Composition of Raw Agricultural Commodity (Ginned Cottonseed) of Glyphosate-Tolerant Cotton (Event GHB614), the Non-Transgenic Counterpart (Coker 312) and Four Commercial Cotton Varieties Grown in Spain in 2008. . Report No. 09 B 002, Bayer CropScience (unpublished).

Rattemeyer-Matschurat, V. (2009). Analysis of Substantial Equivalence of Transgenic and Non-Transgenic Cotton By Means of t-test for Differences, Insect-Resistant Cotton (Event GHB119) vs Non-Transgenic Counterpart (Coker 312). Statistical Report for Studies 08 B 003 and 09 B 003, Bayer CropScience (unpublished).

Oberdörfer, R. (2009).Nutritional Impact Assessment Report on TwinLink™ Cotton. Report No. 08B012, 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 GHB119, this was the non-GM parental line 'Coker 312'. Both lines were grown in two separate studies in 2007 and 2008 (generation $BC_2F_4$ of GHB119) and sampled each season from eight sites in Spain[13] in one of the cotton growing regions in Europe. Only one site was common to both studies. Both the seed that was planted and the seed that was subsequently harvested, were analysed to verify seed identity and purity. In addition 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, there were three main treatments, namely: A) non-GM cotton ('Coker 312'); B) Line GHB119 not sprayed; and C) Line GHB119 sprayed with glufosinate ammonium at the 4-leaf and 12-leaf stages. The trial design was a randomised complete block. In the Andalusian sites, there were 5 plots (replicates) per treatment and in the Catalonia sites there were 3 plots (replicates) per treatment. A number of non-GM commercial cultivars[14] were also planted and sampled at the sites. Plots were harvested by hand from the interior rows of each plot and the cottonseed was ginned at a central site in Spain before transport to Germany for analysis.

[13] 2007: Alcalá de Guadaira, Dos Hermanas (2 sites), Coria del Rio, San José Del Valle (2 sites) – all in Andalucia; Vinyols I Els Arcs, Mont Roig del Camp – both in Catalonia
2008: Alcalá del Rio, Coria del Rio, Alcalá de Guadaira, Jerez de la Frontera (3 sites) – all in Andalucia
Mont Roig del Camp, Cambrils – both in Catalonia

[14] These cultivars were 'Alexandro', 'Crema III', 'Celia', and 'Flora'

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 144 fuzzy cottonseed samples provided results, with up to 57 analytes being obtained from each sample. For each analyte in each year at each site and each of the 3 treatments, 'descriptive statistics' were generated i.e. a mean, standard deviation, minimum value and maximum value were calculated. Figures for these same calculated values were also generated for each analyte by averaging the values over all years and all sites for each treatment. The means and standard deviations thus calculated are presented in Tables 7–11.

For statistical analysis, data were transformed into Statistical Analysis Software[15] (SAS) data sets and analysed using SAS version 8.2. In assessing the significance of any difference between means, a P-value of 0.05 was used. Analysis of Variance (ANOVA) was used for over-all analysis with the factors TREAT, SITE and the respective interaction. Based on the ANOVA model, treatment differences (A vs B and A vs C) were estimated and presented together with 95% confidence intervals (data not shown). In cases of significant TREAT x SITE interactions (where $p<0.05$), treatment comparisons are not valid and a by-site ANOVA was performed individually for each site with factor TREAT followed by t-tests to compare Treatment A vs Treatment B and Treatment A vs Treatment C.

The results for the three treatments were compared to a combined literature range for each analyte, compiled by the Applicant from published literature for commercially available cottonseed[16]. Any statistically significant differences between GHB119 and the 'Coker 312' control were also compared to the tolerance range compiled from the results of the non-GM commercial cultivars, to assess whether the differences were likely to be biologically meaningful.

## 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 total tocopherol) and anti-nutrients (gossypol, phytic acid and cyclopropenoid fatty acids). The results presented in Tables 7–11 show means and standard deviations (SD) averaged from all sites across both years.

### *5.3.1 Proximates and fibre*

Results for levels of proximate and fibre are shown in Table 7. Statistically significant differences between 'Coker 312' and GHB119 (both sprayed and unsprayed) were noted in the over-all analysis for most analytes but in all cases the means of GHB119 were within both the range reported in the literature and the tolerance range. There was a significant TREAT x SITE interaction for fat but the majority of the by-site analyses did not show significant differences between treatments (A vs B and A vs C).

[15] SAS website - http://www.sas.com/technologies/analytics/statistics/stat/index.html
[16] Published literature for cotton included OECD (2004), ILSI (2007) and Codex (2001)

**Table 7: Percentage dry weight (dw) of proximates and fibre in fuzzy cottonseed from 'Coker 312' and GHB119**

| Analyte | Calculated variable | A:'Coker 312' | B: GHB119 not sprayed | C: GHB119 - sprayed | Combined literature range | Tolerance range |
|---|---|---|---|---|---|---|
| Protein (%dw) | Mean | 28.5 | 23.9 | 23.9 | 11.7 – 34.2 | 18.2 – 27.7 |
| | SD | 2.4 | 1.7 | 1.8 | | |
| Fat (%dw) | Mean | 22.0 | 23.2 | 23.6 | 11.8 – 36.3 | 20.1 – 26.2 |
| | SD | 2.6 | 2.2 | 2.0 | | |
| Carbohydrate (%dw)[1] | Mean | 45.4 | 48.9 | 48.6 | 36.4 – 74.4 | 44.3 – 54.4 |
| | SD | 3.5 | 2.9 | 2.9 | | |
| ADF (%dw) | Mean | 38.4 | 39.3 | 39.5 | 29.0 – 66.9 | 34.2 – 42.8 |
| | SD | 2.1 | 1.9 | 1.9 | | |
| NDF (%dw) | Mean | 46.1 | 49.3 | 48.6 | 38.1 – 71.4 | 43.0 – 54.3 |
| | SD | 2.3 | 1.8 | 2.0 | | |
| Ash (%dw) | Mean | 4.1 | 3.9 | 3.9 | 3.2 – 5.0 | 3.3 – 4.4 |
| | SD | 0.3 | 0.2 | 0.2 | | |

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

### *5.3.2 Fatty Acids*

The levels of 17 fatty acids in fuzzy cottonseed from GHB119 and in 'Coker 312' control seed were measured and the means are given in Table 8. These showed the following:

- For lauric and eicosenoic acids, there was no significant difference between the means for GHB119 seeds and the means for 'Coker 312' seeds in the over-all analysis.
- For oleic, myristic, palmitoleic ω7, and arachidic acids the means for GHB119 seeds were significantly higher in the over-all analysis than means for 'Coker 312' seeds but fell within both the range found in the published literature and the tolerance range of the commercial cultivars.
- For linoleic and lignoceric acids the means for GHB119 seeds were significantly lower in the over-all analysis than the means for 'Coker 312' seeds but fell within both the range found in the published literature and the tolerance range of the commercial cultivars.
- There was a significant TREAT x SITE interaction for pentadecanoic, palmitic, palmitoleic ω9, heptadecanoic, stearic, oleic cis isomer, linoleic trans isomer, alpha linolenic and behenic acids. However, the majority of the by-site analyses for each of these did not show significant differences.

**Table 8: Percentage composition, relative to total fat, of major fatty acids in fuzzy cottonseed from 'Coker 312' and GHB119**

| | Calculated variable | A:'Coker 312' | B: GHB119 not sprayed | C: GHB119 - sprayed | Combined literature range | Tolerance range |
|---|---|---|---|---|---|---|
| Lauric acid (C12:0) | Mean | 0.02 | 0.02 | 0.02 | < 0.1 – 0.2 | 0.01 – 0.03 |
| | SD | 0.01 | 0.01 | 0.01 | | |
| Myristic acid (C14:0) | Mean | 0.81 | 0.84 | 0.85 | 0.53 – 1.17 | 0.56 – 1.01 |
| | SD | 0.15 | 0.16 | 0.16 | | |
| Pentadecanoic acid (C15:0) | Mean | 0.03 | 0.02 | 0.02 | Not available | 0.01 – 0.03 |
| | SD | 0.01 | 0.01 | 0.01 | | |
| Palmitic acid (C16:0) | Mean | 23.16 | 23.0 | 22.79 | 21.1 – 29.9 | 20.1 – 24.97 |
| | SD | 1.12 | 0.91 | 0.99 | | |

| | Calculated variable | A:'Coker 312' | B: GHB119 not sprayed | C: GHB119 - sprayed | Combined literature range | Tolerance range |
|---|---|---|---|---|---|---|
| Palmitoleic ω7 acid (C16:1) | Mean | 0.57 | 0.61 | 0.62 | 0.46 – 0.86 | 0.46 – 0.86 |
| | SD | 0.1 | 0.1 | 0.11 | | |
| Palmitoleic ω9 acid (C16:1) | Mean | 0.02 | 0.02 | 0.02 | Not available | 0.01 – 0.03 |
| | SD | 0.01 | 0 | 0 | | |
| Heptadecanoic acid (C17:0) | Mean | 0.1 | 0.11 | 0.11 | <0.1 – 0.1 | 0.08 – 0.13 |
| | SD | 0.02 | 0.01 | 0.01 | | |
| Stearic acid (C18:0) | Mean | 2.86 | 2.9 | 2.94 | 2.15 – 3.4 | 2.41 – 3.28 |
| | SD | 0.37 | 0.38 | 0.39 | | |
| Oleic cis isomer acid (C18:1) | Mean | 0.97 | 0.92 | 0.93 | Not available | 0.67 – 1.7 |
| | SD | 0.28 | 0.18 | 0.18 | | |
| Oleic acid (C18:1) | Mean | 17.43 | 19.73 | 19.64 | 13.4 – 22.0 | 15.83 – 21.12 |
| | SD | 1.01 | 1.05 | 1.13 | | |
| Linoleic acid (C18:2) | Mean | 52.47 | 50.1 | 50.21 | 36.3 – 64.0 | 47.26 – 56.14 |
| | SD | 2.42 | 2.08 | 2.12 | | |
| Linoleic trans isomer acid (C18:2) | Mean | 0.05 | 0.04 | 0.05 | Not available | 0.02 – 0.06 |
| | SD | 0.01 | 0.01 | 0.01 | | |
| Alpha linolenic acid (C8:3)) | Mean | 0.16 | 0.15 | 0.16 | <0.1 – 0.62 | 0.12 – 0.17 |
| | SD | 0.03 | 0.02 | 0.02 | | |
| Arachidic acid (C20:0) | Mean | 0.29 | 0.3 | 0.31 | 0 – 0.48 | 0.2 – 0.33 |
| | SD | 0.06 | 0.05 | 0.06 | | |
| Eicosenoic acid (C20:1) | Mean | 0.06 | 0.06 | 0.06 | <0.1 – 0.1 | 0.04 – 0.07 |
| | SD | 0.01 | 0.01 | 0.01 | | |
| Behenic acid (C22:0) | Mean | 0.1 | 0.11 | 0.12 | 0 – 0.27 | 0.05 – 0.13 |
| | SD | 0.03 | 0.02 | 0.02 | | |
| Lignoceric acid (C24:0) | Mean | 0.08 | 0.07 | 0.07 | 0 – 0.3 | 0.04 – 0.12 |
| | SD | 0.02 | 0.02 | 0.03 | | |

### *5.3.3 Amino acids*

Levels of 18 amino acids were measured in seed from 'Coker 312' and GHB119 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 for the levels are given in Table 9 and can be summarised as follows:

- For alanine, arginine, aspartic acid, cysteine, glutamic acid, glycine, isoleucine, leucine, lysine, phenylalanine, proline, threonine and valine the means for GHB119 seeds were significantly lower in the over-all analysis than the means for 'Coker 312' seeds. These lower levels are reflected in the significantly lower protein levels in GHB 119 compared with 'Coker 312' (refer to Table 7) but the means obtained for the amino acid levels in GHB119 fall within both the range found in the published literature and the tolerance range of the commercial cultivars.
- There was a significant TREAT x SITE interaction for histidine, methionine, serine, tryptophan and tyrosine. The by-site analyses for these amino acids (except histidine which was significantly lower across the majority of sites) did not show a consistent trend.

**Table 9: Percentage dry weight (dw), relative to total dry weight, of amino acids in fuzzy cottonseed from 'Coker 312' and GHB119**

| Amino acid | Calculated variable | A:'Coker 312' | B: GHB119 not sprayed | C: GHB119 - sprayed | Combined literature range | Tolerance range |
|---|---|---|---|---|---|---|
| Alanine | Mean | 1.08 | 0.95 | 0.94 | 0.42 – 1.51 | 0.77 – 1.22 |
| | SD | 0.08 | 0.1 | 0.09 | | |
| Arginine | Mean | 3.4 | 2.74 | 2.66 | 1.05 – 4.4 | 2.08 – 3.43 |
| | SD | 0.34 | 0.41 | 0.36 | | |
| Aspartic acid | Mean | 2.74 | 2.29 | 2.25 | 1.0 – 3.55 | 1.53 – 2.66 |
| | SD | 0.23 | 0.3 | 0.26 | | |
| Cysteine | Mean | 0.85 | 0.75 | 0.71 | 0.16 – 0.86 | 0.44 – 1.34 |
| | SD | 0.1 | 0.11 | 0.09 | | |
| Glutamic acid | Mean | 5.53 | 4.7 | 4.55 | 1.96 – 8.16 | 3.22 – 5.65 |
| | SD | 0.54 | 0.66 | 0.62 | | |
| Glycine | Mean | 1.2 | 1.03 | 1.0 | 0.44 – 1.58 | 0.77 – 1.22 |
| | SD | 0.09 | 0.12 | 0.12 | | |
| Histidine | Mean | 0.88 | 0.76 | 0.74 | 0.31 – 1.03 | 0.66 – 1.0 |
| | SD | 0.07 | 0.08 | 0.09 | | |
| Isoleucine | Mean | 0.97 | 0.83 | 0.81 | 0.35 – 1.17 | 0.66 – 1.0 |
| | SD | 0.07 | 0.09 | 0.09 | | |
| Leucine | Mean | 1.68 | 1.45 | 1.41 | 0.63 – 2.23 | 1.2 – 1.77 |
| | SD | 0.14 | 0.16 | 0.16 | | |
| Lysine | Mean | 1.32 | 1.15 | 1.13 | 0.52 – 1.65 | 0.98 – 1.44 |
| | SD | 0.08 | 0.1 | 0.11 | | |
| Methionine | Mean | 0.59 | 0.49 | 0.45 | 0.15 – 0.54 | 0.22 – 0.44 |
| | SD | 0.32 | 0.24 | 0.21 | | |
| Phenylalanine | Mean | 1.57 | 1.33 | 1.3 | 0.54 – 2.03 | 1.09 – 1.66 |
| | SD | 0.15 | 0.17 | 0.17 | | |
| Proline | Mean | 1.18 | 0.99 | 0.97 | 0.41 – 1.39 | 0.66 – 1.43 |
| | SD | 0.17 | 0.13 | 0.15 | | |
| Serine | Mean | 0.99 | 0.86 | 0.83 | 0.5 – 1.63 | 0.66 – 1.11 |
| | SD | 0.11 | 0.13 | 0.13 | | |
| Threonine | Mean | 0.82 | 0.72 | 0.71 | 0.34 – 1.21 | 0.55 – 1.0 |
| | SD | 0.06 | 0.08 | 0.08 | | |
| Trytophan | Mean | 0.38 | 0.33 | 0.32 | 0.1 – 0.49 | 0.2 – 0.47 |
| | SD | 0.07 | 0.06 | 0.06 | | |
| Tyrosine | Mean | 0.65 | 0.59 | 0.58 | 0.32 – 1.17 | 0.4 – 0.8 |
| | SD | 0.09 | 0.09 | 0.1 | | |
| Valine | Mean | 1.44 | 1.2 | 1.2 | 0.45 – 1.67 | 0.89 – 1.6 |
| | SD | 0.16 | 0.17 | 0.15 | | |

### *5.3.4 Anti-nutrients*

Levels of key anti-nutrients in seeds from GHB119 and 'Coker 312' are given in Table 10. The levels of malvalic and sterculic acids were significantly higher in seeds from GHB119 than from 'Coker 312' but the levels were not outside those found either in the combined literature range or the tolerance range for this analyte. There was a significant TREAT x SITE interaction for free gossypol, total gossypol, phytic acid and dihydrosterculic acid between seeds from GHB119 and those from 'Coker 312'. The by-site analyses indicated that, in the majority of sites, there was no significant difference between seeds from GHB119 and 'Coker 312' except in the case of dihydrosterculic acid which was at a significantly higher level in GHB119 in most sites. The mean level of dihydrosterculic acid in GHB119 fell within both the published literature range and the tolerance range.

**Table 10: Levels of anti-nutrients in fuzzy cottonseed from 'Coker 312' and GHB119**

| Anti-nutrient | Calculated variable | A:'Coker 312' | B: GHB119 not sprayed | C: GHB119 - sprayed | Combined literature range | Tolerance range |
|---|---|---|---|---|---|---|
| Free gossypol (% dw) | Mean | 0.52 | 0.52 | 0.51 | 0.23 – 1.4 | 0.37 – 1.02 |
| | SD | 0.08 | 0.08 | 0.08 | | |
| Total gossypol (% dw) | Mean | 0.72 | 0.73 | 0.77 | 0.46 – 1.99 | 0.61 – 1.3 |
| | SD | 0.13 | 0.12 | 0.12 | | |
| Phytic acid (% dw) | Mean | 1.56 | 1.42 | 1.4 | 0.85 – 2.57 | 0.72 – 2.09 |
| | SD | 0.44 | 0.3 | 0.35 | | |
| Malvalic acid (%tot fat) | Mean | 0.43 | 0.48 | 0.5 | 0.17 – 1.5 | 0.23 – 0.66 |
| | SD | 0.07 | 0.09 | 0.08 | | |
| Sterculic acid (% tot fat) | Mean | 0.19 | 0.25 | 0.25 | 0.12 – 0.92 | 0.16 – 0.29 |
| | SD | 0.04 | 0.04 | 0.04 | | |
| Dihydrosterculic acid (% tot fat) | Mean | 0.18 | 0.26 | 0.27 | 0.11 – 0.5 | 0.17 – 0.38 |
| | SD | 0.04 | 0.04 | 0.04 | | |

### *5.3.5 Minerals and vitamins*

Levels of key minerals and vitamins in seeds from GHB119 and 'Coker 312' are given in Table 11. The level of phosphorus was significantly lower in seeds from GHB119 than in the 'Coker 312' control in the over-all analysis but the mean values fell within both the published literature range and the tolerance range. For all other analytes, there were significant TREAT x SITE interactions. The by-site analyses indicated that for calcium the majority of sites showed a significant difference between the 'Coker 312' control and the unsprayed GHB119. For all of the other analytes the majority of sites did not show a significant difference between the control and GHB119 (both sprayed and unsprayed).

**Table 11: Levels of key minerals and vitamins in fuzzy cottonseed from 'Coker 312' and GHB119.**

| Analyte | Calculated variable | A:'Coker 312' | B: GHB119 not sprayed | C: GHB119 - sprayed | Combined literature range | Tolerance range |
|---|---|---|---|---|---|---|
| Calcium (% dw) | Mean | 0.11 | 0.1 | 0.1 | 0.09 – 0.33 | 0.08 – 0.19 |
| | SD | 0.03 | 0.03 | 0.03 | | |
| Phosphorus (% dw) | Mean | 0.58 | 0.54 | 0.55 | 0.31 – 0.86 | 0.38 – 0.72 |
| | SD | 0.09 | 0.07 | 0.07 | | |
| Potassium (% dw) | Mean | 1.18 | 1.18 | 1.18 | 0.96 – 1.42 | 0.98 – 1.29 |
| | SD | 0.1 | 0.07 | 0.08 | | |

| Analyte | Calculated variable | A:'Coker 312' | B: GHB119 not sprayed | C: GHB119 - sprayed | Combined literature range | Tolerance range |
|---|---|---|---|---|---|---|
| Magnesium (% dw) | Mean | 0.46 | 0.39 | 0.37 | 0.27 – 0.49 | 0.32 – 0.49 |
| | SD | 0.09 | 0.05 | 0.05 | | |
| Iron (mg/kg dw)) | Mean | 54.4 | 47.8 | 46.9 | 23.2 – 160.0 | 23.4 – 71.9 |
| | SD | 11.1 | 11.0 | 10.7 | | |
| Zinc (mg/kg dw) | Mean | 42.1 | 36.7 | 34.2 | 17.8 – 63.0 | 19.9 – 57.0 |
| | SD | 9.7 | 10.0 | 9.2 | | |
| Total tocopherols (mg/kg dw) | Mean | 111.7 | 107.6 | 102.3 | 44.8 – 436* | 82.1 - 140 |
| | SD | 20.7 | 16.6 | 15.5 | | |

***** Reference range for total tocopherols 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 Additional compositional analyses*

In addition to analysing the fuzzy cottonseed in GHB119, the Applicant supplied compositional data for fuzzy cottonseed from 'TwinLink'™. The cotton was grown at a total of seven U.S. sites in Georgia, Arkansas, Mississippi and Texas. In every trial there were three plots of each of 'Coker 315', 'TwinLink'™ unsprayed and 'TwinLink'™ sprayed with glufosinate ammonium herbicide, in a randomised complete block design. In total, 63 fuzzy cottonseed samples from 7 sites were analysed for 50 components.

FSANZ has considered the results from these analyses. As an example, the levels of proximates and fibre for 'TwinLink'™/'Coker 315' are shown in Table 12; these are consistent with the levels for GHB119/'Coker 312' given in Table 7).

**Table 12: Percentage dry weight (dw) of proximates and fibre in fuzzy cottonseed from 'Coker 315' and 'TwinLink'™**

| Analyte | Calculated variable | 'Coker 315' | 'TwinLink' not sprayed | 'TwinLink' sprayed | Combined literature range |
|---|---|---|---|---|---|
| Protein (%dw) | Mean | 23.3 | 22.38 | 22.38 | 11.7 – 34.2 |
| | SD | 2.29 | 2.4 | 2.29 | |
| Fat (%dw) | Mean | 18.4 | 18.53 | 18.76 | 11.8 – 36.3 |
| | SD | 2.4 | 2.38 | 1.83 | |
| Carbohydrate (%dw)[1] | Mean | 54.38 | 55.1 | 54.82 | 36.4 – 74.4 |
| | SD | 2.37 | 2.63 | 2.27 | |
| ADF (%dw) | Mean | 39.64 | 39.53 | 38.92 | 29.0 – 66.9 |
| | SD | 2.64 | 2.89 | 2.6 | |
| NDF (%dw) | Mean | 46.05 | 46.44 | 47.24 | 38.1 – 71.4 |
| | SD | 3.77 | 3.75 | 3.05 | |
| Ash (%dw) | Mean | 3.94 | 4.01 | 4.04 | 3.2 – 5.0 |
| | SD | 0.33 | 0.22 | 0.3 | |

Results for the other analytes for TwinLink'™/'Coker 315' are not shown but were provided by the Applicant. The following is a summary of the results:

- The majority of analytes (including all the amino acids) showed no significant differences between treatment means over all sites.

- There were significant differences between the comparison of 'Coker 315' vs unsprayed 'TwinLink'™ and Coker 315 vs sprayed 'TwinLink'™ for magnesium, cystine, myristic acid and behenic acid. However, all means were in the reference range for commercial cotton seeds.
- Results for lysine were ambiguous as there was a difference between the comparison 'Coker 315' vs 'TwinLink' unsprayed and 'Coker 315' vs 'TwinLink' sprayed.
- For those analytes in which treatment x site interactions occurred (protein, fat, ash, total carbohydrate, NDF, calcium, iron, α-tocopherol, anti-nutrients (except free gossypol) and six of the ten fatty acids, there was no significant difference between seeds of 'Coker 315' and 'TwinLink'™ in the majority of sites.
- Most results for lignoceric acid were below the limit of quantification and therefore a statistical analysis was not valid.

### *5.3.7 Compositional data for processed commodities*

The Applicant obtained comparative compositional data for a range of cottonseed processed commodities (lint, linters, delinted cottonseed, hulls, meal, toasted meal, crude oil and refined deodorised oil) derived from 'TwinLink'™ and 'Coker 315' cotton grown in a single trial in Levelland (Texas).

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 in cottonseed meal and toasted cottonseed meal, the mean levels for both the 'Coker 315' control and 'TwinLink'™ 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 GHB119 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, data were obtained from two further studies using the GM cultivar 'TwinLink'™, which is the product of a conventional cross between line GHB119 and another GM cotton line T304-40. These studies measured various constituents in fuzzy seed as well as in processed products.

For fuzzy cottonseed, across most of the categories but most notably in the amino acids, some significant differences were found in individual analytes between seeds from GHB119 and those of the non-GM control. 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. The mean analyte levels found in seeds from GHB119 fell within the range of natural variation in commercial cotton cultivars. The compositional analysis of seeds from 'TwinLink'™ indicated few significant differences from seeds of the non-GM control.

For processed products derived from 'TwinLink'™ 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 containing event GHB119 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 GHB119 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 GHB119 and these indicate it is equivalent in composition to conventional cotton varieties. The introduction of cotton line GHB119 into the food supply is therefore expected to have little nutritional impact.

The Applicant submitted one animal feeding study with cotton GHB119, the results of which are included below.

## 6.1 Broiler feeding study

**Studies submitted:**

Stafford, J.M. (2008). Broiler Chicken Nutritional Equivalency Study with TwinLink Cotton. Project No. TX99X106, Bayer CropScience (unpublished).

Kowite, W.J. (2008). Production of Cottonseed Samples of TwinLink Cotton, the Non-transgenic Counterpart and a Non-transgenic Commercial Cotton Variety, USA, 2007. Study No. Cp07B004, Bayer CropScience (unpublished).

These types of studies are designed to specifically measure carcass characteristics and are not intended to be toxicity studies.

This 42-day study compared growth, performance and carcass yield of Ross #308 broiler chickens fed diets containing approximately 10% toasted cottonseed meal (refer to Figure 1) from seeds of 'TwinLink'™ (sprayed twice with glufosinate ammonium herbicide) 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 a field trial in Hockley County, Texas (US) and the seed was shipped to Texas A & M University for processing into toasted meal. DNA characterisation, using discriminating PCR analysis, was undertaken of both the raw commodity and the test diets derived from the raw commodity to validate identity and purity.

The broiler study was undertaken using appropriate, internationally recognised Good Laboratory Practice regulations pertinent to the execution of feeding studies. 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. This number of birds satisfied the requirement regarding sufficient statistical power to detect differences between treatment groups.

The Analysis of Variance function in SYSTAT[17] 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).

A number of birds across all treatments exhibited clinical signs or death. These were considered to be associated with the Ross #308 strain used for the study, and not linked to any specific treatment. Birds fed meal derived from 'TwinLink'™ showed significantly higher mean wing weights and mean wing tissue yield than birds fed non-GM meal. No statistically significant effects were detected in any of the other measured parameters that were attributable to the consumption of cottonseed meal derived from 'TwinLink'™ cotton. Broilers consuming a diet containing 'TwinLink'™ 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 'TwinLink'™ 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[18]

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

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

[17] SYSTAT website is at http://www.systat.com/
[18] All website references were current as at 24 May 2010

Arnaut, G., Boets, A., Vanneste, S., Van Rie, J. and Van Houdt, S. (2007) Nucleic acids encoding a novel Cry2Ae *Bacillus thuringiensis* insecticidal protein. US Patent Number 7,265,269 B2. http://www.freepatentsonline.com/7265269.html. Accessed 23 February 2010.

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.

Baum, J.A., Chu, C.-R., Donovan, W.P., Gilmer, A.J. and Rupar, M.J. (2003) Lepidopteran-active *Bacillus thuringiensis* d-endotoxin polynucleotides, compositions, and methods of use. In: Monsanto Technology LLC. eds. 09/661322(US 6,593,293): St Louis, MO (US). http://www.faqs.org/patents/app/20080229449. Accessed 12 April 2010.

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.

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

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.

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. http://www.codexalimentarius.net/web/standard_list.do?lang=en. Accessed 3 May 2010

Crooks, S. (2008) *Australian Honeybee Industry Survey 2006 - 07*. RIRDC Publication No. 08/170, Australian Government Rural Industries Research and Development Corporation, https://rirdc.infoservices.com.au/downloads/08-170.pdf. Accessed 17 March 2010.

D'Halluin, K., De Block, M., Denecke, J., Janssens, J., Leemans, J., Reynaerts, A. and Botterman, J. (1992) The *bar* gene as selectable and screenable marker in plant engineering. *Methods in Enzymology* :415-426.

Damgaard, P.H., Granum, P.E., Bresciani, J., Torregrossa, M.V., Eilenberg, J. and Valentino, L. (1997) Characterization of *Bacillus thuringiensis* isolated from infections in burn wounds. *FEMS Immunology & Medical Microbiology* 18:47-53.

De Almeida, E.R.P., Gossele, V., Muller, C.G., Dockx, J., Reynaerts, A., Botterman, J., Krebbers, E. and Timko, M.P. (1989) Transgenic expression of two marker genes under the control of an *Arabidopsis rbcS* promoter: sequences encoding the Rubisco transit peptide increase expression levels. *Molecular and General Genetics* 218:78-86.

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.

Donovan, W.P., Dankocsik, C.C., Gilbert, M.P., Gawron-Burke, M.C., Groat, R.G. and Carlton, B.C. (1988) Amino acid sequence and entomocidal activity of the P2 crystal protein. *The Journal of Biological Chemistry* 263:561-567.

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. http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1211902590265.htm. Accessed 5 April 2010.

EPA. (2008) Bacillus thuringiensis Cry2Ae in Cotton; Temporary Exemption from the Requirement for a Tolerance. *Federal Register* 73:52591-52594, U.S. Environmental Protection Agency. http://www.epa.gov/EPA-PEST/2008/September/Day-10/p20728.htm. Accessed 1 March 2010.

Federici, B.A. (1999) *Bacillus thuringiensis* in Biological Control. In: Bellows, T.W. and Fisher, T.S. eds. *Handbook of Biological Control*. Chapter 21. Academic Press, San Diego, pp. 575-593.

Fling, M.E., Kopf, J. and Richards, C. (1985) Nucleotide sequence of the transposon Tn7 gene encoding an aminoglycoside-modifying enzyme, 3"(9)-O-nucleotidyltransferase. *Nucleic Acids Research* 13(19):7095-7106.

FSANZ (2002) *Application A436 - Oil and Linters Produced from Insect-Protected Cotton Containing Event 15985*. Report Prepared by Food Standard Australia New Zealand. http://www.foodstandards.gov.au/foodstandards/applications/applicationa436insectprotectedcottonevent15985/index.cfm. Accessed 19 March 2010.

FSANZ (2008) *Application A595 - Food Derived from Insect-Protected Corn Line MON89034*. Report Prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/foodstandards/applications/applicationa595foodd3492.cfm. Accessed 12 April 2010.

FSANZ (2010) *Application A1028 - Food Derived from Insect-Protected and Herbicide-Tolerant Cotton Line T304-40: Safety Assessment Report*. Report prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm. Accessed 20 May 2010.

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.

Green, M., Heumann, M., Sokolow, R., Foster, L.R., Bryant, R. and Skeels, M. (1990) Public health implications of the microbial pesticide *Bacillus thuringiensis*: an epidemiological study, Oregon 1985 - 86. *American Journal of Public Health* 80:848-852.

Hajdukiewicz, P., Svab, Z. and Maliga, P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Molecular Biology* 25:989-994.

Harpster, M.H., Townsend, J.A., Jones, J.D.C., Bedbrook, J. and Dunsmuir, P. (1988) Relative strengths of the 35S cauliflower mosaic virus 1', 2' and nopaline synthase promoters in transformed tobacco, sugarbeet and oilseed rape callus tissue. *Molecular and General Genetics* 212:182-190.

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.

Hernandez, E., Ramisse, F., Ducoureau, J.-P., Cruel, T. and Cavallo, J.-D. (1998) *Bacillus thuringiensis* subsp. *konkukian* (serotype H34) superinfection: case report and experimental evidence of pathogenicity in immunosuppressed mice. *Journal of Clinical Microbiology* 36:2138-2139.

Hernández-Rodríguez, C.S., Van Vliet, A., Bautsoens, N., Van Rie, J. and Ferré, J. (2008) Specific binding of *Bacillus thuringiensis* Cry2A insecticidal proteins to a common site in the midgut of *Helicoverpa* species. *Applied and Environmental Microbiology* 74:7654-7659.

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* d-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. 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*. available online at http://www.cropcomposition.org/cgi-perl/search_ora.cgi. Accessed 13 March 2007.

Jackson, S.G., Goodbrand, R.B., Ahmed, R. and Kasatiya, S. (1995) *Bacillus cereus* and *Bacillus thuringiensis* isolated in a gastroeneritis outbreak investigation. *Letters in Applied Microbiology* 21:103-105.

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.

Kepiro, J. (2008) *Experimental Use Permit Request for "Bacillus thuringiensis Cry2Ae Insecticidal protein as Expressed in Cotton Plants" , and "Bacillus thuringiensis Cry1Ab and Cry2Ae Insecticidal Proteins as Expressed in Combined Trait Cotton (Twinlink Cotton) Plants"*. Replacement Pages for 264-EUP-RUG submitted to U.S. Environmental Protection Agency by Bayer CropScience LP - BioScience. http://www.regulations.gov/search/Regs/contentStreamer?objectId=0900006480621ee2&disposition=attachment&contentType=pdf. Accessed 10 May 2010.

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.

Krell, R. (1996) *Value-Added Products from BeeKeeping*. Chapter 3: Pollen, FAO Agricultural Services Bulletin No. 124, Food and Agriculture Organization of the United Nations. http://www.fao.org/docrep/w0076e/w0076e00.htm#con. Accessed 9 February 2010.

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.

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.

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

Moar, W.J., Trumble, J.T., Hice, R.H. and Backman, P.A. (1994) Insecticidal activity of the CryIIA protein from the NRD-12 isolate of *Bacillus thuringiensis* subsp. *kurstaki* expressed in *Escherichia coli* and *Bacillus thuringiensis* and in a leaf-colonizing strain of *Bacillus cereus*. *Applied and Environmental Microbiology* 60:896-902.

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. Accessed 29 March 2010.

Murakami, T., Anzai, H., Imai, S., Satoh, A., Nagaoka, K. and Thompson, 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.

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. Accessed 10 February 2010.

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. Accessed 22 February 2010.

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. Accessed 22 February 2010.

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://appli1.oecd.org/olis/2007doc.nsf/linkto/env-jm-mono(2007)14. Accessed 23 February 2010.

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. Accessed 9 February 2010.

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. Accessed 9 February 2010.

Oka, A., Sugisaki, H. and Takanami, M. (1981) Nucleotide sequence of the kanamycin resistance transposon Tn903. *Journal of Molecular Biology* 147:217-226.

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.

Reynaerts, A. and De Sonville, A. (2002) Method for Agrobacterium mediated transformation of cotton. US Patent Number 6,483,013 B). http://www.freepatentsonline.com/6483013.html. Accessed 26 February 2010.

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. Accessed 24 February 2010.

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.

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

Valadares de Amorim, G., Whittome, B., Shore, B. and Levin, D.B. (2001) Identification of *Bacillus thuringiensis* subsp. *kurstaki* strain HD1-like bacteria from environmental and human samples after aerial spraying of Victoria, British Columbia, Canada, with Foray 48B. *Applied and Environmental Microbiology* 67:1035-1043.

Verdaguer, B., de Kochko, A., Beachy, R.N. and Fauquet, C. (1996) Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosiac virus (CVMV) promoter. *Plant Molecular Biology* 31:1129-1139.

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.

Widner, W.R. and Whiteley, H.R. (1990) Location of the dipteran specificity region in a lepidopteran-dipteran crystal protein from *Bacillus thuringiensis*. *Journal of Bacteriology* 172:2826-2832.

Yu, J., Tan, L., Liu, Y. and Pang, Y. (2002) Phylogenetic analysis of *Bacillus thuringiensis* based on PCR amplified fragment polymorphisms of flagellin genes. *Current Microbiology* 45:139-143.

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

Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J.J., Farrand, S.K. and Winans, S.C. (2000) The bases of crown gall tumorigenesis. *Journal of Bacteriology* 182:3885-3895.