Bayer CropScience

Application to FSANZ to Vary Food Standard 1.5.2 to Include the Dual Herbicide Tolerant Cotton (*G. hirsutum*) Event GHB811

Prepared by

Submitted April 2017

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDED ELECTRONIC DOCUMENTS

Node B.1 (a) Herrmann, K.M. (1995) The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. The Plant Cell, 7:907-919. Document no. M-269843-01.

Node A.1 (a) Lebrun, M., Sailland, A., Freyssinet, G. (1997) 5-enol pyruvylshikimate-e-phosphate synthase mutee, gene codant pour cette proteine et plantes transformees contenant ce gene. Patent Application: WO9704103-A 1. Document no. M-216526-01.

Node A.2 (a), (i)
Organisation for Economic Co-Operation and Development (OECD) (1997) Series on harmonization of regulatory oversight in biotechnology No.6 - Consensus document on information used in the assessment of environmental applications involving Pseudomonas. Document no. M-357528-01.

Node A.3 (a) Description of the GHB811 cotton transformation methodology. Unpublished Bayer Crop Science Report. Document no. M-543362-01.

Node A.1 (a) Boudec, P.; Rodgers, M.; Dumas, F.; Sailland, A.; Bourdon, H. (2001) Mutated hydroxyphenylpyruvate dioxygenase, DNA sequence and isolation of plants which contain such a gene and which are tolerant to herbicides. US Patent US6245968B1. Document no. M-229534-01.

Node A.3 (b), (i) and (ii)
Description of vector pTSIH09. Unpublished Bayer Crop Science Report. Document no. M-388224-02.

Node A.3 (c), (i) (CCI)
Detailed insert characterization and confirmation of the absence of vector backbone sequence in cotton GHB811. Unpublished Bayer Crop Science Report. Document no. M-572036-01. (**Contains confidential commercial information**)

Node A.3 (c), (iii) (CCI)
DNA sequence determination of the transgenic and insertion loci of cotton GHB811. Unpublished Bayer Crop Science Report. Document no. M-533573-01. (**Contains confidential commercial information**)

Node A.3 (c), (v) (CCI)
Bioinformatics analysis of the GHB811 cotton insertion locus. Unpublished Bayer Crop Science report. Document no. M-581222-01. (**Contains confidential commercial information**)

Node A.3 (c), (v) (CCI)
GHB811 cotton - Identification of Open Reading Frames and homology search of sequences ≥30 amino acids to known allergens and toxins. Unpublished Bayer Crop Science Report. Document no. M-575144-01. (**Contains confidential commercial information**)

Node A.3 (e), (i) (CCI)

Structural stability analysis of cotton GHB811. Unpublished Bayer Crop Science Report. Document no. M-548778-01. (**Contains confidential commercial information**)

Node A.3 (e), (ii)

GHB811 cotton - Inheritance of the insert over three generations. Amended analytical report. Unpublished Bayer Crop Science Report. Document no. M-547925-02.

Node A.3 (e), (ii)

2016) GHB811 cotton – Protein expression analyses of field samples grown in the USA during 2015. Unpublished Bayer Crop Science Report. Document no. M-574232-01.

Node B.1 (a) Brownlee, J. M.; Johnson-Winters, K.; Harrison,D H T.; Moran G. R (2004) Structure of the Ferrous Form of (4-Hydroxyphenyl)pyruvate Dioxygenase from Streptomyces avermitilis in Complex with the Therapeutic Herbicide, NTBC. Biochemistry 43:6370-6377. Document no. M-358228-01.

Node B.1 (a) Fritze, I. M., Linden, L., Freigang, J., Auerbach, G., Huber, R., Steinbach, S. (2004) The crystal structures of *Zea mays* and *Arabidopsis* 4-hydroxyphenylpyruvate dioxygenase. *Plant Physiology* 134: 1388 – 1400. Document no. M-359884-01.

Node B.1 (a) Pallett, K. E.; Cramp, S. M.; Little, J. P.; Veerasekaran, P.; Crudace, A. J.; Slater, A. E. (2001) Isoxaflutole: the background to its discovery and the basis of its herbicidal properties. *Pest Manag. Sci.* 57:133-142. Document no. M-357534-01.

Node B.1 (a) Dayan, F. E.; Duke, S. O.; Sauldubois, A.; Singh, N.; McCurdy, C.; Cantrell, C. (2007) p-Hydroxyphenylpyruvate dioxygenase is a herbicidal target site for beta-triketones from Leptospermum scoparium. *Phytochemistry* 68:2004-2014. Document no. M-357499-01.

Node B.1 (a) Fischer, K. (2008) Generation of a herbicide-resistant 4-hydroxyphenylpyruvate dioxygenase from *Arabidopsis thaliana* and *Pseudomonas fluorescens*: Saturation mutagenesis and biochemical characterization Scientific work in support of Bachelor of Science. University of Mainz. Document no M-359847-01.

Node B.1 (a) Healy, M. L.; Schoenbrunn, E. (2006) Biochemical characterization of 5-enolpyruvylshikimate-3-phoaphate synthase (wtEPSPS) and its mutated glyphosate-resistant form (2mEPSPS). Research thesis, The University of Kansas. Document no. M-267370-01.

Node B.1 (a) Characterization of 2mEPSPS protein purified from GHB811 cotton and comparability with the recombinant 2mEPSPS protein batch 1417_2mEPSPS. Unpublished Bayer Crop Science Report. Document no. M-568145-01.

Node B.1 (a) Characterization of the recombinant 2mEPSPS protein batch 1417_2mEPSPS. Unpublished Bayer Crop Science Report. Document no. M-497839-01.

Node B.1 (a) Characterization of HPPD W336 protein purified from GHB811 cotton and comparability with the recombinant HPPD W336 protein batch 1411_HPPD W336. Unpublished Bayer Crop Science Report. Document no. M-576569-01.

Node B.1 (a) Characterization of the recombinant HPPD W336 protein batch no 1411_HPPD W336. Unpublished Bayer Crop Science Report. Document no. M-497842-01.

Node B.5 (c) GHB811 cotton: Production, processing, and analysis of resultant fractions, 2014/2015. Unpublished Bayer Crop Science Report. Document no. M-574125-01.

Node B.1 (d) GHB811 cotton – Identification of open reading frames and homology search of sequences of ≥30 amino acides to allergens and toxins. Unpublished Bayer Crop Science Report. Document no. M-575144-01.

Node B.2 (a) (i) 2mEPSPS protein - Amino acid sequence homology search with known allergens and known toxins. Unpublished Bayer Crop Science Report. Document no. M-445681-04.

Node B.2 (a) (i) HPPD W336 protein - Amino acid sequence homology search with known toxins. Unpublished Bayer Crop Science Report. Document no. M-445678-04.

Node B.2 (b) (i) Endogenous food allergens in corn. Unpublished Bayer Crop Science Position Paper. Document no. M-402872-01.

Node B.2 (b) (i) Safety assessment of the double mutant 5-enol-pyruvylshikimate-3-phosphate synthase (2mepsps) protein. Unpublished Bayer Crop Science Report. Document no M-278169-01.

Node A.3 (c) (iii) (CCI) Description of the base pair sequence of the double mutated maize 5-enol pyruvylshikimate-3-phosphate synthase gene (2m EPSPS). Unpublished Bayer Crop Science Report. Document no. M-234184-01. **(Contains confidential commercial information)**.

Node B.5 (a) GHB811 cotton - Composition assessment of GHB811 cotton grown in the USA during 2014 and 2015. Unpublished Bayer Crop Science Report. Document no. M-566678-01.

Executive Summary

Bayer CropScience Pty Ltd seeks to vary FSANZ Standard 1.5.2 to allow the use of genetically modified cotton (*Gossypium hirsutum*) derived from transformation event GHB811 *G. hirsutum* in the Australian and New Zealand food industries. Four food products are derived from cotton: oil, meal, hulls and linters. Refined oil is the primary food product consumed by humans in Australia, with the other cotton food products, as well as whole cottonseeds, used as components of animal feed.

Bayer CropScience has developed a dual-herbicide tolerant line of GM cotton (*G. hirsutum*) that will be commercialized in the USA and Brazil and possibly other cotton cultivation countries in the future. Planting double-herbicide tolerant cotton GHB811 varieties provides growers with new options for weed control using isoxaflutole (IFT) and/or glyphosate herbicide. Glyphosate is widely used in cotton and other agricultural production systems. IFT herbicide offers and alternative weed control option for the cotton grower to help manage problem weed species and as an alternative mode of action tool to help slow the spread of herbicide resistant weeds. With IFT, a new mode of action is introduced in cotton that is efficacious against many weeds currently found in cotton fields.

GHB811 cotton was developed through *Agrobacterium*-mediated transformation using the vector pTSIH09 containing *hppdPfW336-1Pa* and *2mepsps* expression cassettes. The OECD identifier is BCS-GH811-4.

(i) The double mutant 5-enol pyruvylshikimate-3-phosphate synthase (*2mepsps*) gene that encodes for the 2mEPSPS protein. The *2mepsps* coding sequence was developed by introducing two point mutations to the wild-type *epsps* gene cloned from maize (*Zea mays*). Expression of the 2mEPSPS protein confers tolerance to glyphosate herbicides. FSANZ has previously assessed the 2mEPSPS protein, as expressed by the *2mepsps* gene, in the Bayer CropScience applications for GlyTol cotton (A614) and FG72 soy bean (A1051).

(ii) The *hppdPf W336* gene encodes for the HPPD W336 protein. The *hppdPf W336* coding sequence was developed by introducing a single point mutation to the wild type *hppd* gene derived from *Pseudomonas fluorescens*. Expression of the HPPD W336 protein confers tolerance to isoxaflutole herbicides. The *hppdPf W336* gene has been used to confer HPPD inhibitor tolerant properties to soy bean in the past. FSANZ has previously assessed the HPPD W336 protein, as expressed by the *hppdPf W336* gene, in the Bayer CropScience application for FG72 soy bean (A1051).

Cotton is primarily used worldwide for its lint. Lint is produced on the seed coat, and is spun into fine strong threads. Only the United States and a few other countries have developed major commercial uses for the seed. Raw unprocessed cottonseed may be fed to ruminants in the form of cottonseed meal and hulls or the seed can be processed for oil, the primary component consumed by humans. Linters, the short fibers that remain on the hulls after the removal of the lint have both edible and non-edible use.

The incorporation of the GHB811 transgenic locus in the *G. hirsutum* genome and the safety of proteins expressed by introduced genes, *hppdPf W336* and *2mepsps* have been characterized according to international standards for the safety assessment of biotechnology products. This information is included with this application to support the food safety of the 2mEPSPS and HPPD W336 proteins. Open pollinated *G. hirsutum* varieties containing the GHB811 event will be grown commercially in the cotton producing areas of the USA and Brazil.

Molecular characterization determined that a single copy of the complete T-DNA of the pTSIH09 plasmid was inserted at a single locus of the cotton GHB811 genome. These data also demonstrated the absence of vector backbone sequences in cotton GHB811 gDNA. The DNA sequence of the cotton GHB811 transgenic locus and the corresponding insertion locus was determined. Molecular characterization analysis also demonstrated inheritance and stability of the insert across multiple generations.

Bioinformatics analysis of the full DNA sequence revealed no evidence supporting cryptic gene expression or unintended effects resulting from the genetic modification.

Food safety evaluation of the 2mEPSPS and HPPD W336 proteins was undertaken utilising guidance provided by Codex (2009). No health-related adverse effects have been associated with the proteins.

5-enolpyruvylshiklmate-3-phosphate synthase (EPSPS) catalyzes the penultimate step of the shikimate pathway, which is responsible for the synthesis of aromatic amino acids and other aromatic compounds in plants, fungi and microorganisms including apicomplexan parasites (Herrmann, K. M.; 1995; M-269843-01; published; Node B.1 (a)). As such, it has been shown that EPSPS enzymes are ubiquitous in nature and are present in food and feed derived from plant and microbial sources. No health-related adverse effects have been associated with these proteins.

The *2mepsps* gene was generated by introducing mutations into the *epsps* gene from maize (*Z mays* L.) that result in two amino acid substitutions. The modified EPSPS (2mEPSPS) enzyme has a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of glyphosate herbicides (Lebrun, M. *et al*., 1997; M-216526-01; Node A.1 (a)). Since the 2mEPSPS protein is derived from maize and has only two amino acid modifications, the safety profile of the novel protein is expected to remain unchanged relative to its wild-type counterpart. EPSPS proteins are present in food and feed from plant and microbial sources with good safety records. Therefore, EPSPS proteins have a history of safe use. The 2mEPSPS protein has been assessed previously by FSANZ in association with approval of the herbicide tolerant cotton event GHB614 (FSANZ A614), and the herbicide tolerant soy bean event FG72 (FSANZ A1051). As food safety of this protein has been established previously, the information provided for the protein within this application will be limited to studies confirming its amino acid sequence and up to date data to confirm lack of amino acid sequence homology with known toxins and allergens.

The coding sequence of the 4-hydroxyphenylpyruvate dioxygenase (HPPD) protein was isolated from the *Pseudomonas fluorescens* strain A32. *P. fluorescens* is a Gram-negative, rod-shaped, motile, asporogenous, aerobic bacterium. *P. fluorescens*, is ubiquitous in the environment, including soil, water and food (OECD; 1997; M-357528-01; Node A.1 (a), (i)). It has many beneficial uses in agriculture, human health and bioremediation. It is not described as allergenic, toxic or pathogenic to healthy humans and animals and has an overall history of safe use. The HPPD W336 protein has no amino acid sequence homology to known allergens and is rapidly degraded in simulated gastric fluid and simulated intestinal fluid assays. The HPPD W336 protein has no amino acid sequence similarity to known toxins and exhibited no effects in acute oral mouse toxicity tests. The protein is known to have a good history of safe use. The HPPD W336 protein too has been assessed for food safety by FSANZ within the approved herbicide tolerant soy bean event FG72 (FSANZ A1051). As food safety for this protein has been established previously, information for this protein within this application will be limited to studies confirming the amino acid sequence of the protein and up to date data to confirm lack of amino acid sequence homology with known toxins and allergens.

The nature of N-glycosylation sites, heat stability and degradation in simulated digestive environments have been established previously for both of these proteins. In addition to this effects of the proteins have been independently tested within acute oral mouse toxicity testing, the associated data which has been presented to FSANZ in association with the approval of FG72 soybean (A1051). It is therefore concluded that GHB811 cotton has negligible impact on the nutritional value of foods derived from cotton.

Part 1 General Information on the Application

1.1 Applicant Details

(a) Applicant (individual organisation's) name

Bayer CropScience Pty Ltd

(b) Name of contact person

(c) Address (street and postal)

Bayer CropScience Pty Ltd
Level 1, 8 Redfern Road
Hawthorn East
Victoria 3123
Australia

(d) Telephone numbers

(e) Email address

(f) Nature of applicant's business

Seeds and traits, biotechnology.

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

Not applicable.

1.2 Purpose of the Application

This application, on behalf of Bayer CropScience Pty Ltd, seeks to vary FSANZ Standard 1.5.2 to allow the use of genetically modified cotton (*G. hirsutum*) derived from transformation event GHB811 cotton in the Australian and New Zealand food industries.

Four food products are derived from cotton: oil, meal, hulls and linters. Refined oil is the primary food product consumed by humans in Australia, with the other cotton food products, as well as whole cottonseeds, used as components of animal feed.

Cotton varieties containing event GHB811 will be approved as a single event for food approval in the major cotton product receival countries of the world. It is anticipated that food products derived from cotton containing this event will enter the Australian and New Zealand food supply via local production and imports from major cotton producing countries such as the United States and Brazil.

1.3 Justification for the Application

The GHB811 transformation event introduced two genes to the *G. hirsutum* genome. These genes confer two novel traits; tolerance to the broad spectrum herbicides glyphosate and isoxaflutole. Cotton varieties containing the GHB811 event will be produced commercially in the major cotton producing countries of the world.

Advantages of GHB811 cotton

The novel traits expressed by cotton varieties containing event GHB811 provide several potential benefits over conventional cotton varieties and other transgenic cotton currently in cultivation. These include:

- Glyphosate is a broad spectrum, post-emergence weed control system that provides an alternative to pre-emergent and residually active compounds, and encourages herbicide use on an as-needed basis.
- Isoxaflutole is also a broad spectrum, post-emergence weed control system that provides an alternative herbicidal mode of action to both pre-emergent, residually active compounds and glyphosate which also allows for herbicide rotation within cotton farming systems.
- Broad spectrum weed control reduces cultivation needs, reducing on-farm fuel consumption, decreasing CO_2 emissions and also importantly improving soil health (Brookes, G. and Barfoot, P.; 2016; http://www.tandfonline.com/doi/full/10.1080/21645698.2016.1192754).

Note:

(a) Any public health and safety issues related to the proposed change including details of target groups and population groups that may be adversely affected
(b) Any consumer choice issues related to the proposed change
(c) Any evidence that the food industry generally or other specific companies have an interest in, or support, the proposed change.

In relation to points (a), (b) and (c) above, the data contained within this submission indicates the general safety of GHB811 cotton-derived foods and their close similarity to non-genetically modified (GM) comparators that have been used in studies. From the work conducted there is no indication that there are public health or safety issues related to the proposed change to Standard 1.5.2 of the Food Standards Code. The section below discussing food safety of GHB811 cotton goes into further detail in this respect.

Consumer choice with respect to the proposed change is anticipated to be dealt with by FSANZ via their assessment of the data included in this package. It should be noted that GHB811 cotton when used in breeding systems to deliver cotton-derived food products will result in the primary food product – cotton seed oil – which contains novel proteins which are below the limit of quantification. This food item therefore does not result in the need for labelling to differentiate it from cotton seed oil derived from non-GM cotton varieties.

As GHB811 *G. hirsutum* is still in the developmental stage with Bayer CropScience, there is no specific information available to indicate that the food industry have interest in, or support, the proposed change to the Standard 1.5.2. However, due to reasonably rapid uptake by the farming community of GM cotton in past years and the impact that this has had on the price per tonne of seed for crushing to oil and meal it may be anticipated that the food industry generally support technology that leads to lower commodity prices for the cotton seed oil that they wish to purchase, use in food production and on-sell. Equally use of lower-priced, cotton seed meal as a feed for animals would be welcomed by animal producers and in turn would have an impact on the price of food derived from livestock for consumers.

Food safety

Tolerance to glyphosate is achieved through expression of a 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) encoded by the *2mepsps* gene derived from *Zea mays* (corn). The *epsps* gene has been widely used in the genetic modification of a number of crop species. EPSPS derived from *Z. mays* has a long history of safe use in agriculture. 5-enolpyruvylshikimate-3-phosphate synthase has been successfully used to engender herbicide tolerance to a significant array of crops for more than 20 years. These crops have been in commercial production over the past decade and longer. Of the EPSPS group, CP4 EPSPS has been the most extensively used and well characterised. FSANZ has assessed several EPSPS previously, including 2mEPSPS (GHB614 cotton, A614) (see Table 2, Section A.2(a)(i)), and have not identified public health or safety concerns.

Tolerance to HPPD inhibitors, such as isoxaflutole, is achieved through expression of HPPD W336 encoded by the *hppdPf W336* gene derived from *Pseudomonas fluorescens*. The HPPD W336 protein has also been assessed by FSANZ previously. No public health or safety concerns were identified associated with the HPPD W336 protein expressed by the Bayer CropScience event FG72 (A1051). In the review of FG72 soy bean, FSANZ stated: "FSANZ has completed a comprehensive safety assessment of food derived from soybean line FG72... This assessment included consideration of (i) the genetic modification to the plant; (ii) the potential toxicity and allergenicity of the novel proteins; and (iii) the composition of soybean line FG72 compared with that of conventional soybean cultivars. No public health and safety concerns have been identified in this assessment.
On the basis of the available evidence, including detailed studies provided by the Applicant, food derived from soybean line FG72 is considered as safe and wholesome as food derived from other commercial soybean cultivars." (FSANZ, 2011).

Information is provided in this application to support claims that the 2mEPSPS and HPPD W336 proteins expressed by the GHB811 event share no characteristics consistent with toxins or allergens, and potential for mammalian toxicity has been addressed previously for these proteins. Compositional and nutritional analyses demonstrate that food derived from cotton containing event GHB811 is as safe and nutritious as food derived from conventional cotton varieties.

The USA have submitted for FDA approval of GHB811 cotton in April 2017. Submissions for food and feed approval in other importing countries will follow. Information on the global submission status can be provided to update FSANZ during the application process.

1.4 Regulatory impact information

Costs and benefits, and impacts on trade

Varying FSANZ Standard 1.5.2 to include commercial cotton varieties containing event GHB811 is unlikely to have a detrimental impact on the Australian cotton or food industries. At present, the US and Brazilian cotton may be the source for many imported food products

on the domestic market. Once cotton varieties containing the GHB811 event are launched for commercial production in the US and Brazil, as well as potentially in other parts of the world, food products derived from cotton containing this event may enter the domestic food supply.

If the cotton event GHB811 is not incorporated into the FSANZ Standards, this could have wide ranging impacts on the price of food products containing ingredients derived from cotton. These would arise from the need to source other cotton varieties that do not contain the GHB811 event. These products may attract a premium price that must be met by the manufacturer, with those costs eventually passed on to the consumer. This would be compounded by the costs of segregating GHB811 cotton products from other cotton products, where trading partners are willing to comply with this requirement. Other factors to consider include disruptions to the food supply, and the significant costs of recalling food products if the GHB811 event were to be distributed in the local food supply.

Varying the FSANZ Standards to include GHB811 will contribute to maintaining stable food prices, consumer choice in the marketplace, and decreased production costs for transgenic cotton varieties in the longer term.

The potential trade implications of not including cotton event GHB811 cotton in the FSANZ Standards are significant. Segregating GHB811 cotton products from other cotton products has compliance and identification requirements that are difficult and costly to meet. The US and Brazil are major trading partners of Australia, and approved transgenic crops are considered to be substantially equivalent to conventional crops. Therefore, in the US and Brazil, it is unlikely that segregation or labelling transgenic cotton crops or their products will occur. Products containing event GHB811 imported into Australia from the US or Brazil, or other trading partners with similar treatments of transgenic crops, may need to be removed from sale. This could expose Australia to disputes with trading partners at the World Trade Organisation.

1.5 Information to support the application

All of the relevant information to support the application is supplied within this summary and the associated electronic dossier that has been supplied to FSANZ. The relevant studies are listed in the “List of Appended Electronic Documents” above, and suitable literature references are provided in a reference list at the end of this document. To navigate the electronic dossier a direction to which “Node” of the dossier the document may be found under is supplied.

1.6 Assessment Procedure

We consider that the appropriate assessment for this application is the General Procedure since the 2mEPSPS and HPPD W336 proteins have been evaluated by FSANZ previously.

1.7 Confidential Commercial Information

Information in the Bayer CropScience reports provided in Nodes A.3 (c), (i) (CCI) (Document M-572036-01), A.3 (c), (iii) (CCI) (Documents M-533573-01, M-234184-01), A.3, (c), (v) (CCI) (Document nos. M-581222-01 and M-575144-01), A.3 (e), (i) (CCI) (Document M-548778-01) contain confidential commercial information. A formal request for this information to be treated as such has been submitted to FSANZ.

1.8 Other Confidential Information

Bayer CropScience requests that versions of supporting documents submitted with this application that have privacy information removed only are provided to any interested members of the public upon completion of the FSANZ review. The documents included in this request include: M-543362-01, M-388224-02, M-572036-01, M-533573-01, M-581222-01, M-575144-01, M-548778-01, M-547925-02, M-574232-01, M-568145-01, M-497839-01, M-576569-01, M-497842-01, M-574125-01, M-575144-01, M-445681-04, M-445678-04, M-402872-01, M-278169-01, M-234184-01 and M-566678-01.

1.9 Exclusive Capturable Commercial Benefit (ECCB)

The application is expected to confer an ECCB upon Bayer CropScience since it will contribute to facilitating commercial activities with GHB811 cotton firstly in the USA and Brazil, followed possibly by other cotton producing countries.

1.10 International and Other Standards

The Bayer CropScience reports and studies included in the information supporting this application have been conducted according to international standards. In the safety assessment of biotechnology products, Bayer CropScience refers primarily to the *Codex Alimentarius* Commission weight-of-evidence approach (CAC, 2009), and the relevant Codex Standard is:

Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants. CAC/GL 45-2003. Adopted in 2003, Annexes II and III adopted in 2008. (CAC, 2009).

Other guidelines and recommendations are also considered including those of the World Health Organisation (WHO), the United Nations Food and Agriculture Organisation (FAO), the United States Food and Drug Administration (US-FDA), the United States Environment Protection Agency (US-EPA), and the European Food Safety Agency (EFSA) (see CAC, 2009 above; EFSA, 2011; FAO/WHO, 2001; US-FDA, 2012).

1.11 Statutory Declaration

Included in the application to FSANZ, which is appended as an electronic document at Node 1K within the DVD which contains the submission.

1.12 Checklist for Standards Related to New Foods

APPLICATION REQUIREMENT CHECKLIST	SECTION IN THIS APPLICATION	PAGE NUMBER
General Requirements (Application Handbook section 3.1)		
Form of application		
Applicant details	1.1	14
Purpose of the application	1.2	14
Justification of the application	1.3	15
Regulatory impact information	1.4	16
Information to support the application	1.5 Parts A, B, C and D	17
Assessment procedure	1.6	17
Confidential Commercial Information	1.7	18
Other Confidential Information	1.8	18
Exclusive Capturable Commercial Benefit	1.9	18
International and other standards	1.10	18
Statutory Declaration	1.11	See Node K
Checklist for Standards Related to New Foods	1.12	19
Foods Produced Using Gene Technology (Application Handbook section 3.5.1)		
Nature and identity of the genetically modified food	A.1	21 – 22
History of use of host and donor organisms	A.2	22 – 33
The nature of the genetic modification	A.3	33 – 93
Characterisation and safety assessment of new substances	B.1	93 – 117
New proteins	B.2	117 - 124
Other (non-protein) new substances	B.3	124 - 125
Novel herbicide metabolites in GM herbicide-tolerant plants	B.4	125 - 127
Compositional analyses of the food produced using gene technology	B.5	127 – 141
Information related to the nutritional impact of	C	141

the food produced using gene technology		
Other information	D	141

Part A Technical Information on the Food Produced Using Gene Technology

A.1 Nature and Identity of the Genetically Modified Food

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

The GM organism is cultivated cotton (*Gossypium hirsutum*) transformed with the GHB811 event. Seeds of cotton variety Coker 312 were germinated on Murashige & Skoog (MS) medium. Hypocotyl segments were dissected from the cotton seedlings and were transformed with the transformation vector pTSIH09 using a cotton hypocotyl *Agrobacterium tumefaciens* (*A. tumefaciens*) transformation method (M-543362-01; Dossier Node A.3 (a)).

The GHB811 event introduced two genes to the *G. hirsutum* genome:

(i) The *2mepsps* gene encodes for the 2mEPSPS protein. The modified EPSPS (2mEPSPS) enzyme has a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of glyphosate herbicides (Lebrun, M. *et al*., 1997; M-216526-01; Dossier Node A.1 (a)). The gene is derived from *Zea mays* L. (corn). Therefore, the plants bearing this gene become tolerant to glyphosate herbicides (Lebrun, M. *et al*., 1997; M-216526-01; Dossier Node A.1 (a)). The *2mepsps* gene has been used to confer glyphosate tolerant properties to crops including maize, cotton, canola and soybean (Herouet-Guicheney *et al*., 2009). FSANZ has previously assessed the 2mEPSPS protein, as expressed by the *2mepsps* gene, in the Bayer CropScience applications for GlyTol cotton (A614) and FG72 soy bean (A1051).

(ii) The *hppdPf W336* gene encodes for the HPPD W336 protein. The *hppdPf W336* coding sequence was developed by introducing a single point mutation to the wild type *hppd* gene derived from *Pseudomonas fluorescens*. Expression of the HPPD W336 protein confers tolerance to HPPD inhibitors, such as isoxaflutole herbicides. The *hppdPf W336* gene has been used to confer HPPD inhibitor tolerant properties to soy bean in the past. FSANZ has previously assessed the HPPD W336 protein, as expressed by the *hppdPf W336* gene, in the Bayer CropScience application for FG72 soy bean (A1051).

The coding sequence of 5-enol pyruvylshikimate-3-phosphate synthase (*epsps*) gene was isolated from maize (*Zea Mays* L.) Two amino acids were substituted (threonine by isoleucine at position 102 and proline by serine at position 106) (Lebrun, M. *et al*., 1997; M-216526-01; Dossier Node A.1 (a)). These modifications confer to the protein a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of the herbicide. Therefore, the plants expressing this modified protein become tolerant to glyphosate herbicides (Lebrun, M. *et al*., 1997; M-216526-01; Dossier Node A.1 (a)). The modified protein is designated as 2mEPSPS.

The coding sequence of the 4-hydroxyphenylpyruvate dioxygenase (HPPD) protein was isolated from the *Pseudomonas fluorescens* strain A32. One amino acid was substituted (glycine at position 336 with tryptophan) to improve the tolerance against HPPD inhibitors such as the herbicide isoxaflutole. The modified protein is designated as HPPD W336 (Boudec, P.; et al.; 2001; M-229534-01; Dossier Node A.1 (a)).

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

The transformation event is named “GHB811”, and cotton transformed with this event will be referred to as GHB811 cotton. The OECD Unique identifier of GHB811 cotton is BCS-GH811-4.

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

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

A.2 History and Use of the Host and Donor Organisms

The common and scientific names of the host and donor organisms must be stated.

The taxonomic classifications of the organisms from which the genetic elements of GHB811 cotton are derived, are presented below in Table 1.

Cotton

The word ‘cotton’ is used in this document to refer primarily to *G. hirsutum*, however, generally ‘cotton’ refers to four species in the genus *Gossypium* (Malvaceae) - *G. hirsutum* L., *G. barbadense* L., *G. arboreum* L. and *G. herbaceum* L. - that were domesticated independently as source of textile fibre (Brubaker et al. 1999). Today, *G hirsutum* and *G. barbadense* are the major cultivated cotton species, with *G. hirsutum* accounting for 90% of world production (Jenkins 2003; OGTR 2016). The host species, *G. hirsutum* belongs to the extensive genus *Gossypium*, which contains 43 species. *Gossypium hirsutum* L. is commonly known as upland cotton, American cotton or Mexican cotton – indicating its origin in the West (OGTR, 2008). No *Gossypium* species are recognised as problematic weeds in Australia, either agriculturally or environmentally (Lazarides et al. 1997; Tothill et al. 1982). Cotton has no relatives that are problematic weeds (Keeler et al. 1996), although locally *G. sturtianum* can be weedy (Lazarides et al. 1997; OGTR, 2016).

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

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

Pseudomonas fluorescens Migula 1895 (type strain ATCC 13525; taxonomy ID: 136843; Skerman, 1980)*, Pseudomonas putida* and *Pseudomonas chlororaphis* are closely related to

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

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

The *hppdPf W336* gene coding sequence is under the control of the sequence including the promoter region of the Cassava Vein Mosaic Virus (CVMV; Verdaguer *et al.,* 1996), and by the optimized transit peptide as described by Lebrun *et al.* (1996), and terminated by the 3′ untranslated region of the histone H4 gene of *Arabidopsis thaliana* (Chabouté *et al*., 1987).

The plant species *Arabidopsis thaliana,* common name mouse-ear cress, is member of the family Brassicaceae. *Zea mays*, more commonly know as maize or corn, is a member of the Poaceae family. *Helianthus annuus*, or sunflower, is a member of the family Asteraceae. These plant species are not considered to cause disease in humans, plants or animals.

Cassava vein mosaic virus (CVMV) is a plant pathogenic virus of the family Caulimoviridae. Caulimoviruses are spherical or bacilliform plant viruses containing a circular, double-stranded (ds) DNA genome of 7.1 to 8.2 kb. The CVMV is a pararetrovirus that is only known to infect cassava in Brazil. It is not known to cause disease in animals or humans.

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

Table 1 Taxonomy of the donor organisms from which the genetic elements of GHB811 cotton are derived

GENETIC ELEMENT	DONOR ORGANISM TAXONOMY							
	Kingdom	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Scientific Name**	**Common Name**
Plant Genome								
Genomic DNA	Plantae	Streptophyta	Magnoliophyta	Malvales	Malvaceae	*Gossypium*	*Gossypium hirsutum* L. (2n=52)	Cultivated cotton
Gene Construct								
ThistonAt	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress
hppdPfW336-1Pa	Bacteria	Proteobacteria	*Gammaproteobacteria*	Pseudomonadales	Pseudomonadaceae	Pseudomonas	*Pseudomonas fluorescens*	Strain ATCC 13525; taxonomy ID 136843[A]
TPotpY-1Pa	Plantae	Anthophyta	Liliopsida	Poales	Poaceae	Zea	*Zea mays*	corn
	Plantae	Magnoliophyta	Magnoliopsida	Asterales	Asteraceae	Helianthus	*Helianthus annuus*	sunflower
Pcsvmv	Viruses	-	-	-	Caulimoviridae	"Cassava vein mosaic-like viruses"	Cassava vein mosaic virus	CsVMV
lox	Viruses	-	-	Caudovirales	Myoviridae	"Puna like virus"	Bacteriophage P1	
Ph4a748	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress
intron1 h3At	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress
TPotpC	Plantae	Anthophyta	Liliopsida	Poales	Poaceae	Zea	*Zea mays*	corn
	Plantae	Tracheophyta	Magnoliopsida	Asterales	Asteraceae	Helianthus	*Helianthus annuus*	sunflower
2mepsps	Plantae	Anthophyta	Liliopsida	Poales	Poaceae	Zea	*Zea mays*	corn
ThistonAt	Plantae	Magnoliophyta	Magnoliopsida	Brassicales	Brassicaceae	Arabidopsis	*Arabidopsis thaliana*	mouse-ear cress
lox	Viruses	-	-	Caudovirales	Myoviridae	"Puna like virus"	Bacteriophage P1	

[A] Skerman, 1980. Approved lists of bacterial names. *Int. J. Syst. Bacteriol.* 30:225-420. Available at http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=bacname (accessed on January 4, 2017).

Where information relating to an organism has been included in previous safety assessments prepared by FSANZ, it is not necessary to provide any further information. Where an organism has not been considered previously by FSANZ, the following information must be provided. A partial package of data has been provided to FSANZ GHB811 cotton as the proteins 2mEPSPS and HPPD W336 have been considered for food safety approval by FSANZ previously in the submissions for GHB614 cotton and FG72 soy bean, respectively. Full protein safety data sets were provided for 2mEPSPS and HPPD W336 within the GHB614 (A614) and FG72 (A1051) dossiers, respectively. Data requirements as per the March 2016 Application Handbook are addressed here.

(a) For the donor organism(s) from which the genetic elements are derived:

(i) Any known pathogenicity, toxicity or allergenicity of relevance to the food;

Cotton (*Gossypium hirsutum)*

The host organism, cultivated cotton (*G. hirsutum*), is an established agricultural field crop that is grown as a source of fibre, food and feed. Cotton has been cultivated for millennia in many parts of the world, primarily for lint which is spun into yarn. Cotton is not known to be capable of causing disease or other ill health in people, plants or animals, except in cases of excessive consumption of cotton tissues, particularly the seeds. Plants commonly produce toxins and allergens that serve as a natural defence against pests and pathogens. Seeds and other cotton plant tissues contain toxic and anti-nutritional secondary defence chemicals including gossypol, cyclopropenoid fatty acids (dihydrosterculic, sterculic and malvalic acids) and phytic acid. Despite the natural presence of these compounds, cotton has a long history of safe use.

Cottonseed oil is the only product of cotton that represents a major component of human food, and it is an important vegetable oil source. Cottonseed oil intended for human consumption is highly processed to reduce its toxicological properties (Jones and King, 1993; Bailey *et al*., 1966; Harris *et al*., 1964; Levi *et al*., 1967; Phelps *et al*., 1965). Highly processed edible oils contain virtually no proteins, indicating the minimal allergenic potential of cottonseed oil used for human consumption. This hypothesis is supported by clinical trials demonstrating a correlation between the absence of water-soluble allergens in cottonseed oil with a lack of clinical allergy observations after its consumption (Hefle and Taylor, 1999; Bernton *et al*., 1940; 1949). The absence of proteins in processed cottonseed oil also suggests that cottonseed oil from any transgenic variety should be as safe for human consumption as cottonseed oil from conventional cotton varieties.

2mepsps gene and 2mEPSPS protein

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

The maize source organism is a safe crop plant widely used for food and feed with little pathogenic, toxic, or allergenic effects on humans and animals. The 2*mepsps* gene is composed of the same essential nucleic acids found in any food or feed DNA, which is commonly consumed as part of human or animal diets. Decades of research have indicated that dietary DNA poses no direct toxicity to human health. The EPSPS proteins are ubiquitous in nature, widely expressed in food and feed crops (e.g. soybean, tomato, maize). No health-related adverse effects have been associated with these proteins. Since the 2mEPSPS protein is derived from maize and has only two amino acid modifications, the safety profile of the novel protein is expected to remain unchanged relative to its wild-type counterpart. The 2mEPSPS protein is highly homologous to, and shares similar molecular weight and functionalities with other shikimate synthase proteins which have been

demonstrated to be non-toxic and non-allergenic over the years through consumption. Its identity with the wt EPSPS enzyme is greater than 99.5%.

hppdPf W336 gene

Pathogenicity to humans

The gene *hppdPf W336* in event GHB811 is derived from *P. fluorescens. P. fluorescens* can be an opportunistic pathogen in immunocompromised patients (McKellar, 1982). Some cases of septicemia have been reported due to *P. fluorescens* contamination of transfused blood and blood products, given its ability to grow at 5°C (Gibb *et al*., 1995, Puckett *et al*., 1992). Some *P. fluorescens* strains were also reported to create biofilms on compounded sterile products like catheters and have led to rare infections in immunocompromised populations (Gershman *et al.,* 2008). However, the general virulence of *P. fluorescens* is low due to its inability to multiply rapidly at body temperature and having to compete with defense mechanisms of the host (Liu, 1964).

Pathogenicity to animals

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

Pathogenicity to plants

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

Allergenicity

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

Table 2 Gazetted FSANZ Standards for events encoding for the expression of HPPD W336 and 2mEPSPS (and EPSPS family) proteins

CROP	EVENTS/LINES EVALUATED	APP
Canola	CP4 EPSPS from *A. tumefaciens* + GOX from *Ochrobactrum anthropi* (A363) CP4 EPSPS (A1071)	A363, A1071
Cotton	CP4 EPSPS from *A. tumefaciens* (A355 and A553) or 2mEPSPS from *Zea mays* (A614)	A355, A553, A614
Lucerne	CP4 EPSPS from *A. tumefaciens*	A575
Maize	CP4 EPSPS from *A. tumefaciens* or 2mEPSPS from *Zea mays*	A362, A416, A548, A1066, A1097, A1112
Soybean	CP4 EPSPS from *A. tumefaciens* (A338, A592, A1049) 2mEPSPS (A1073) HPPD W336 and 2mEPSPS from *A. tumefaciens* (A1051)	A338, A592, A1049, A1051, A1073
Sugarbeet	CP4 EPSPS from *A. tumefaciens*	A378, A525
Wheat	CP4 EPSPS from *A. tumefaciens*	A524

Regulatory sequences
The promoter and terminator sequences used in GHB811 are derived from common plants or plant pathogens. These genetic elements constitute a minute component of their respective genomes, no genes that may be implicated in human disease, allergies or toxic effects have been transferred. Many of these organisms from which these elements are derived are model species in plant science with a history of safe use. These elements are described in Table 3, Section A.3(c)(i).

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

*Cotton (*Gossypium hirsutum)
The host organism, cultivated cotton (*G. hirsutum*), is an established agricultural field crop that is grown as a source of fibre, food and feed and has a long history of safe use. It has been cultivated for millennia in many parts of the world primarily for its lint which is spun into yarn. The cultivation of cotton and the manufacture of cotton fabrics developed independently in both the Eastern and Western Hemispheres. One of the oldest records of cotton textiles, dating back about 5,000 years, was found in the Indus River Valley in what is now Pakistan. Excavations in Peru and Mexico have uncovered cotton cloth identified as being 4,500 to 7,000 years old. Cotton fabrics have also been found in the remains of some of the ancient civilisations of Egypt and in the ruins of Indian pueblos of the southwestern United States, dating back at least 2,000 years.

The presence of gossypol and cyclopropenoid fatty acids limits the use of cottonseed as a protein supplement in animal feed, except for cattle that are able to detoxify these chemicals in the rumen. Feeding raw oilseeds, and especially cottonseed, is a method used to provide concentrated nutrients, particularly energy, to high producing dairy cows. In Australia, cotton seed and meal are used as high protein sources in the dairy industry surrounding cotton growing districts. For other animals, the use of cottonseed as stockfeed is limited to a relatively small proportion of the diet and it must be introduced gradually to avoid potential toxic effects. Inactivation or removal of gossypol and cyclopropenoid fatty acids during processing enables the use of some cotton seed meal for catfish, poultry and swine (OGTR, 2002; 2008).

Although cotton seed meal is not used for human consumption in Australia or New Zealand, it has been approved for use in human food in some countries where it is derived from gossypol-free varieties of cotton or after processing to remove the gossypol. Human consumption of cotton seed meal is primarily in Central American countries and India where it is used as a low cost, high quality protein ingredient (Franck 1989; Ensminger et al. 1990).

Edible grades of cottonseed linter fibre, containing more than 99% total fibre, are also used for human consumption. The product is a pure white, flavourless, odourless flour that is chemically stable and will not react with other ingredients. It is used in many food products including baked goods, dressings, snacks and processed meats. Linter fibre is also used to improve the viscosity of dressings and is commonly used to bind solids in pharmaceutical preparations such as tablets. Since linter fibre is white and is one of the purest forms of fibre, it has obvious advantages over other sources of edible fibre such as wood pulp. Linter pulp is also used in the production of such diverse items as sausage casings, cellulose lacquers for use on furniture, metal, and in fingernail polishes, decorative laminates, industrial and automotive filters, battery separators, and printed electrical circuit boards for use in the computer and electronics industry (NCPA, 1999).

Cottonseed oil is the only product of cotton that represents a major component of human food and it is an important vegetable oil source. Up until World War II, cottonseed oil was the

major vegetable oil produced in the US. It now ranks third in volume behind soybean oil and corn oil (NCPA, 1999), and is considered to be a premium quality oil due to its balance in unsaturated fatty acids and high tocopherol (Vitamin E) content and stability when used as a frying oil. Cottonseed oil is also used in shortening and salad dressing, and snack foods including crackers and cookies (NCCA, 1999).

Cottonseed oil intended for human consumption is highly processed to reduce its toxicological properties (Jones and King, 1965; Bailey *et al*., 1966; Harris *et al*., 1964; Levi *et al*., 1967; Phelps *et al*., 1965). Highly processed edible oils contain virtually no proteins, indicating a minimal allergenic potential of cottonseed oil for human consumption. Therefore, cottonseed oil is not considered to contain toxins or anti-nutritional components of concern for human health.

Further information on *Gossypium hirsutum* and its uses can be found in the OECD consensus document on the biology of cotton (OECD, 2008).

2mepsps gene

The coding sequence for the 2mEPSPS protein was isolated from *Zea mays* (maize). Maize is one of the few major crops that are indigenous to the Western Hemisphere and it is grown in nearly all areas of the world (Hallauer *et al*., 1988). There are many food/feed and industrial products that contain ingredients derived from maize. It is an important crop in human and animal nutrition because of its high levels of starch, protein, oil and other nutritionally valuable components. Consequently, maize has a very long history of safe use.
FSANZ has previously evaluated other transgenic crops that express EPSPS proteins (Table 2) and determined the protein indicates no potential for allergenicity or toxicity in humans.

hppdPf W336 gene

The *hppdPf W336* gene was isolated from *Pseudomonas fluorescens* strain A32. *Pseudomonas fluorescens* are ubiquitous bacterium in the natural environment and are frequently present in water, soil and plant rhizosphere (Bossis *et al*., 2000). The bacterium can be isolated from water, animals, human clinical specimens, the hospital environment, and spoiled foodstuffs such as fish and meat. The survival of *P. fluorescens* is affected by number of biotic and abiotic factors such as soil density, temperature, pH and humidity (OECD, 1997).
The natural properties of *P. fluorescens* are exploited in agriculture for plant growth-promotion (Fliessbach *et al*., 2009; OECD, 1997) and pest control. As a growth control agent, the bacterium can enhance plant growth through production of siderophores, which efficiently complex environmental iron rendering it unavailable to other organisms of the soil microflora. As a biopesticide, *P. fluorescens* is able to prevent the growth of frost-forming bacteria on leaves and blossoms of crops and fruits (Compant *et al*., 2005; Raaijmakers *et al*., 2006; US-EPA. 2008a), and prevent damping off diseases caused by fungi (Haas and Defago, 2005; Thrane *et al*., 2001; Voisard *et al*., 1989) and nematodes (Hamid *et al*., 2003) when used as a seed treatment. Naturally occurring strains of *P. fluorescens* have been registered commercially for the control of frost injury and fire blight on pear (Wilson and Lindow, 1993). Since 1992, four products containing *P. fluorescens* strains as active ingredients were approved by US-EPA (US-EPA, 2008b). The US-EPA recognized that this bacterial active ingredient is not expected to cause any adverse health effects in humans, based on various studies that found no evidence that these bacteria are harmful to mammals (US-EPA. 2008a). The US-EPA also established a tolerance exemption for residues of *P. fluorescens* in or on the raw agricultural commodity mushrooms (US-EPA, 1994). The pesticidal activity of *P. fluorescens* is attributed to three mechanisms: competition for an ecological niche or a substrate, production of inhibitory chemicals and induction of systemic resistance in host plants to a broad spectrum of pathogens (Compant *et al*., 2005; Haas and Defago, 2005).

In other applications, strains of *P. fluorescens* have been genetically modified to encapsulate crystal δ-endotoxins (Cry proteins) from the bacterium *Bacillus thuringiensis* (Bt) (Downing *et al*., 2000; Peng *et al*., 2003). The Cry proteins encapsulated by *P. fluorescens* showed high insecticidal activity and retained this activity for two to three times longer than Bt formulations (Peng *et al*., 2003). In pharmaceutical uses, *P. fluorescens* produces the antibiotic pseudomonic acid (also called mupirocin), which is used to prevent *Staphylococcus aureus* infections (Hothersall *et al*., 2007; Tacconelli *et al*., 2003). Further, in addition to the metabolic diversity of *P. fluorescens*, it may be used in bioremediation applications. The bacterium is able to degrade a wide variety of compounds, including 3-chlorobenzoic acid, naphthalene, phenathrene, fluorene and fluoranthene, chlorinated aliphatic hydrocarbons, styrene, pure hydrocarbons and crude oil (OECD, 1997).
In summary, the source of the *hppdPf W336* gene is ubiquitous in the environment, including soil, water and food. It has many beneficial uses in agriculture, human health and bioremediation. Despite this widespread presence, it is not described as allergenic, toxic or pathogenic to healthy humans and animals.

Regulatory sequences
The promoter and terminator sequences used in GHB811 are derived from common plants or plant pathogens. These genetic elements constitute a minute component of their respective genomes, no genes that may be implicated in human disease, allergies or toxic effects have been transferred. Many of the organisms from which these elements are derived are model species in plant science with a history of safe use.

(b) For the host organism into which the genes were transferred:

(i) Its history of safe use for food

Cotton is primarily grown as a fibre crop. It is harvested as “seed cotton” which is then “ginned” to separate the seed and lint. The long “lint” fibres are further processed by spinning to produce yard that is knitted or woven into fabrics. Cotton fabrics, used in clothing, upholstery, towels and other household products, are made from cotton lint (OGTR, 2008).

The ginned *G. hirsutum* seed is covered in short, fuzzy fibres, known as “linters”. These must be removed before the seed can be used for planting or crushed for oil. The linters are produced as first-cut or second-cut linters. The first-cut linters have a longer fibre length and are used in the production of mattresses, furniture upholstery and mops. The second-cut linters have a much shorter fibre length 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. In the chemical industry the second-cut linters are used with other compounds to produce cellulose derivatives such as cellulose acetate, nitrocellulose and a wide range of other compounds (Gregory *et al*., 1999; OGTR, 2008). G. hirsutum ginned seed comprises 17% crude oil, 45% meal, 10% linters and 28% hulls (Smith, 1995; OGTR, 2008).

De-linted cotton seed (i.e. seed with no lint or linters) is processed into oil, meal and hulls (Cherry and Leffler, 1984; OGTR, 2008). The processing of cotton seed oil involves a series of steps including heating, addition of sodium hydroxide, bleaching with clay, filtering and treating with steam under vacuum (OECD, 2004; OGTR, 2008). Cotton seed oil has been in common use since the middle of the nineteenth century and achieved GRAS (Generally Recognised as Safe) status under the United States Federal Food Drug and Cosmetic Act because of its common use prior to 1958 (ANZFA, 2002; OGTR, 2008). It is used in a variety of products including edible vegetable oils and margarine, soap and plastics (Franck, 1987; OGTR, 2008).

Cotton seed meal is the product remaining once the oil has been removed by crushing and can contain up to 41% protein (Smith, 1995; OGTR, 2008). Cotton seed, or meal, flour or hulls derived from it, is used in food products and for animal feed, but this is limited by the presence of natural toxicants in the seeds (gossypol and cyclopropenoid fatty acids). Although cotton seed meal is not used for human consumption in Australia or New Zealand, it has been approved for use in human food in the USA and other countries, when derived from gossypol-free varieties of cotton or after processing to remove the gossypol. The FAO and WHO permit up to 0.6 µg/mg (600 ppm) free gossypol in edible cotton seed products, whereas the FDA has a lower limit of 450 ppm (Lusas and Jividen, 1987; OGTR, 2008). Human consumption of cotton seed meal is reported mainly in central American countries and India where it is used as a low cost, high quality protein ingredient (Frank, 1987; OGTR, 2008).

(ii) The part of the organism typically used as food

Cotton seeds are processed into three major products, oil, meal and linters. Two of these products are used as food here in Australia and New Zealand – the oil and linters. 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. Subsequent alkali treatment and deodorisation steps are likely to remove any last detectable traces of protein in the refined oil. Deodorisation also greatly reduces the cyclopropenoid fatty acid content.
Cotton linters are short fibres that remain after the long fibres have been removed at the ginning process for textile manufacture. Linters consist of nearly pure (> 99%) cellulose and are used in both chemical and high fibre dietary products. Food uses include casings for processed meats, and as a viscosity enhancer (thickener) in ice cream, salad dressings and toothpaste. (FSANZ, A1094).

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

See the information under Section A.2 (b)(ii) above.

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

Most harvested cotton is formed into modules, covered and stored in the field until it can be ginned. At the gin, cottonseed is cleaned, delinted and hulled, and crushed to extract oil from the meal and hulls. This process is described below and illustrated in .

<u>Cleaning and delinting</u>
In the first steps of processing, cottonseeds are cleaned to remove any plant debris or sand picked up in the fields or during handling. Various combinations of revolving screens, shaker screens, and pneumatic equipment or vacuum pipes are used for this purpose. The cottonseeds are then conveyed to delinting machines to remove the short fibres known as linters. These machines consist of a series of circular saws that project through a set of steel ribs on a horizontal revolving shaft. As cottonseeds fall on the closely spaced ribs, the saws revolve and cut off the linters. The linters are removed from the saws by brushes or air and collected and pressed into bales. Cottonseed may be run through delinting machines once, in which case the resulting linters are known as "mill run." Most mills run the cottonseed through twice and produce "first-cut" and "second-cut" linters. First cuts consist of the longer, more resilient fibres, while second cuts are made up of short fibres or fuzz. Some mills employ an abrasive delinting process that involves a rubbing action to remove the linters.

Some high-capacity gins can produce as many as sixty 500-pound bales of lint per hour (NCPA, 1999).

Hulling

After the linters are removed, the cottonseed proceeds to hulling. The protective hull surrounding cottonseed kernels is relatively tough. To loosen the hulls from the kernels, the cottonseed is passed through a hulling machine that consists of a series of knives. The cottonseed is then passed through a series of shakers to separate the hulls from the kernels. Good separation is necessary for the efficient production of quality oil and meal. After separation, the hulls are ready for marketing, and the kernels, or meats, are ready for oil extraction (NCPA, 1999).

Oil extraction

Oil is removed from the seed by mechanical screw presses, solvent extraction, or a combination of both. In both processes, meats pass from the huller through a series of heavy cast iron rollers arranged one above the other which revolve at high speed. The rollers reduce the meats to thin flakes.

For screw pressing, the flakes move from the rollers to a cooker or conditioner that reduces moisture content to a low level. From the cooker they flow directly into the press which has a screw or worm revolving inside a horizontal steel barrel. Flakes enter one end of the barrel where they are exposed to very high pressure (as much as 10 to 12 tons per square inch) created by the revolving screw. Oil is forced from the meats and flows through small openings in the barrel of the press into a receiving chamber below. From there it is piped through a filter which removes any remaining fine particles of meats and into storage tanks. The extracted flakes are formed into a ribbon of cake that passes through the barrel. After cooling, the cake is ground into meal or further processed into pellets. The meal produced by the screw press process has about 3-4% residual oil remaining.

The newest technology for oil extraction uses an expander to prepare the meats for oil extraction. After the preparation step, expanded collets are exposed to an organic solvent that extracts the oil. The mixture of oil and solvent is then passed through a series of evaporators and stills to separate the solvent, which is recovered and re-used. Solvent is also recovered from extracted meats by use of a desolventiser. These meats may then be toasted and ground into meal or processed into pellets. Meal produced by solvent processes usually contains 1-3% residual fat. Large, efficient processing plants employ this expander-solvent method to produce higher quality products at a lower cost compared to the older technology.

Oil Refining

The crude cottonseed oil obtained from either of the oil extraction processes is further processed to become food grade oil. The refinement process, detailed below in Figure 2 (source NCPA, 1999), consists of:

- Treatment with alkaline water solutions (caustic soda or sodium hydroxide) to remove acidic compounds such as free fatty acids. The alkali combines with a portion of the oil to form what is known as soapstock. The soapstock, together with impurities that may be present, is then separated from the oil by means of a high-speed centrifuge.
- Bleaching: removal of undesirable colour pigments is achieved by adding an absorbent material (bleaching clay). The adsorbent is activated by heating and the oil is then filtered to remove the adsorbent along with the pigments.
- Winterisation: refined cottonseed cooking oil turns cloudy at 4.4 - 10°C and becomes solid at a little below 0°C. To control this, the oil is reduced in temperature to 3.3 – 4.4°C. A portion of the oil, known as stearine, crystallises or solidifies and is separated from the liquid oil by a filter process. The stearine is used in shortening and margarine products.

- Deodorization: all cottonseed oil, whether winterised or not, is finally deodorised by exposing it to steam under a partial vacuum to remove unwanted flavours and to further purify it before use. Cyclopropenoid fatty acids are largely deactivated or removed from the oil by hydrogenation or during deodorization at 230-235°C.

The by-products of oil refining are used for off-grade oil or soapstock (NCPA, 2000).

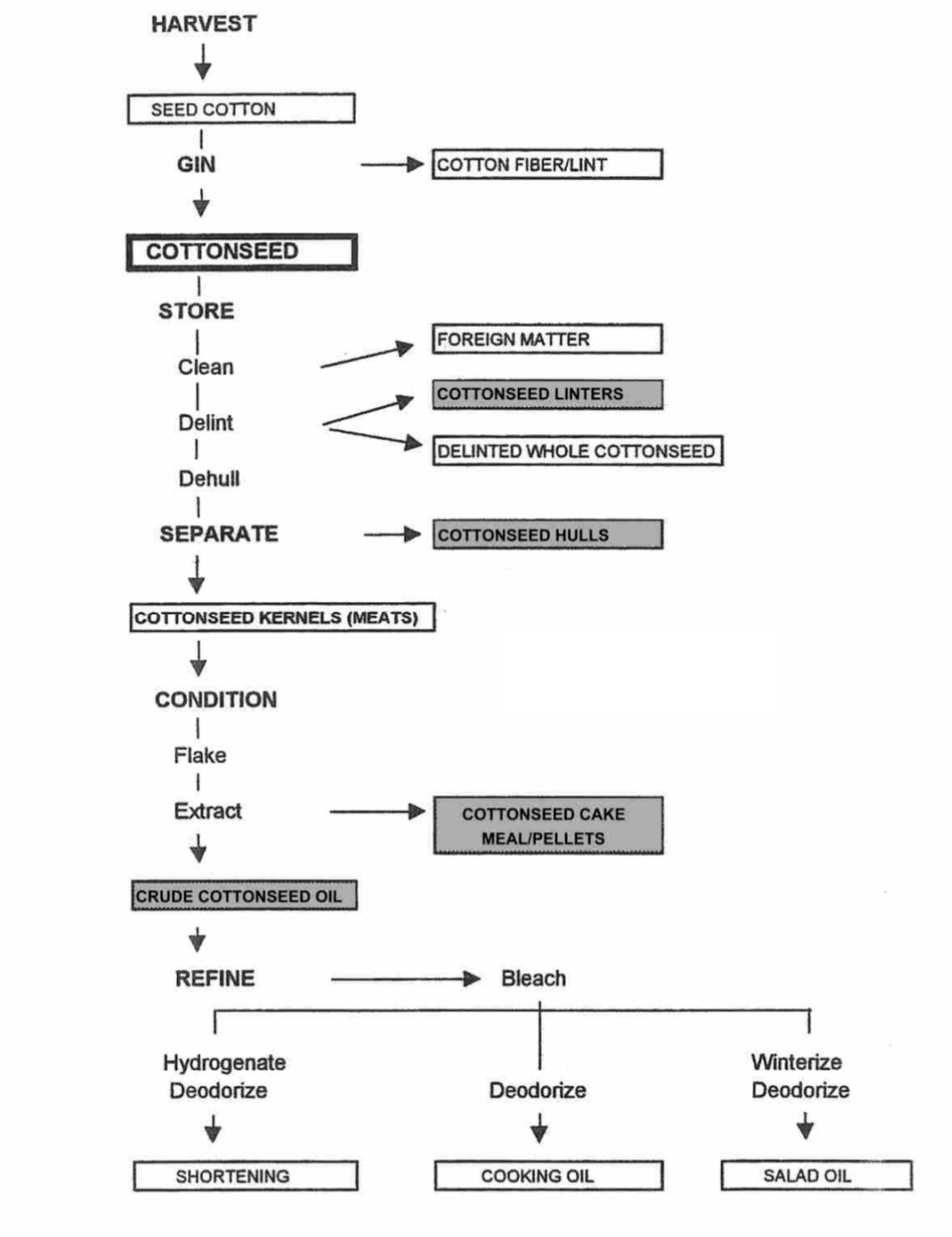


Figure 1 Diagram of the processing steps involved in food production from cottonseed. Source: NCPA, 2000

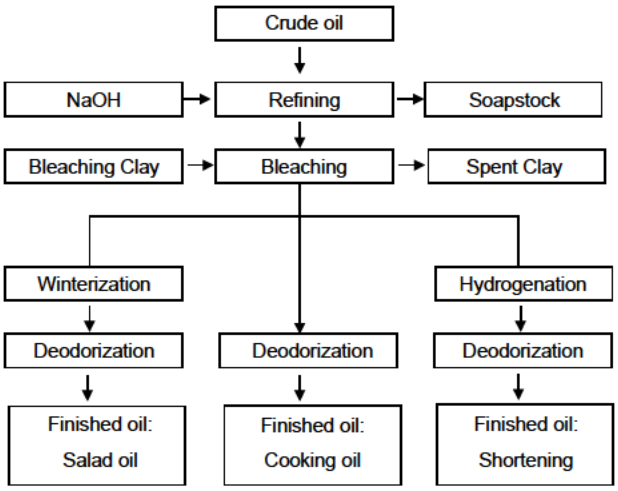


Figure 2 Diagram of the processing steps involved in producing finished oil products from crude oil derived from cottonseed. Source: NCPA, 2000

A.3 The Nature of the Genetic Modification

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

Seeds of cotton variety Coker 312 were germinated on Murashige & Skoog (MS) medium. Hypocotyl segments were dissected from the cotton seedlings and were transformed with the transformation vector pTSIH09 using a cotton hypocotyl *Agrobacterium tumefaciens* (*A. tumefaciens*) transformation method (M-543362-01; Dossier Node A.3 (a)).

See Table 3 below for a description of the vector M-388224-02; Node A.3 (b), (i) and (ii)).

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

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

The genetic components comprising GHB811 *G. hirsutum* are detailed in Table 3 below, and in M-388224-02; Node A.3 (b), (i) and (ii)). The sequence between nucleotide positions 1 – 6856 bp represents the intended transgenic locus, with the remaining sequence comprising the vector backbone. These components are shown in Figure 3 in Section A.3(b)(ii) below.

Table 3 Genetic elements comprising the pTSIH09 vector used in GHB811 *G. hirsutum*

GENETIC ELEMENT	NT POSITION	SIZE (KB)	ORIENTATION	DESCRIPTION AND FUNCTION	REFERENCE
RB	1 – 25			Right border region of the T-DNA of *Agrobacterium tumefaciens.*	Zambryski, 1988
	26 - 82			Synthetic polylinker derived sequences	
ThistonAt	83 - 749		Counter clockwise	Sequence including the 3' untranslated region of the histone H4 gene of *Arabidopsis thaliana*.	Chabouté et al, 1987
	750 - 765			Synthetic polylinker derived sequences	
hppdPfW336-1Pa gene	766 - 1842		Counter clockwise	Coding sequence of the 4-hydroxyphenylpyruvate dioxygenase gene of Pseudomonas fluorescens strain A32 modified by the replacement of the smino acid Glycine 336 with a Tryptophan adapted to cotton codon usage.	Boudec *et al*., 2001
TPotpY-1Pa	1843- 2214		Counter clockwise	Coding sequence of an optimized transit peptide derivative (position 55 changed into Tyr), containing sequence of the RuBisCO small subunit genes of *Zea mays* and *Helianthus annuus*, adapted for cotton codon usage.	Lebrun et al., 1996
	2215 - 2222			Synthetic polylinker derived sequences	
Pcsvmv	2223 - 2735		Counter clockwise	Sequence including the promoter region of the Cassava Vein Mosaic Virus.	Verdaguer *et al*., 1996
	2736 - 2795			Synthetic polylinker derived sequences	
lox	2796 - 2829		Clockwise	Sequence including the 34 bp recognition sequence for the Cre recombinase of bacteriophage P1.	Hoess and Abremski, 1985
	2830 - 2833			Synthetic polylinker derived sequences	
Ph4a748	2834 - 3750		Clockwise	Sequence including the promoter region of the histone H4 gene of *Arabidopsis thaliana*.	Chabouté *et al*., 1987
	3751 - 3789			Synthetic polylinker derived sequences	
Intron1 h3At	3790 - 4255		Clockwise	First intron of gene II of the histone H3.III variant of *Arabidopsis thaliana*.	Chaubet *et al*., 1992
	4256 - 4268			Synthetic polylinker derived sequences	
TPotpC	4269 - 4640		Clockwise	Coding sequence of the optimized transit peptide, containing sequence of the RuBisCO small subunit genes of *Zea Mays* and *Helianthus annuus*.	Lebrun *et al*., 1996
2mepsps gene	4641 - 5978		Clockwise	Coding sequence of the double-mutant 5-enol-pyruvylshikimate-3-phosphate synthase gene of *Zea mays*.	Lebrun *et al*., 1997
	5979 - 5998			Synthetic polylinker derived sequences	
ThistonAt	5999 - 6665		Clockwise	Sequence including the 3' untranslated region of the histone H4 gene of *Arabidopsis thaliana*.	Chabouté *et al*., 1987
	6666 - 6669			Synthetic polylinker derived sequences	
lox	6670 - 6703		Clockwise	Sequence including the 34 bp recognition sequence for the Cre recombinase of bacteriophage P1.	Hoess and Abremski, 1985
	6704 - 6831			Synthetic polylinker derived sequences	
LB	6832 - 6856			Left border repeat from the T-DNA of *Agrobacterium tumefaciens*.	Zambryski, 1988
	6857 - 7161			Ti-plasmid sequences of pTiAch5 flanking the left border repeat.	Zhu *et al*., 2000

GENETIC ELEMENT	NT POSITION	SIZE (KB)	ORIENTATION	DESCRIPTION AND FUNCTION	REFERENCE
aadA	7162 - 8946		Counter clockwise	Fragment including the aminoglycoside adenyltransferase gene of *Escherichia coli*.	Fling *et al*., 1985
ORI pVS1	8947 - 11736			Fragment including the origin of replication from the *Pseudomonas* plasmid pVS1 for replication in *Agrobacterium tumefaciens*.	Hajdukiewicz *et al*., 1994
ORI ColE1	11737 - 12893			Fragment including the origin of replication from the plasmid pBR322 for replication in *Escherichia coli*.	Bolivar *et al*., 1977
	12894 - 13099			Ti-plasmid and sequences of pTiAch5 flanking the right border repeat.	Zhu *et al*., 2000

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

A vector map of pTSIH09, containing the genetic elements described above in Table 3, is presented below in Figure 3, and in M-388224-02; Dossier Node A.3 (b), (i) and (ii)). Vector pTSIH09 is derived from pGSC1700 and pUC19. The locations of restriction sites within the transgenic locus are shown below in Figure 6.

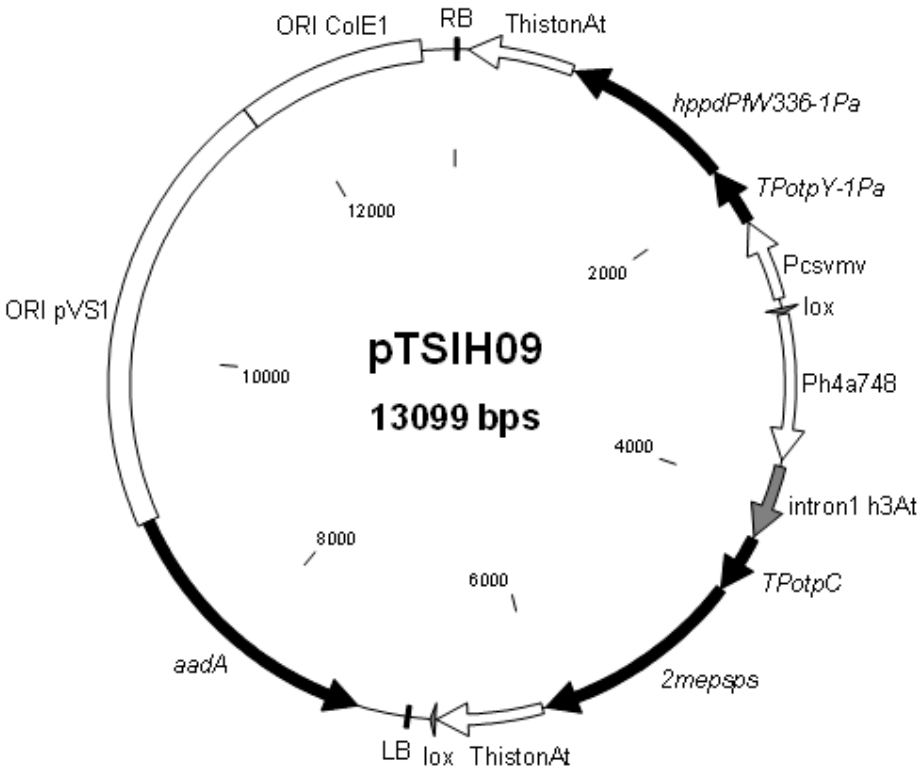


Figure 3 Map of plasmid vector pTSIH09 used in GHB811 *G. hirsutum*

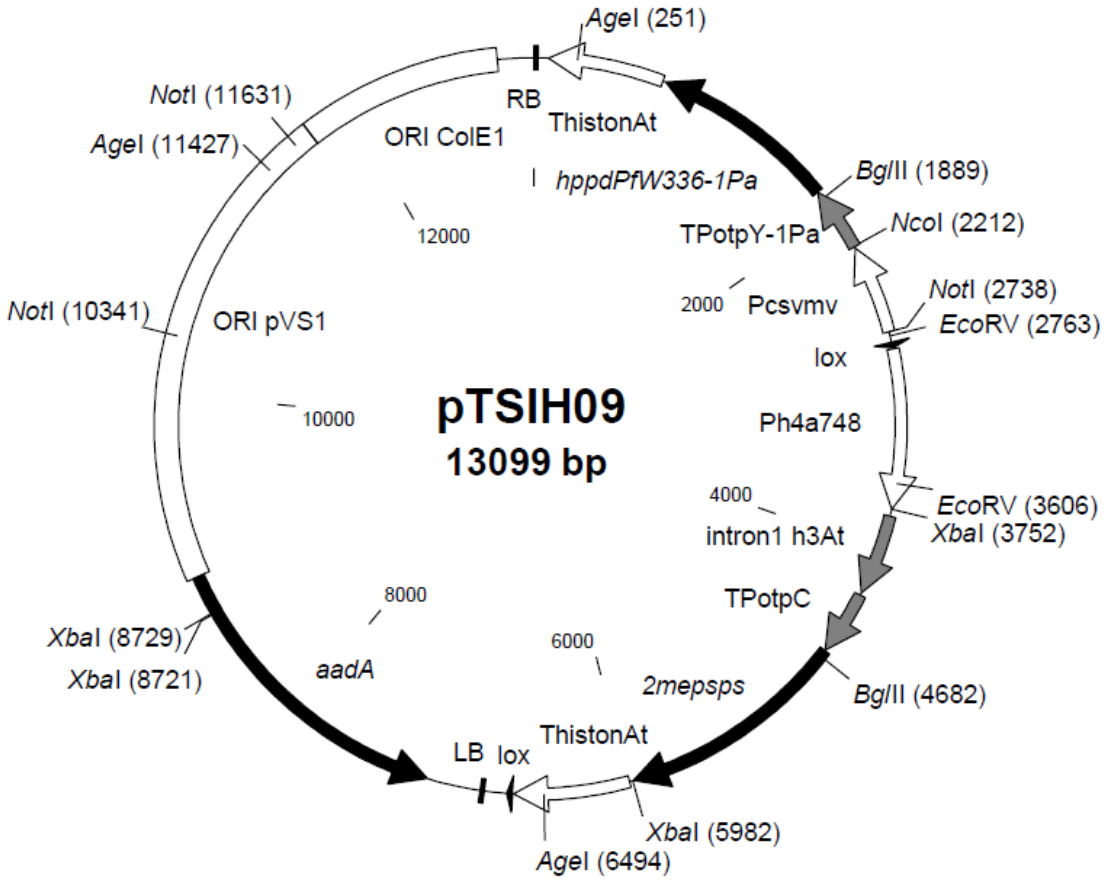


Figure 4 Map of transformation vector pTSIH09 with indication of restriction enzymes used for plasmid confirmation

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

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

The transgenic locus of cotton GHB811 (T1 generation) was characterized by means of Southern blot analysis on genomic DNA (gDNA) prepared from leaf material M-572036-01; Node A.3 (c), (i) (CCI)).

A set of restriction enzymes were chosen to produce different restriction fragments containing portions of the insert and adjacent genomic DNA for each enzyme, which generated a specific banding pattern on the Southern blots. The selection and design of probes used in this study allowed the investigation of the T-DNA insert organization. Probes used in this study are summarized in Table 4. Probes covering the different features of the transgenic cassettes (P001 to P008) as well as the probe covering the complete T-DNA region were used (P009) (Figure 5).

Figure 6 shows the expected fragments for a complete single copy of the complete T-DNA integration in a single locus of GHB811 cotton genome. Expected and obtained hybridization

fragments are listed in Table 5. The hybridization results to characterize the T-DNA insertion in GHB811 cotton are presented in Figure 7 to Figure 15.

Each membrane contained one negative control, which showed no hybridization with any of the probes used (Figure 7 to Figure 15, lane 13). Consequently, the absence of any background hybridization was demonstrated for all the probes used.

Similarly, each membrane contained a positive control. This positive control, consisting of pTSIH09 plasmid DNA, was digested with *Hin*cII and an equimolar amount was spiked in *Hin*cII digested gDNA from the non-GM counterpart. For each of the probes used, the expected fragments were detected (Figure 7 to Figure 15, lane 14), confirming that the applied experimental conditions allowed specific hybridization of the used probes with the target sequences.

***a. Restriction digestion with* AflII**

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, digestion of the insert with the *Afl*II restriction enzyme should generate two integration fragments and three internal fragments (Figure 6). The 5' integration fragment is expected to be greater than 1905 bp and contains genomic DNA flanking the 5' end of the insert, the right border (RB) sequence and the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene. The internal fragment with expected length of 1417 bp contains a small part of the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, the *hppdPfW336-1Pa* gene and the TPotpY-1Pa sequence. The internal fragment with expected length of 2222 bp contains a small part of the TPotpY-1Pa sequence, the Pcsvmv and Ph4a748 promoter sequences, a lox recognition site, the intron1 h3At sequence and a small part of the TPotpC sequence. The 1682 bp internal fragment contains the TPotpC sequence, the *2mepsps* gene sequence and a small part of the ThistonAt terminator sequence following the *2mepsps* gene. The 3' integration fragment is expected to be greater than 2102 bp and contains the ThistonAt terminator sequence following the *2mepsps* gene, the lox recognition site, left border (LB) sequences and genomic DNA flanking the 3' end of the insert.

Hybridization of the *Afl*II digested cotton GHB811 gDNA with *hppdPfW336-1Pa* probe results in one band of 1417 bp (Figure 8, lane 3). This band corresponds to one of the internal fragments and was also observed with the TPotpY-1Pa probe and the T-DNA probe, as expected (Figure 9 and Figure 15, lane 3). Hybridization of the *Afl*II digested cotton GHB811 gDNA with Pcsvmv, Ph4a748 + lox, intron1 h3At probe resulted in one band of 2222 bp (Figure 10 to Figure 12, lane 3). This band corresponds to one of the internal fragments and was also observed with the TPotpC probe, the TPotpY-1Pa probe and the T-DNA probe, as expected (Figure 9, Figure 13, and Figure 15, lane 3). Hybridization of the *Afl*II digested cotton GHB811 gDNA with *2mepsps* probe resulted in one band of 1682 bp (Figure 14, lane 3). This band corresponds to one of the internal fragments and was also observed with the TPotpC probe and the T-DNA probe, as expected (Figure 13 and Figure 15, lane 3). In addition, two strong bands with lengths of 2500 bp and 4400 bp were observed after hybridization with the ThistonAt probe and the T-DNA probe (Figure 7 and Figure 15, lane 3). These bands correspond to the two integration fragments. Yet, with this experimental setup, it was not possible to determine which of these two fragments represented the 5' or the 3' integration fragment.

In conclusion, the *Afl*II restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

b. Restriction digestion with Bsp*HI*

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, digestion of the insert with the *Bsp*HI restriction enzyme produces two integration fragments (Figure 6). The 5' integration fragment is expected to be 1365 bp and contains genomic DNA flanking the 5' end of the insert, the RB sequence, the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene and a part of the *hppdPfW336-1Pa* gene. The 3' integration fragment is expected to be greater than 7084 bp and contains part of the *hppdPfW336-1Pa* gene, the TPotpY-1Pa sequence, the Pcsvmv promoter sequence, a lox recognition site, the Ph4a748 promoter sequence, the intron1 h3At sequence, the TPotpC sequence, the *2mepsps* gene sequence, the ThistonAt terminator sequence following the *2mepsps* gene, the lox recognition site, the LB sequences and the genomic DNA flanking the 3' end of the insert.

Hybridization with the TPotpY-1Pa, Pcsvmv, Ph4a748 + lox, intron1 h3At, TPotpC and *2mepsps* probes (Figure 9 to Figure 14, lane 4) showed, as expected, the presence of only one band of 8300 bp which corresponds to the 3' integration fragment. Hybridization with the ThistonAt, *hppdPfW336-1Pa* and T-DNA probes (Figure 7, Figure 8 and Figure 15, lane 4) showed besides the band of 8300 bp, a second band of 1365 bp which corresponds to the 5' integration fragment.

In conclusion, the *Bsp*HI restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

c. Restriction digestion with Psi*I/*Sap*I*

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the double digestion of the insert with the *Psi*I and *Sap*I restriction enzymes produces two integration fragments and one internal fragment (Figure 6). The 5' integration fragment is expected to be 3423 bp and contains genomic DNA flanking the 5' end of the insert, the RB sequences, the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, the *hppdPfW336-1Pa* gene, the TPotpY-1Pa sequence, the Pcsvmv promoter sequence, a lox recognition site and a part of the Ph4a748 promoter sequence. The internal fragment with expected length of 1588 bp contains a part of the Ph4a748 promoter sequence, the intron1 h3At sequence, the TPotpC sequence and a part of the *2mepsps* gene sequence. The 3' integration fragment is expected to be 2644 bp and contains part of the *2mepsps* gene sequence, the ThistonAt terminator sequence following the *2mepsps* gene, the lox recognition site, the LB sequences and the genomic DNA flanking the 3' end of the insert.

Hybridization of the *Psi/Sap*I digested cotton GHB811 gDNA with the intron1 h3At and the TPotpC probe (Figure 12 and Figure 13, lane 5) showed the presence of one band of 1588 bp. This band corresponds to the internal fragment and was also observed with the Ph4a748 + lox probe, the *2mepsps* probe and the T-DNA probe, as expected (Figure 11, Figure 14 and Figure 15, lane 5). In addition, two strong bands with lengths of 2644 bp and 3423 bp are observed after hybridization with several probes. These bands correspond to the two integration fragments. The 3423 bp band, corresponding to the 5' integration fragment was observed with ThistonAt, *hppdPfW336-1Pa*, TPotpY-1Pa, Pcsvmv, Ph4a748 + lox, and T-DNA probes (Figure 7 to Figure 11, and Figure 15, lane 5), as expected. The 2644 bp band, corresponding to the 3' integration fragment, was observed with the ThistonAt, the *2mepsps* and the T-DNA probes (Figure 7, Figure 14 and Figure 15, lane 5), as expected.

In conclusion, the *Psi*I/*Sap*I restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

d. Restriction digestion with PvuI

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the digestion of the insert with the *Pvu*I restriction enzyme produces two internal fragments (Figure 6). The integration fragments are not expected to be observed because there is no or only very small overlap with the selected probes. The internal fragment with expected length of 5784 bp contains the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, the *hppdPfW336-1Pa* gene, the TPotpY-1Pa sequence, the Pcsvmv promoter sequence, a lox recognition site, the Ph4a748 promoter sequence, the intron1 h3At sequence, the TPotpC sequence and a part of the *2mepsps* gene sequence. The internal fragment with expected length of 921 bp contains a part of the *2mepsps* gene sequence, the ThistonAt terminator sequence following the *2mepsps* gene and the lox recognition site.

Hybridization of the *Pvu*I digested cotton GHB811 gDNA with the ThistonAt, *2mepsps* and T-DNA probes, resulted in a band which has an estimated size of 1050 bp. This size differed more than 10 % from the expected size of 921 bp (Figure 7, Figure 14, and Figure 15, lane 6). The fact that the size of this fragment (1050 bp) corresponds to the summation of the 921 bp internal fragment and a 136 bp 3' integration fragment demonstrates that this fragment is due to an incomplete digestion of a second, nearby *Pvu*I restriction site located on the 3' flanking sequence. Hybridization of the *Pvu*I digested cotton GHB811 gDNA with all feature probes and the T-DNA showed the presence of a band corresponding with the 5784 bp internal fragment (Figure 7 to Figure 15, lane 6).

In conclusion, the *Pvu*I restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

e. Restriction digestion with SacI

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the digestion of the insert with the *Sac*I restriction enzyme produces two integration fragments and one internal fragment (Figure 6). The 5' integration fragment is expected to be greater than 2740 bp and contains genomic DNA flanking the 5' end of the insert, the RB sequence, the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, and a part of the *hppdPfW336-1Pa* gene. The internal fragment with expected length of 1119 bp contains a part of the *hppdPfW336-1Pa* gene, the TPotpY-1Pa sequence and part of the Pcsvmv promoter sequence. The 3' integration fragment is expected to be 4494 bp and contains part of the Pcsvmv promoter sequence, the lox recognition site, the Ph4a748 promoter sequence, the intron1 h3At sequence, the TPotpC sequence, the *2mepsps* gene sequence, the ThistonAt terminator sequence following the *2mepsps* gene, the LB sequences and the genomic DNA flanking the 3' end of the insert.

Hybridization of the *Sac*I digested cotton GHB811 gDNA with the TPotpY-1Pa probe (Figure 9, lane 7) showed one fragment of 1119 bp. This band corresponds to the internal fragment and was also observed with the *hppdPfW336-1Pa* probe, the Pcsvmv probe and the T-DNA probe as expected (Figure 8, Figure 10 and Figure 15, lane 7). In addition, two bands with lengths of 4494 bp and >10 000 bp are observed after hybridization with several probes. These bands correspond to the two integration fragments. As the > 10 000 bp band is only observed with probes ThistonAt, *hppdPfW336-1Pa* and the T-DNA probe (Figure 7, Figure 8 and Figure 15, lane 7), this fragment corresponds to the 5' integration fragment. The

4494 bp band, corresponding to the 3' integration fragment is obtained with the ThistonAt, Pcsvmv, Ph4a748 + lox, intron1 h3At, TPotpC, the *2mepsps* and the T-DNA probes (Figure 7 and Figure 10 to Figure 15, lane 7), as expected.

Hybridization of the *Sac*I digested cotton GHB811 gDNA with the ThistonAt, Ph4a748 + lox, intron1 h3At, TPotpC, the *2mepsps* and the T-DNA probes, which are all probes fully overlapping the *Sac*I 3' integration fragment, resulted also in a weak fragment of 4900 bp (Figure 7, Figure 11 to Figure 15, lane 7). The weakness of the signal, combined with the fact that the size of this fragment (4900 bp) corresponds to the approximate summation of the 4494 bp 3' integration fragment and a 374 bp fragment in the 3' flanking sequence demonstrates that this fragment is due to incomplete digestion of the *Sac*I restriction site in the 3' flanking sequence (Figure 6).

In conclusion, the *Sac*I restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

f. Restriction digestion with ScaI

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the digestion of the insert with the *Sca*I restriction enzymes produces three internal fragments and a 3' integration fragment (Figure 6). The 5' integration fragment is not expected to be observed because there is no overlap with the selected probes. The internal fragment with expected length of 1890 bp contains the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, the *hppdPfW336-1Pa* gene and a small part of the TPotpY-1Pa sequence. The internal fragment with expected length of 342 bp contains the TPotpY-1Pa sequence and a small part of the Pcsvmv promoter sequence. The internal fragment with expected length of 2925 bp contains the Pcsvmv promoter sequence, a lox recognition site, the Ph4a748 promoter sequence, the intron1 h3At sequence, the TPotpC sequence and a part of the *2mepsps* gene sequence. The 3' integration fragment is expected to be greater than 2950 bp and contains a part of the *2mepsps* gene, the ThistonAt terminator sequence following the *2mepsps* gene, the lox recognition site, LB sequences and the genomic DNA flanking the 3' end of the insert.

Hybridization of the *Sca*I digested cotton GHB811 gDNA with Pcsvmv probe, the Ph4a748 + lox probe, the intron1 h3At probe and the TPotpC probe results in one band of 2925 bp (Figure 10 to Figure 13, lane 8). This band corresponds to one of the internal fragments and was also observed with the *2mepsps* probe and the T-DNA probe, as expected (Figure 14 and Figure 15, lane 8). Hybridization of the *Sca*I digested cotton GHB811 gDNA with *hppdPfW336-1Pa* probe resulted in one band of 1890 bp (Figure 8, lane 8). This band corresponds to one of the internal fragments and was also observed with the ThistonAt probe, the TPotpY-1Pa probe and the T-DNA probe, as expected (Figure 7, Figure 9 and Figure 15, lane 8). Hybridization of the *Sca*I digested cotton GHB811 gDNA with TPotpY-1Pa probe resulted, besides the 1890 bp fragment, in another fragment of 342 bp (Figure 9, lane 8). This band corresponds to the third internal fragment. This fragment is not observed using the T-DNA probe (Figure 15, lane 8), as expected, because of the small overlap with the probe in proportion to the large size of the probe. In addition, hybridization with the ThistonAt, the *2mepsps* and the T-DNA probes (Figure 7, Figure 14 and Figure 15, lane 8) resulted in a band with length of 5900 bp. This band corresponds to the 3' integration fragment.

In conclusion, the *Sca*I restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

g. Restriction digestion with PacI

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the digestion of the insert with the *Pac*I restriction enzymes produces a 3' integration fragment (Figure 6). The 5' integration fragment is not expected to be observed because there is no overlap with the selected probes. This 3' integration fragment is expected to be 7399 bp and contains the whole T-DNA starting from the ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene until the LB sequence and a part of the genomic DNA flanking the 3' end of the insert.

Hybridization of the *Pac*I digested cotton GHB811 gDNA with all feature probes and the T-DNA probe (Figure 7 to Figure 15, lane 9) showed the presence of a band with length of 7399 bp. This band corresponds to the 3' integration fragment.

In conclusion, the *Pac*I restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

h. Restriction digestion with StyI

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the digestion of the insert with the *Sty*I restriction enzyme produces two integration fragments and one internal fragment (Figure 6). The 5' integration fragment is expected to be greater than 3405 bp and contains the genomic DNA flanking the 5' end of the insert, RB sequences, ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, the *hppdPfW336-1Pa* gene and the TPotpY-1Pa sequence. The internal fragment with expected length of 2882 bp contains the Pcsvmv promoter sequence, a lox recognition site, the Ph4a748 promoter sequence, the intron1 h3At sequence, the TPotpC sequence and a part of the *2mepsps* gene sequence. The 3' integration fragment is expected to be 2356 bp and contains a part of the *2mepsps* gene, the ThistonAt terminator sequence following the *2mepsps* gene, the lox recognition site, LB sequences and the genomic DNA flanking the 3' end of the insert.

Hybridization of the *Sty*I digested cotton GHB811 gDNA with the Pcsvmv probe, the Ph4a748 + lox probe, the intron1 h3At probe and the TPotpC probe resulted in one band of 2882 bp (Figure 10 to Figure 13, lane 10). This band corresponds to the internal fragment and was also observed with the *2mepsps* probe and the T-DNA probe (Figure 14 and Figure 15, lane 10), as expected. In addition, two strong bands with lengths of 3400 bp and 2356 bp were observed after hybridization with several probes. The obtained fragment size of 3400 bp was smaller than the expected fragment size (>3405 bp) due to rounding of values. These bands correspond to the two integration fragments. As the 3400 bp band was only observed with probes ThistonAt, *hppdPfW336-1Pa*, TPotpY-1Pa and the T-DNA probe, this band corresponds to the 5' integration fragment (Figure 7 to Figure 9, and Figure 15, lane 10). The 2356 bp band was obtained with the ThistonAt, the *2mepsps* and the T-DNA probes and not with the other probes (Figure 7, Figure 14 and Figure 15, lane 10). Therefore, this band corresponds to the 3' integration fragment.

In conclusion, the *Sty*I restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

i. Restriction digestion with Hin*dIII*

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the digestion of the insert with the *Hin*dIII restriction enzyme produces two integration fragments and one internal fragment (Figure 6). The 5' integration fragment is expected to be greater than 3978 bp and contains the genomic DNA flanking the 5' end of the insert, RB sequences, ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, the *hppdPfW336-1Pa* gene, the TPotpY-1Pa sequence and the Pcsvmv promoter sequence. The internal fragment with expected length of 979 bp contains a lox recognition site and the Ph4a748 promoter sequence. The 3' integration fragment is expected to be 4013 bp and contains the intron1 h3At sequence, the TPotpC sequence, the *2mepsps* gene sequence, the ThistonAt terminator sequence following the *2mepsps* gene, the lox recognition site, LB sequences and the genomic DNA flanking the 3' end of the insert.

Hybridization of the *Hin*dIII restriction digestion of cotton GHB811 with the Ph4a748 + lox probe and the T-DNA probe confirmed the presence of the 979 bp internal fragment (Figure 11 and Figure 15, lane 11). Hybridization of the *Hin*dIII restriction digestion of cotton GHB811 with probes intron1 h3At, TPotpC and *2mepsps* resulted in one band of 4013 bp (Figure 12 to Figure 14, lane 11). This band corresponds to the 3' integration fragment and was also observed with the ThistonAt probe and the T-DNA probe (Figure 7 and Figure 15, lane 11), as expected. Hybridization of the *Hin*dIII restriction digestion of cotton GHB811 gDNA with probes ThistonAt, *hppdPfW336-1Pa*, TPotpY-1Pa, Pcsvmv and the T-DNA probe resulted in another band of 4300 bp (Figure 7 to Figure 10, and Figure 15, lane 11) which corresponds to the 5' integration fragment.

Hybridization of the *Hin*dIII restriction digestion of cotton GHB811 with the *hppdPfW336-1Pa* probe also resulted in a weak 5200 bp fragment (Figure 8, lane 11). The weakness of the signal, combined with the fact that the size of this fragment (5200 bp) corresponds to the approximate summation of the 4300 bp 5' integration fragment and a 979 bp internal fragment demonstrates that this fragment is due to incomplete digestion of the *Hin*dIII restriction site between the Pcsvmv and lox recognition site (Figure 8).

In conclusion, the *Hin*dIII restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

j. Restriction digestion with Ase*I*

Based on the presented single copy integration model for the T-DNA region of the pTSIH09 plasmid, the digestion of the insert with the *Ase*I restriction enzyme produces two integration fragments and one internal fragment (Figure 6). The 5' integration fragment is expected to be 3508 bp and contains the genomic DNA flanking the 5' end of the insert, RB sequences, ThistonAt terminator sequence following the *hppdPfW336-1Pa* gene, the *hppdPfW336-1Pa* gene, the TPotpY-1Pa sequence and the Pcsvmv promoter sequence. The internal fragment with expected length of 2462 bp contains a lox recognition site, the Ph4a748 promoter sequence, the intron1 h3At sequence, the TPotpC sequence and a part of the *2mepsps* gene sequence. The 3' integration fragment is expected to be 2081 bp and contains a part of the *2mepsps* gene sequence, the ThistonAt terminator sequence following the *2mepsps* gene, the lox recognition site, LB sequences and the genomic DNA flanking the 3' end of the insert.

Hybridization of the *Ase*I restriction digestion of cotton GHB811 with the Ph4a748 + lox probe, the intron1 h3At probe and the TPotpC probe resulted in one band of 2462 bp (Figure 11 to Figure 13, lane 12). This band corresponds to the internal fragment and was also

observed with the *2mepsps* probe and the T-DNA probe (Figure 14 and Figure 15, lane 12), as expected. Hybridization of the *Ase*I restriction digestion of cotton GHB811 gDNA with probes *hppdPfW336-1Pa*, TPotpY-1Pa and Pcsvmv resulted in one band of 3508 bp (Figure 8 to Figure 10, lane 12). This band corresponds to the 5' integration fragment and was also observed with the ThistonAt and the T-DNA probes (Figure 7 and Figure 15, lane 12), as expected. Hybridization of the *Ase*I restriction digestion of cotton GHB811 with probes ThistonAt, *2mepsps* and the T-DNA probe resulted in another band of 2081 bp (Figure 7, Figure 14 and Figure 15, lane 12) which corresponds to the 3' integration fragment.

In conclusion, the *Ase*I restriction digestion of the cotton GHB811 gDNA and subsequent hybridization with all used probes confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09.

Conclusion

Digestions of gDNA from cotton GHB811 by a set of restriction enzymes and subsequent hybridizations with the different probes that spanned the complete T-DNA of pTSIH09 confirmed the integration model based on a single copy of the complete T-DNA region of pTSIH09. A few weak additional bands were observed in hybridization of the *Sac*I or *Hin*dIII digested cotton GHB811 gDNA. The weakness of the signal, combined with the fact that the size of those bands corresponds to the approximate summation of two fragments produced by complete digestion, demonstrates that those additional weak bands are due to incomplete digestion of the *Sac*I *or Hin*dIII restriction site. Therefore, it was demonstrated that a single copy of the complete T-DNA of the pTSIH09 plasmid was inserted at a single locus of the cotton GHB811 genome.

Table 4: Information on the probes used

Probe ID	Description	Primer pair/ Restriction digest	Primer sequence (5'→3')	Primer position on pTSIH09 (bp)	Size probe (bp)	Overlap between probe
P001	ThistonAt	GLPA449°	CCCGATCAAATCTGAGGGAC	83 → 102 and 6665 → 6646	646 6583* 7808*	NA
		GLPA456°	CTGGGTTTCTCACTTAAGCG	728 → 709 and 6020 → 6039		
P002	*hppdPfW336-1Pa*	GLPA457	AAACGGGTCCCATGAGAGTC	882 → 901	939	NA
		GLPA312	CTATGGGACTCATGGGTTTC	1820 → 1801		
P003	TPotpY-1Pa	GLPA459	ACCTCCGTTGCTAACATTCC	1855 → 1874	338	NA
		GLPA460	TTGCCACTGTTTCACGTACC	2192 → 2173		
P004	Pcsvmv	GLPA473	CAAATGCCGAACTTGGTTCC	2303 → 2322	439	NA
		GLPA474	GGCCGCGAAGGTAATTATCC	2741 → 2722		
P005	Ph4a748 + lox	GLPA472	CCCTGTTATCCCTAAAGCTTATTAA TATAAC	2770 → 2800	997	NA
		GLPA462	CGTGGGATCCTCTAGAGTCG	3762 → 3747**		
P006	intron1 h3At	GLPA073	TCAGGCGAAGAACAGGTATG	3785 → 3804	507	NA
		GLPA463	ACTGAGGAGGAGATCGAAGC	4291 → 4272		
P007	TPotpC	GLPA464	GCTTCGATCTCCTCCTCAGT	4272 → 4291	363	NA
		GLPA465	GATCCTTCCGCCGTTGCTGA	4634 → 4615		
P008	*2mepsps*	GLPA075	GCGCCGAGGAGATCGTGCTGC	4648 → 4668	1312	NA
		GLPA076	CTCAGCACATCGAAGTAGTC	5959 → 5940		
P009	T-DNA	GLPA467	AAGGCCCGATCAAATCTGAG	79 → 98	6700	NA
		GLPA468	GTGCCGTAATGCCGTAATGC	6778 → 6759		
P010	Vector backbone - *aadA*	GLPA032	GCCGCCGCTGCCGCTTTGC	6853 → 6871	1990	395 bp
		GLPA352	AGATCCTTGACCCGCAGTTG	8842 → 8823		
P011	Vector backbone - ORI pVS1	GLPA180	GAACCGAACAGGCTTATGTC	8448 → 8467	2354	
		GLPA469	GCGTGGTGTTTAACCGAATG	10801 → 10782		261 bp
P012	Vector backbone - ORI ColE1	GLPA470	TCCGCTACGAGCTTCCAGAC	10541 → 10560	2559	
		GLPA161	TGTCGCGTGTGAATAAGTCGC	13099 → 13079		

NA means not applicable
* Two additional PCR products of 6583 bp and 7808 bp can be generated with these primers. Only the fragment of interest (646 bp) was produced.
** Part of the GLPA462 does not bind on pTSIH09
° These primers amplify two identical regions

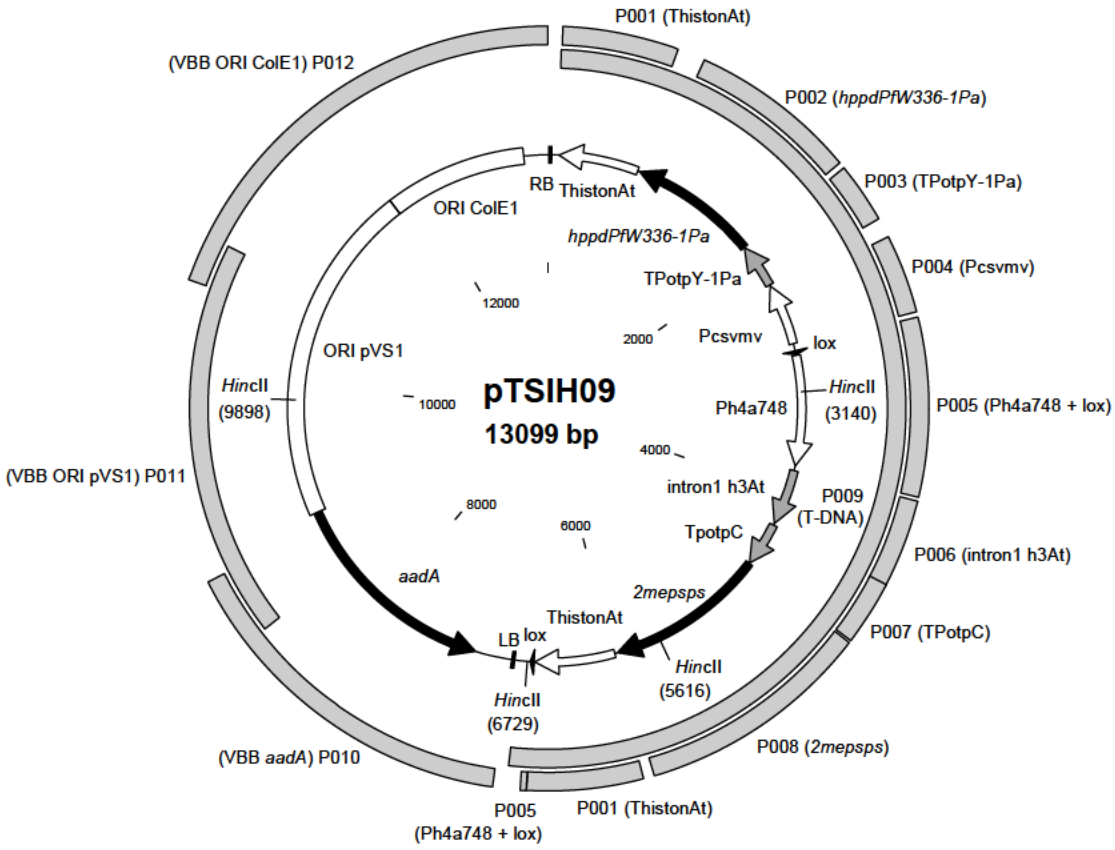


Figure 5 Map of transformation vector pTSIH09 with indication of restriction enzymes and probes used for Southern blot analysis

The indicated restriction enzyme positions between brackets refer to the first base after the cleavage site of the restriction enzyme.

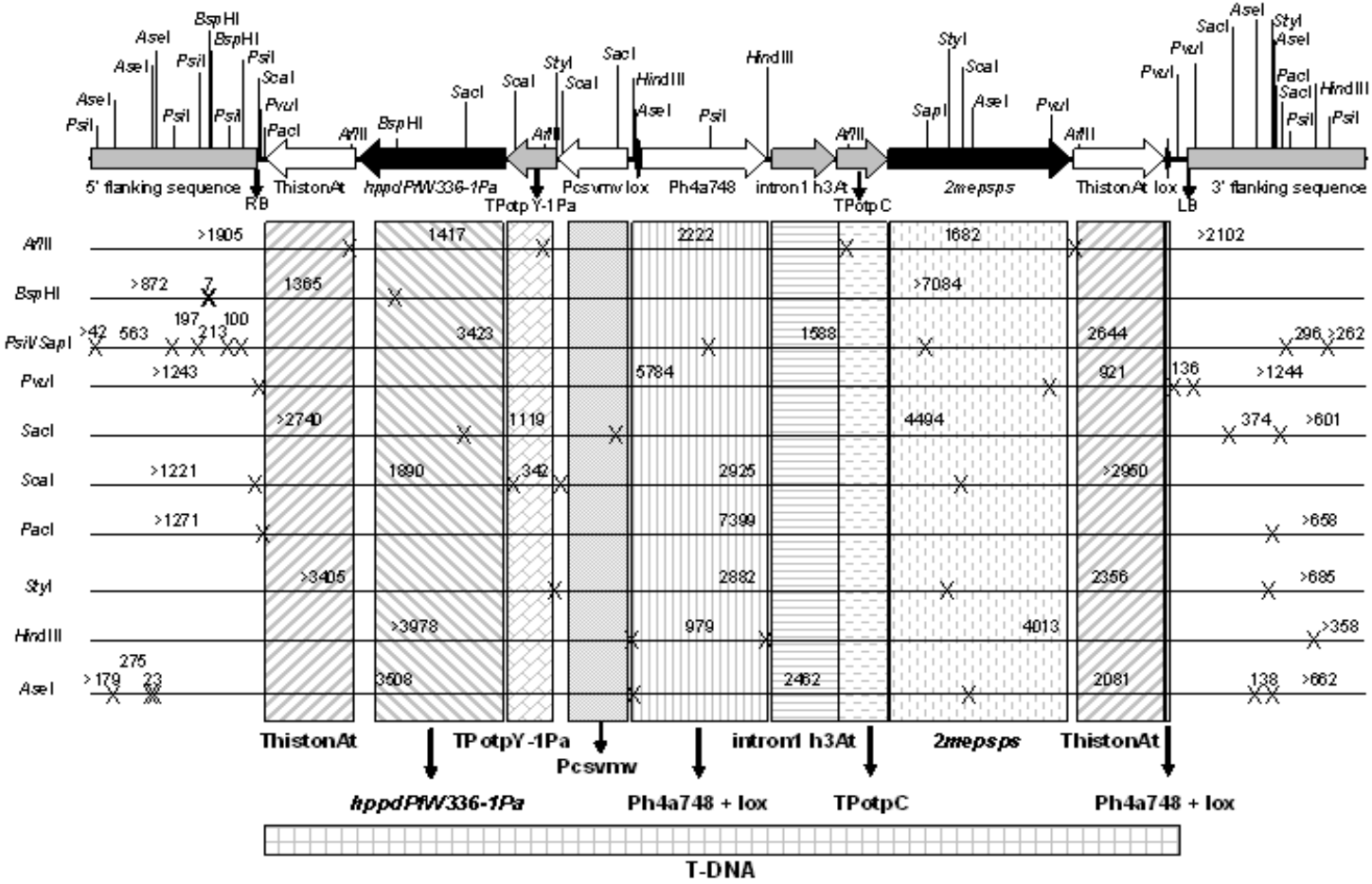


Figure 6: Schematic overview of the cotton GHB811 trangenic locus sequence (with indication of the different restriction enzymes and probes used for the cotton GHB811 insert characterization and expected fragment size (bp).

Table 5: Expected and obtained hybridization fragments determined for the insert characterization of cotton GHB811

Sample	Expected fragment sizes (bp)[a]	Obtained fragment sizes (bp)	Fragment description	H1/THT063B/28-F3		H1/THT063B/29-F4		H1/THT063B/30-F3		H1/THT063B/31-F2		H2/THT063B/30-F3		H2/THT063B/31-F2		H3/THT063B/31-F2		H2/THT063B/29-F2		H2/THT063B/28-F2	
				P001-3 ThistonAt		P002-1 *hppdPfW336-1Pa*		P003-3 TPotpY-1Pa		P004-3 Pcsvmv		P005-2 Ph4a748 + lox		P006-2 intron1 h3At		P007-3 TPotpC		P008-2 *2mepsps*		P009-7 T-DNA	
				Figure 9		Figure 10		Figure 11		Figure 12		Figure 13		Figure 14		Figure 15		Figure 16		Figure 17	
				Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
GHB811 - *Afl*II	> 1905	2500[$]	5' integration fragment	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	1417	1417	internal fragment	Yes**(17)	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	2222	2222	internal fragment	No	No	No	No	Yes**(64)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes**(79)	Yes	No	No	Yes	Yes
	1682	1682	internal fragment	Yes**(13)	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	> 2102	4400[$]	3' integration fragment	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	No	No	No	No	No	No	Yes	Yes
GHB811 - *Bsp*HI	1365	1365	5' integration fragment	Yes	Yes	Yes**(169)	Yes	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	> 7084	8300	3' integration fragment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GHB811 – *Psi*I/ *Sap*I	3423	3423	5' integration fragment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	Yes	Yes
	1588	1588	internal fragment	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2644	2644	3' integration fragment	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	No	No	No	No	Yes	Yes	Yes	Yes
GHB811 – *Pvu*I	> 1243	/	5' integration fragment	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
	5784	5784	Internal fragment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	921	1050 [c]	Internal fragment	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	No	No	No	No	Yes**(126)	Yes	Yes	Yes
	136		3' integration fragment	No		No		No		No		No		No		No		No		Yes**(24)	

Sample	Expected fragment sizes (bp)[a]	Obtained fragment sizes (bp)	Fragment description	H1/THT063B/28-F3		H1/THT063B/29-F4		H1/THT063B/30-F3		H1/THT063B/31-F2		H2/THT063B/30-F3		H2/THT063B/31-F2		H3/THT063B/31-F2		H2/THT063B/29-F2		H2/THT063B/28-F2	
				P001-3 ThistonAt		P002-1 *hppdPfW336-1Pa*		P003-3 TPotpY-1Pa		P004-3 Pcsvmv		P005-2 Ph4a748 + lox		P006-2 intron1 h3At		P007-3 TPotpC		P008-2 *2mepsps*		P009-7 T-DNA	
				Figure 9		Figure 10		Figure 11		Figure 12		Figure 13		Figure 14		Figure 15		Figure 16		Figure 17	
				Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
GHB811 – *Sac*I	> 2740	> 10 000	5' integration fragment	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	1119	1119	Internal fragment	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	Yes	Yes
	4494	4494	3' integration fragment	Yes	Yes	No	No	No	No	Yes** (76)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	/	4900	additional fragment (partial digestion)	No	Yes	No	No	No	No	No	No	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
GHB811 – *Sca*I	> 1221	/	5' integration fragment	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
	1890	1890	Internal fragment	Yes	Yes	Yes	Yes	Yes**(63)	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	342*	342*	Internal fragment	No	No	No	No	Yes*	Yes	No	No	No	No	No	No	No	No	No	No	Yes*	No
	2925	2925	Internal fragment	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	> 2950	5900	3' integration fragment	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	No	No	No	No	Yes	Yes	Yes	Yes
GHB811 – *Pac*I	> 1271	/	5' integration fragment	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
	7399	7399	3' integration fragment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GHB811 – *Sty*I	> 3405[b]	3400[b]	5' integration fragment	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes
	2882	2882	Internal fragment	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2356	2356	3' integration fragment	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	No	No	No	No	Yes	Yes	Yes	Yes
GHB811 – *Hin*dIII	> 3978	4300	5' integration fragment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes**(15)	No	No	No	No	No	No	No	Yes	Yes
	979	979	Internal fragment	No	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No	No	Yes	Yes

Sample	Expected fragment sizes (bp)[a]	Obtained fragment sizes (bp)	Fragment description	H1/THT063B/28-F3		H1/THT063B/29-F4		H1/THT063B/30-F3		H1/THT063B/31-F2		H2/THT063B/30-F3		H2/THT063B/31-F2		H3/THT063B/31-F2		H2/THT063B/29-F2		H2/THT063B/28-F2	
				P001-3 ThistonAt		P002-1 *hppdPfW336-1Pa*		P003-3 TPotpY-1Pa		P004-3 Pcsvmv		P005-2 Ph4a748 + lox		P006-2 intron1 h3At		P007-3 TPotpC		P008-2 *2mepsps*		P009-7 T-DNA	
				Figure 9		Figure 10		Figure 11		Figure 12		Figure 13		Figure 14		Figure 15		Figure 16		Figure 17	
				Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
	4013	4013	3' integration fragment	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	/	5200	additional fragment (partial digestion)	No	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No
GHB811 – *Ase*I	3508	3508	5' integration fragment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes**(22)	No	No	No	No	No	No	No	Yes	Yes
	2462	2462	Internal fragment	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2081	2081	3' integration fragment	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	No	No	No	No	Yes	Yes	Yes	Yes
non-GM counterpart (Coker 312) - *Bsp*HI	/	/	negative control	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
non-GM counterpart - *Hin*cII + one equimolar amount pTSIH09 - *Hin*cII	6341	6341	positive control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	Yes	Yes
	2476	2476	positive control	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	1113	1113	positive control	Yes	Yes	No	No	No	No	No	No	Yes**(34)	No	No	No	No	No	Yes	Yes	Yes	Yes
	3169	3169	positive control	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes**(50)	Yes

Exp. : expected; Obt.: obtained;

[a] Expected fragment sizes are based on the full sequence of GHB811 as determined in

[b] The obtained fragment size is smaller than the expected fragment size because of rounding of the values

[c] Instead of the expected fragment of 921 bp in the *Pvu*I digest, a 1050 bp fragment was obtained, which was the result of incomplete digestion at the *Pvu*I restriction site neighboring the LB region. The size of this fragment (1050 bp) corresponds to the summation of the 921 bp internal fragment and a 136 bp 3' integration fragment.

* Based on technical limitations of the Southern blotting technique, these fragments might be too small for visualization

** Due to a small overlap with the probe, these fragments are not always be visible. The size of the overlap is indicated between brackets.

° The obtained fragment size exceeded the 10 % range of the expected fragment size.

[$] With this experimental setup, it is not possible to assign this fragment to either the 5' or 3' integration fragment

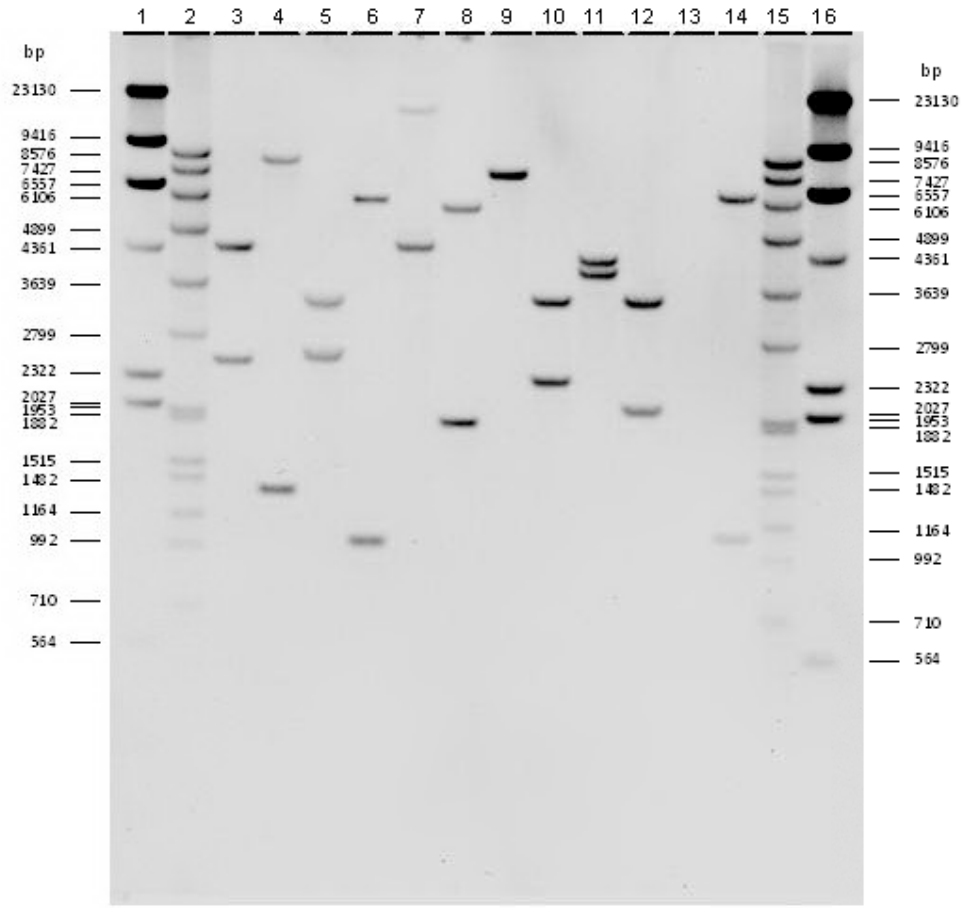


Figure 7: Hybridization performed with a ThistonAt probe (P001) to determine the insert organization of cotton GHB811

Digital image: H1/THT063B/28-F3

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 ThistonAt sequence (P001-3).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

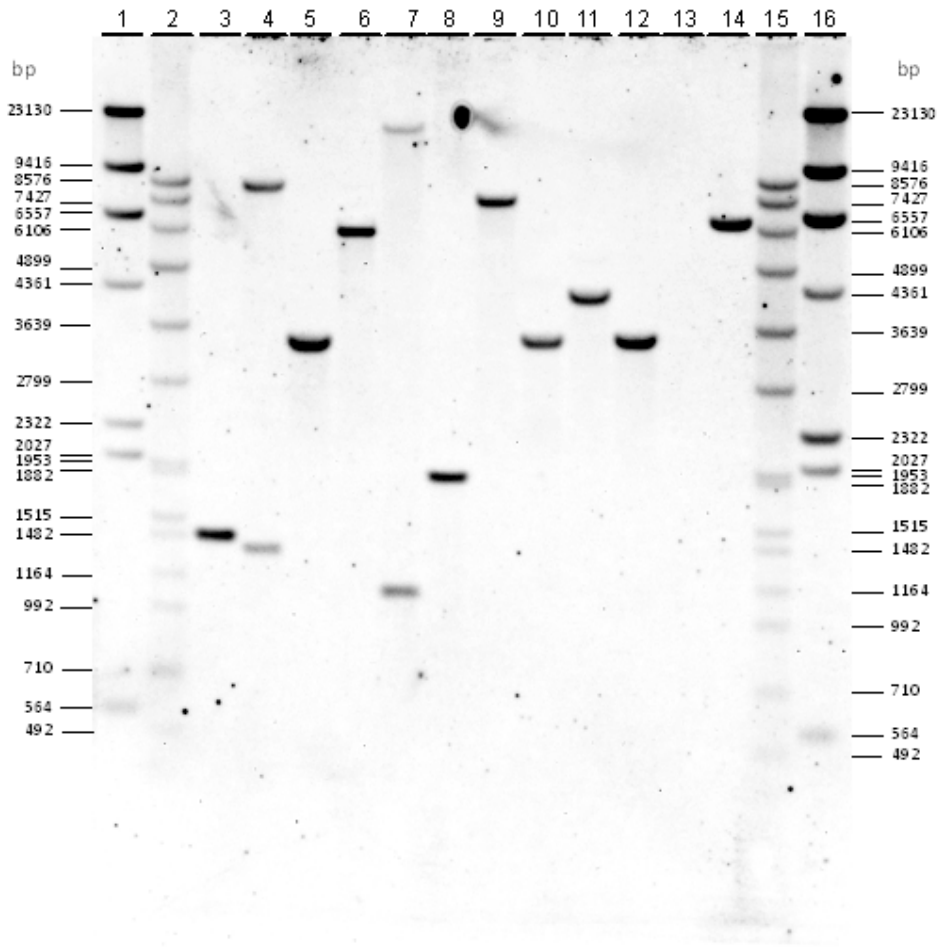


Figure 8: Hybridization performed with a *hppdPfW336-1Pa* probe (P002) to determine the insert organization of cotton GHB811

Digital image ID: H1/THT063B/29-F4

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 *hppdPfW336-1Pa* sequence (P002-1).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

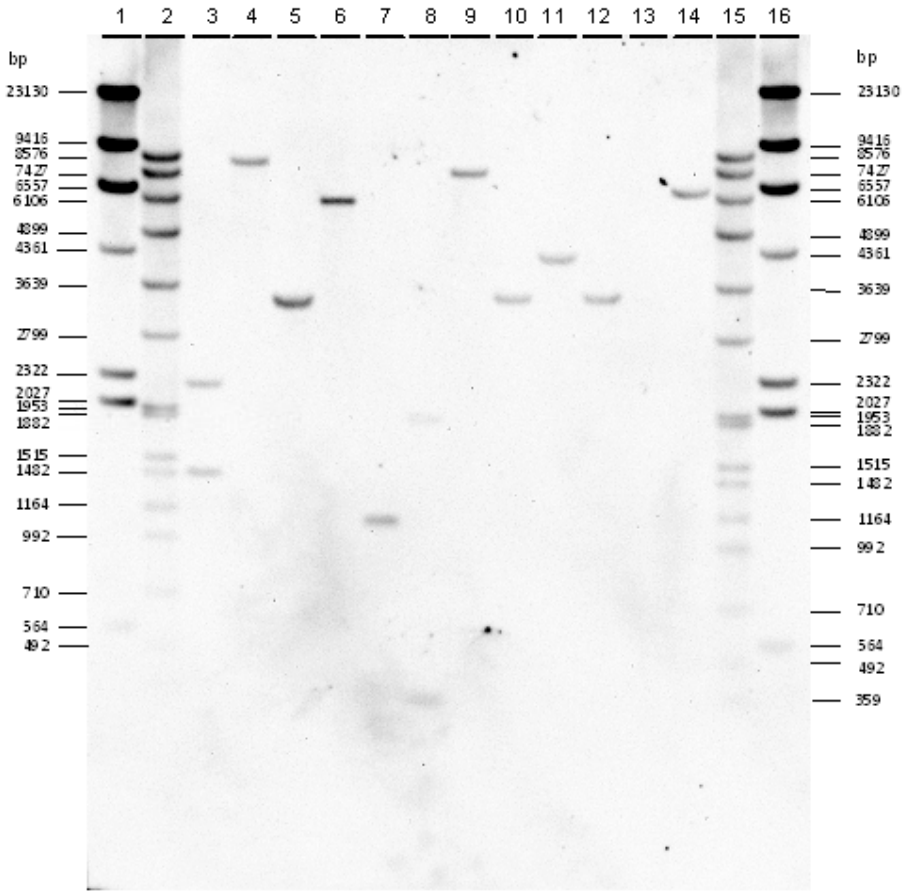


Figure 9: Hybridization performed with a TPotpY-1Pa probe (P003) to determine the insert organization of cotton GHB811

Digital image ID: H1/THT063B/30-F3

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 TPotpY-1Pa sequence (P003-3).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

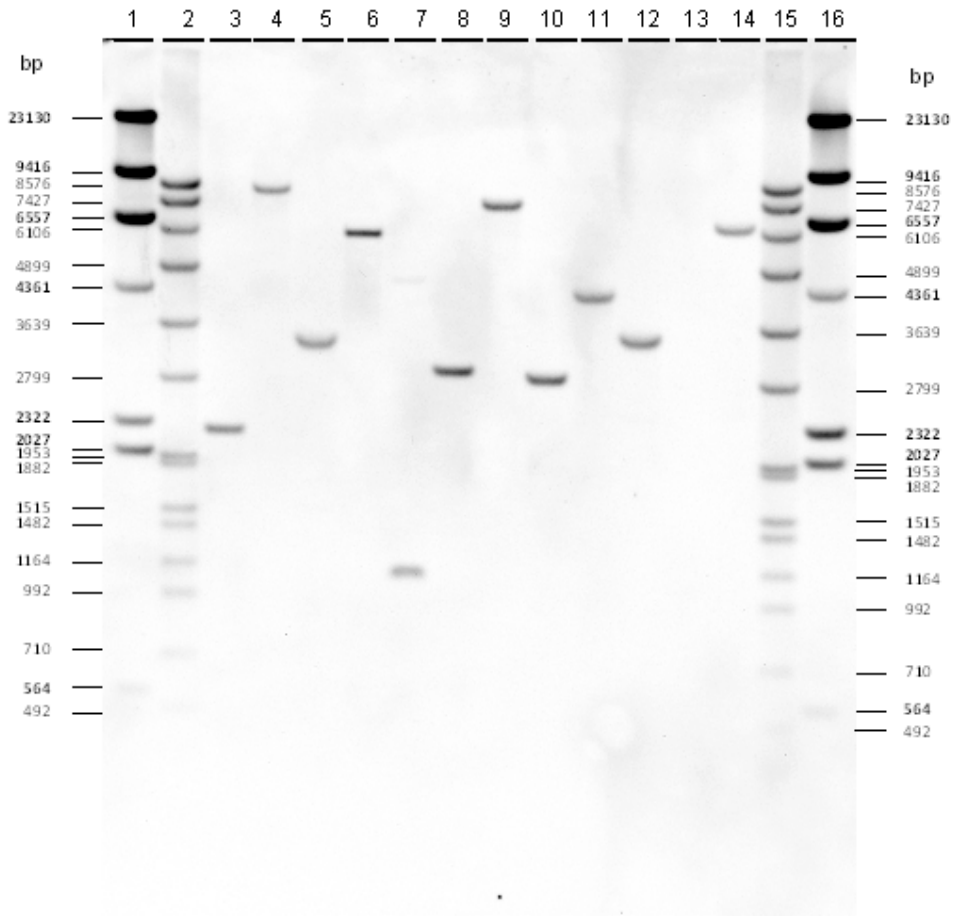


Figure 10: Hybridization performed with a Pcsvmv probe (P004) to determine the insert organization of cotton GHB811

Digital image ID: H1/THT063B/31-F2

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 Pcsvmv sequence (P004-3).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

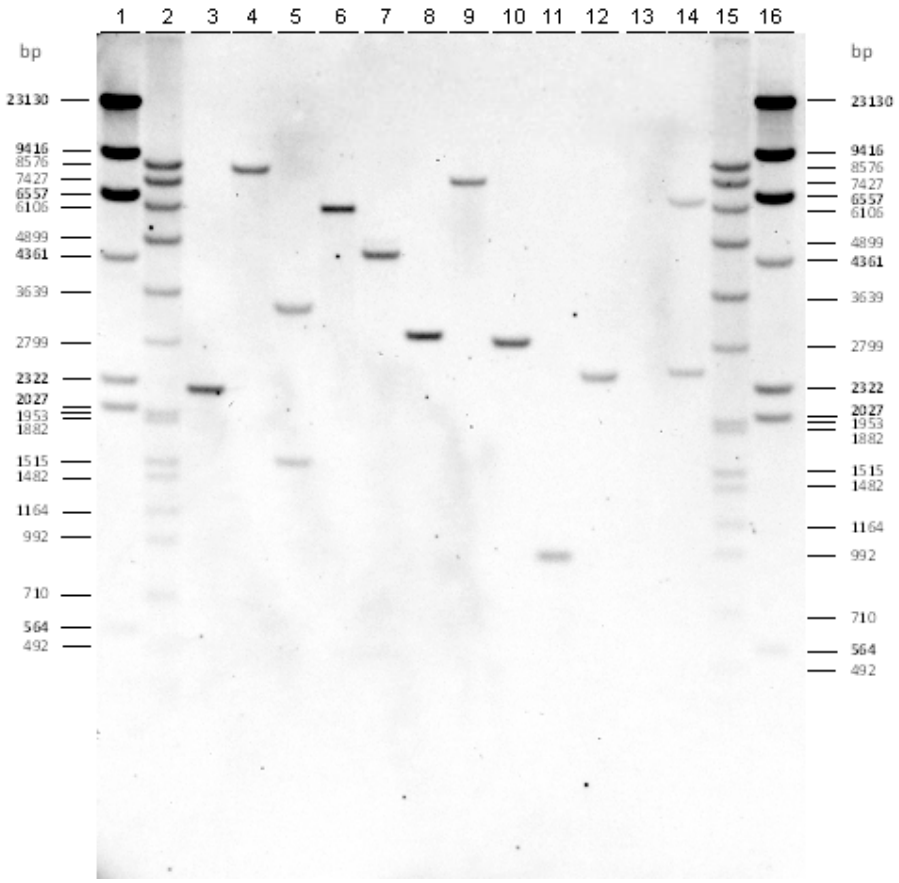


Figure 11: Hybridization performed with a Ph4a748 + lox (P005) to determine the insert organization of cotton GHB811

Digital image ID: H2/THT063B/30-F3

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 Ph4a748 + lox recognition site (P005-2).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

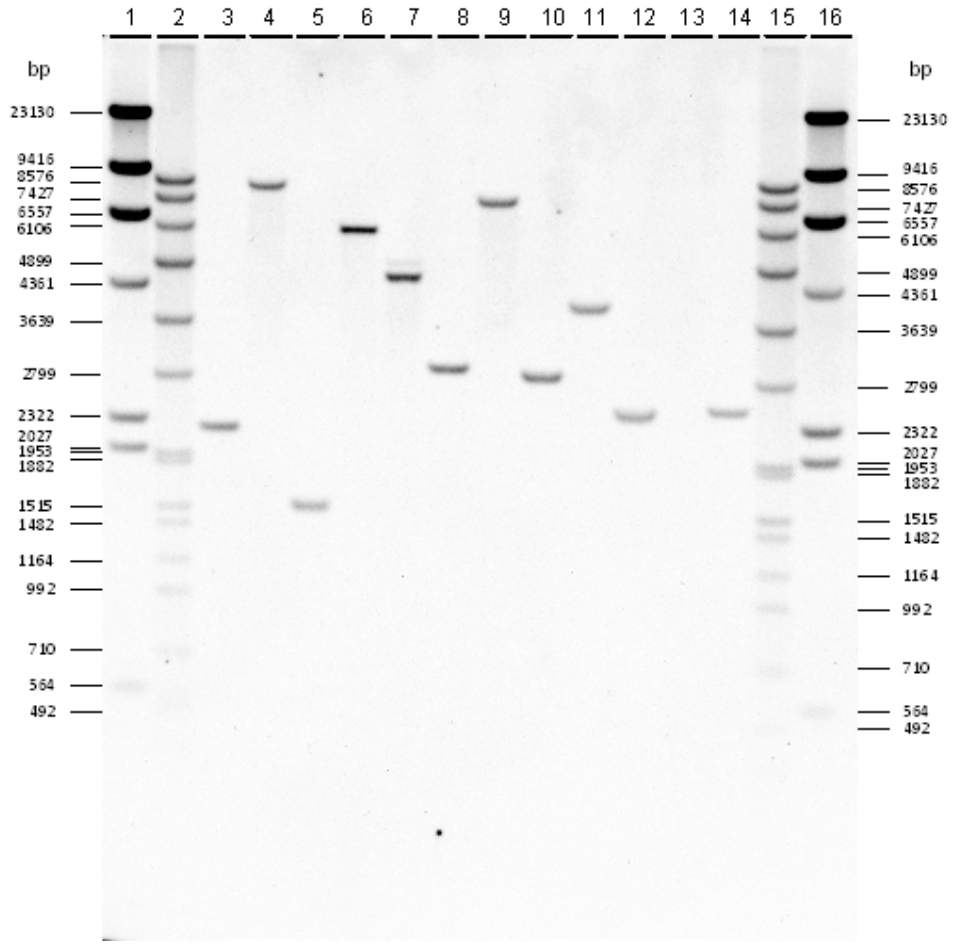


Figure 12: Hybridization performed with a intron1 h3At probe (P006) to determine the insert organization of cotton GHB811

Digital image ID: H2/THT063B/31-F2

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 intron1 h3At sequence (P006-2).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

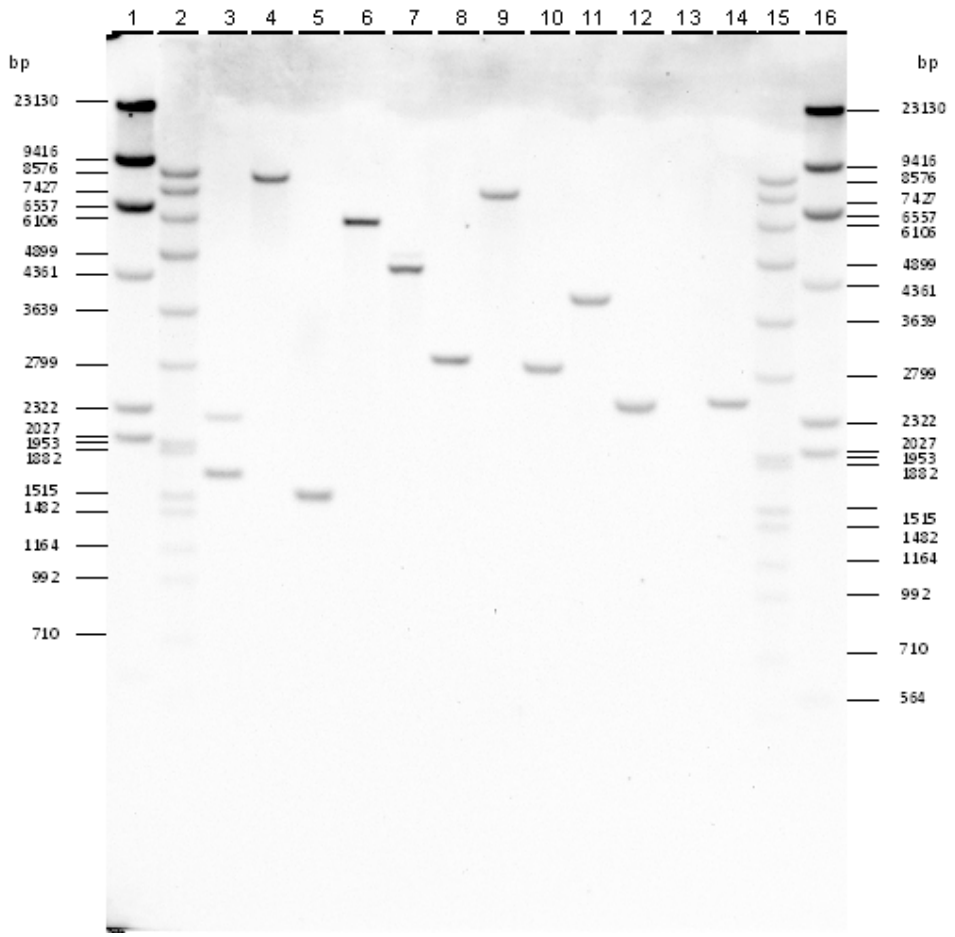


Figure 13: Hybridization performed with a TPotpC probe (P007) to determine the insert organization of cotton GHB811

Digital image ID: H3/THT063B/31-F2

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 TPotpC sequence (P007-3).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

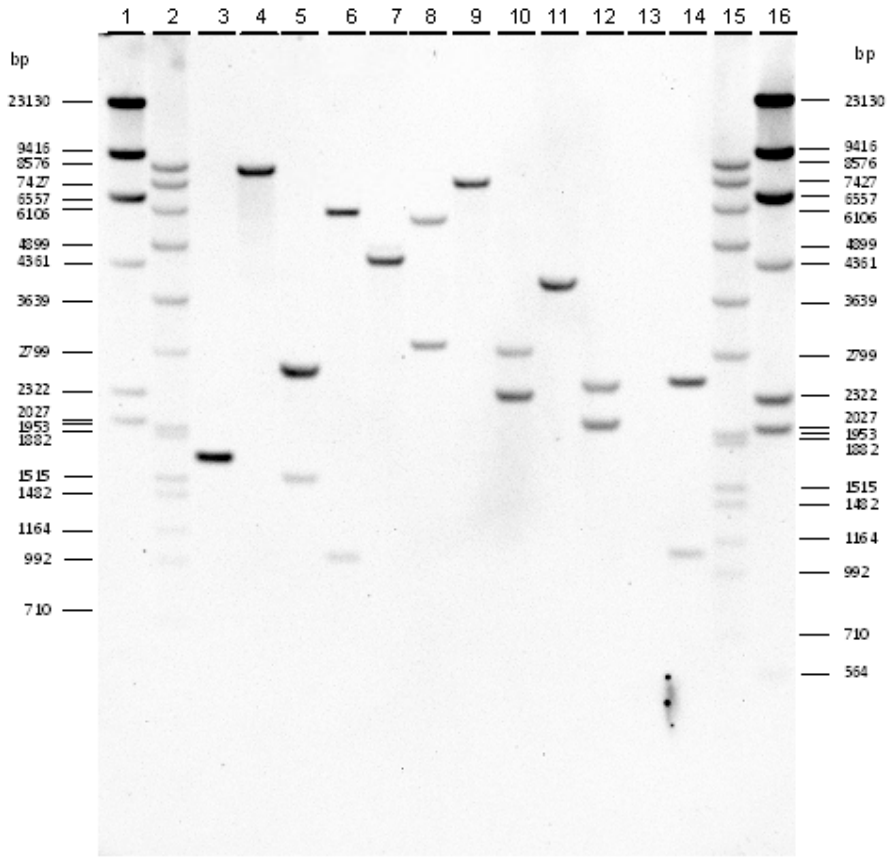


Figure 14: Hybridization performed with a *2mepsps* probe (P008) to determine the insert organization of cotton GHB811

Digital image ID: H2/THT063B/29-F2

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 *2mepsps* sequence (P008-2).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

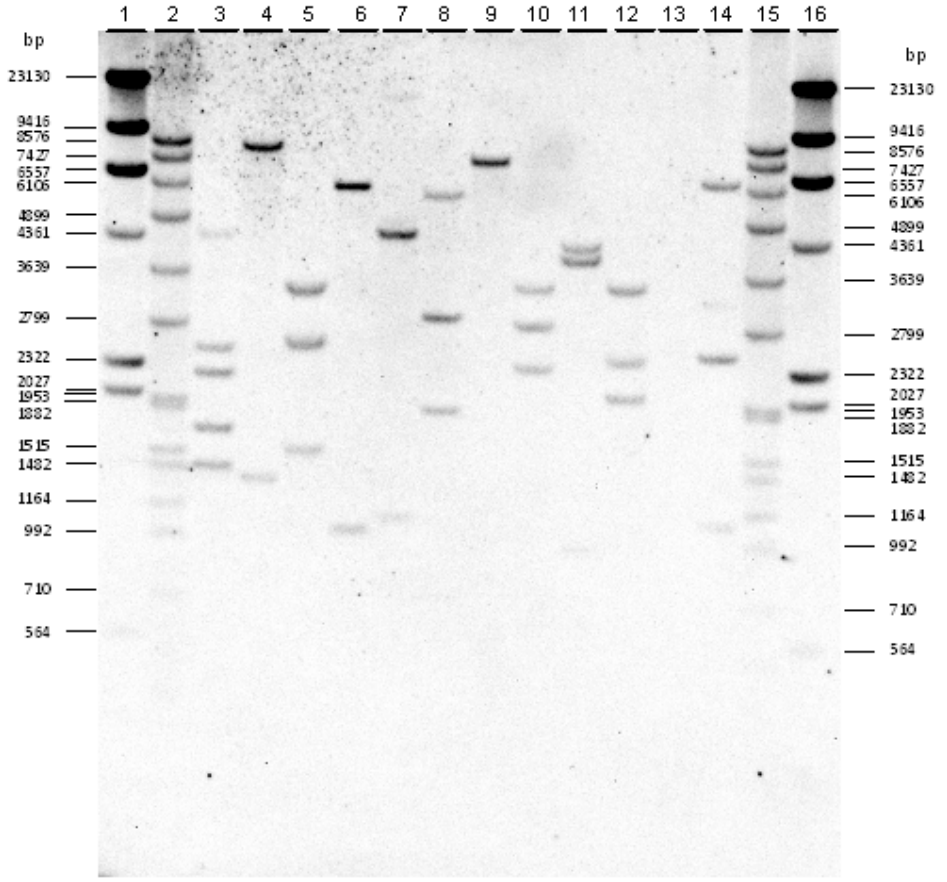


Figure 15: Hybridization performed with a T-DNA probe (P009) to determine the insert organization of cotton GHB811

Digital image ID: H2/THT063B/28-F2

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with different restriction enzymes and hybridized with a probe corresponding to the cotton GHB811 T-DNA region (P009-7).

Lane 1: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3 µg gDNA from cotton GHB811 - *Afl*II digested
Lane 4: 3 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 5: 3 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 6: 3 µg gDNA from cotton GHB811 - *Pvu*I digested
Lane 7: 3 µg gDNA from cotton GHB811 - *Sac*I digested
Lane 8: 3 µg gDNA from cotton GHB811 - *Sca*I digested
Lane 9: 3 µg gDNA from cotton GHB811 - *Pac*I digested
Lane 10: 3 µg gDNA from cotton GHB811 - *Sty*I digested
Lane 11: 3 µg gDNA from cotton GHB811 - *Hin*dIII digested
Lane 12: 3 µg gDNA from cotton GHB811 - *Ase*I digested
Lane 13: 3 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 14: 3 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 15: 3 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 16: 3 µg gDNA from non-GM counterpart - *Hin*dIII digested + 7.5 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

The absence of vector backbone sequences

The presence or absence of vector backbone sequences in cotton GHB811 was assessed by means of Southern blot analysis ; M-572036-01; Node A.3 (c), (i) (CCI)).

Genomic DNA (gDNA) from the cotton GHB811 T1 and BC2F3 generations was digested with restriction enzymes and subjected to Southern blot analysis using probes (P010 –P012) that collectively spanned the complete plasmid vector backbone (Figure 5). The selection and design of probes used in this study allowed for the assessment of the presence or absence of vector backbone sequences in the cotton GHB811 genome. Probes used in this study are summarized in Table 4 (above). Expected and obtained hybridization fragments are listed in Table 6 and Table 7. The hybridization results in the T1 and BC2F3 generations of cotton GHB811 are presented in Figure 16 to Figure 21.

Each membrane contained one negative control, which showed no hybridization with any of the probes used (Figure 16 to Figure 21, lane 5). Consequently, the absence of any background hybridization was demonstrated for all the probes used.

Similarly, each membrane contained a positive control. This positive control, consisting of pTSIH09 plasmid DNA, was digested with *Hin*cII and an equimolar amount was spiked in *Hin*cII digested gDNA from the non-GM counterpart. For each of the probes used, the expected fragments were detected (Figure 16 to Figure 21, lane 7), confirming that the applied experimental conditions allowed specific hybridization of the used probes with the target sequences.

Additionally, a supplementary positive control was used. This additional positive control, consisting of pTSIH09 plasmid DNA, was digested with *Hin*cII and a $1/10^{th}$ equimolar amount was spiked in *Hin*cII digested gDNA from the non-GM counterpart. Both positive controls, supplemented with one or $1/10^{th}$ equimolar amount of the *Hin*cII digested gDNA from the non-GM counterpart, showed the expected hybridization fragments after hybridization with the vector backbone probes (Figure 16 to Figure 21, lanes 6 and 7). This demonstrated that the hybridizations were performed in conditions allowing detection of the possible presence of vector backbone in cotton GHB811 genome. After hybridization with the T-DNA probe, the expected fragments were obtained in the positive control spiked with one equimolar amount of pTSIH09 digested plasmid DNA (data not shown). This demonstrated that the hybridizations were performed in conditions allowing detection of the presence of T-DNA sequences.

Hybridization of the digested cotton GHB811 gDNA samples with the vector backbone probes resulted in no hybridization fragments, as expected (Figure 16 to Figure 21, lanes 3 and 4). This demonstrated the absence of vector backbone sequences in cotton GHB811 gDNA. When hybridizing the same membranes with the T-DNA probe, all expected fragments were obtained (data not shown). This demonstrated that an ample amount of a sufficient quality of digested cotton GHB811 gDNA was loaded on the gels to enable detection of vector backbone sequences in cotton GHB811.

Table 6: Expected and obtained hybridization fragments determined for the vector backbone assessment in T1 generation of cotton GHB811

Sample	T-DNA or plasmid fragment sizes (bp)	Fragment description	Membrane M/THT063B/12				Membrane M/THT063B/13				Membrane M/THT063B/14			
			P010-1 Vector backbone probe		P009-1 T-DNA probe		P011-1 Vector backbone probe		P009-1 T-DNA probe		P012-1 Vector backbone probe		P009-1 T-DNA probe	
			Figure 18				Figure 19				Figure 20			
			Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
GHB811- *Bsp*HI	1365	5' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	8300*	3' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
GHB811- *Psi*I/*Sap*I	3423	5' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	1588	Internal fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	2644	3' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
non-GM counterpart (Coker 312) - *Bsp*HI	/	/	/	/	/	/	/	/	/	/	/	/	/	/
non-GM counterpart - *Hin*cII + 0.1 equimolar amount pTSIH09 - *Hin*cII	6341	positive control	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2476	positive control	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	No°
	1113	positive control	No	No	Yes	No°	No	No	Yes	No°	No	No	Yes	No°
	3169	positive control	Yes	Yes	Yes**(50)	No	Yes	Yes	Yes**(50)	No	No	No	Yes**(50)	No
non-GM counterpart - *Hin*cII + 1 equimolar amount pTSIH09 - *Hin*cII	6341	positive control	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2476	positive control	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	1113	positive control	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	3169	positive control	Yes	Yes	Yes**(50)	No	Yes	Yes	Yes**(50)	No	No	No	Yes**(50)	No

* Fragment sizes as determined in the detailed insert characterization and vector backbone assessment in this study

** Due to a small overlap with the probe, these fragments may not be visible. The size of the overlap is indicated between brackets.

° Although not all expected fragments of positive control containing 1/10th equimolar amount of pTSIH09 were obtained after hybridization with the T-DNA probe, all expected fragments were obtained after hybridization with the vector backbone probes. Therefore, hybridization conditions were appropriate to detect vector backbone sequence.

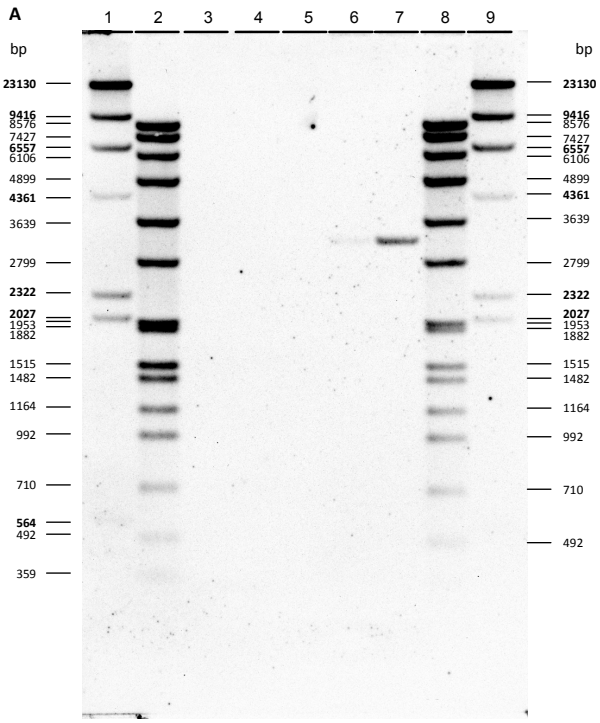
Table 7: Expected and obtained hybridization fragments determined for the vector backbone assessment in BC2F3 generation of cotton GHB811

Sample	T-DNA or plasmid fragment sizes (bp)	Fragment description	Membrane M/THT063B/23				Membrane M/THT063B/24				Membrane M/THT063B/25			
			P010-1 Vector backbone probe		P009-5 T-DNA probe		P011-2 Vector backbone probe		P009-2 T-DNA probe		P012-3 Vector backbone probe		P009-5 T-DNA probe	
			Figure 21				Figure 22				Figure 23			
			Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.	Exp.	Obt.
GHB811- *Bsp*HI	1365	5' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	8300*	3' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
GHB811- *Psi*I/*Sap*I	3423	5' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	1588	Internal fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	2644	3' integration fragment	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
non-GM counterpart (Coker 312) - *Bsp*HI	/	/	/	/	/	/	/	/	/	/	/	/	/	/
non-GM counterpart - *Hin*cII + 0.1 equimolar amount pTSIH09 - *Hin*cII	6341	positive control	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2476	positive control	No	No	Yes	Yes	No	No	Yes	No°	No	No	Yes	Yes
	1113	positive control	No	No	Yes	Yes	No	No	Yes	No°	No	No	Yes	Yes
	3169	positive control	Yes	Yes	Yes**(50)	No	Yes	Yes	Yes**(50)	No	No	No	Yes**(50)	No
non-GM counterpart - *Hin*cII + 1 equimolar amount pTSIH09 - *Hin*cII	6341	positive control	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2476	positive control	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	1113	positive control	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
	3169	positive control	Yes	Yes	Yes**(50)	No	Yes	Yes	Yes**(50)	No	No	No	Yes**(50)	No

* Fragment sizes as determined in the detailed insert characterization and vector backbone assessment in this study

** Due to a small overlap with the probe, these fragments may not be visible. The size of the overlap is indicated between brackets.

° Although not all expected fragments of positive control containing 1/10th equimolar amount of pTSIH09 were obtained after hybridization with the T-DNA probe, all expected fragments were obtained after hybridization with the vector backbone probes. Therefore, hybridization conditions were appropriate to detect vector backbone sequence.

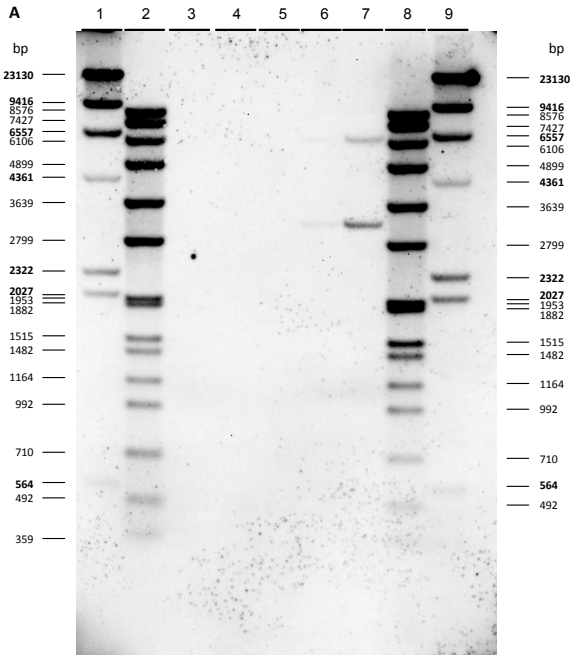


H1/THT063B/12-F6: Vector backbone probe P010-1

Figure 16: Hybridization performed with a vector backbone probe covering the *aadA* sequence (P010) to assess the vector backbone presence in the T1 generation of cotton GHB811

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with restriction enzymes *Bsp*HI and *Psi*I/*Sap*I and hybridized with a vector backbone probe (P010-1) and with the T-DNA probe (P009-1) (data not shown).

Lane 1: 3.5 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3.5 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3.5 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 4: 3.5 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 5: 3.5 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 6: 3.5 µg gDNA from non-GM counterpart - *Hin*cII digested + 1/10th of an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 7: 3.5 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 8: 3.5 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 9: 3.5 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

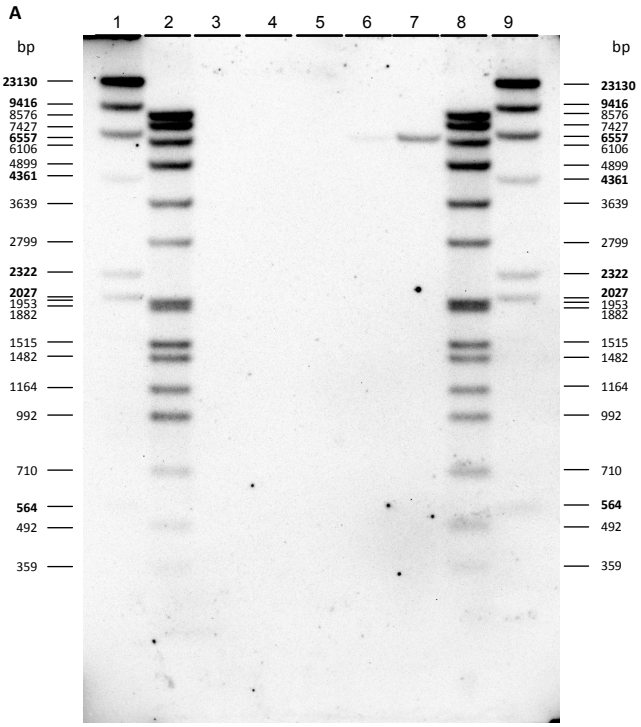


H1/THT063B/13-F4: Vector backbone probe P011-1

Figure 17: Hybridization performed with a vector backbone probe covering the ORI pVS1 region (P011) to assess the vector backbone presence in the T1 generation of cotton GHB811

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with restriction enzymes *Bsp*HI and *Psi*I/*Sap*I and hybridized with a vector backbone probe (P011-1) and with the T-DNA probe (P009-1) (data not shown).

Lane 1: 3.5 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3.5 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3.5 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 4: 3.5 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 5: 3.5 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 6: 3.5 µg gDNA from non-GM counterpart - *Hin*cII digested + 1/10th of an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 7: 3.5 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 8: 3.5 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 9: 3.5 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

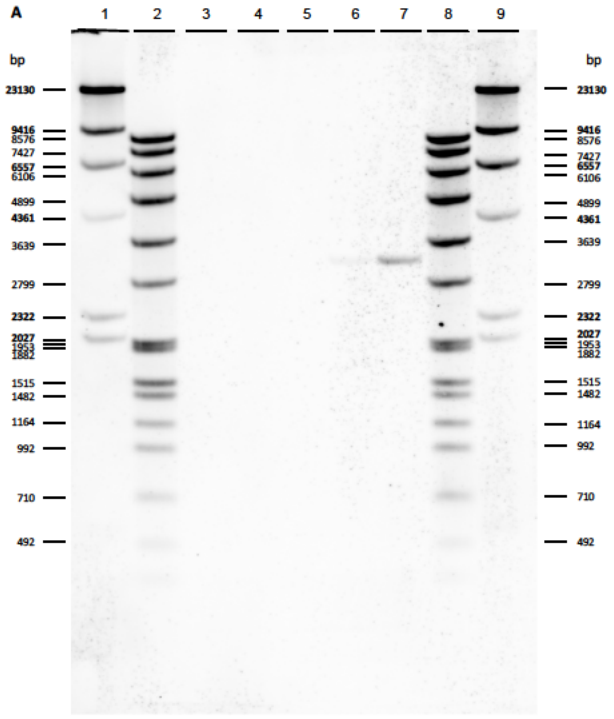


H1/THT063B/14-F4: Vector backbone probe P012-1

Figure 18: Hybridization performed with a vector backbone probe covering the ORI ColE1 region (P012) to assess the vector backbone presence in the T1 generation of cotton GHB811

gDNA was isolated from cotton GHB811 plants (T1 generation) and from the non-GM counterpart plants. The gDNA samples were digested with restriction enzymes *Bsp*HI and *Psi*I/*Sap*I and hybridized with a vector backbone probe (P012-1) and with the T-DNA probe (P009-1) (data not shown).

Lane 1: 3.5 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 3.5 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 3.5 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 4: 3.5 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 5: 3.5 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 6: 3.5 µg gDNA from non-GM counterpart - *Hin*cII digested + 1/10th of an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 7: 3.5 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 8: 3.5 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 9: 3.5 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

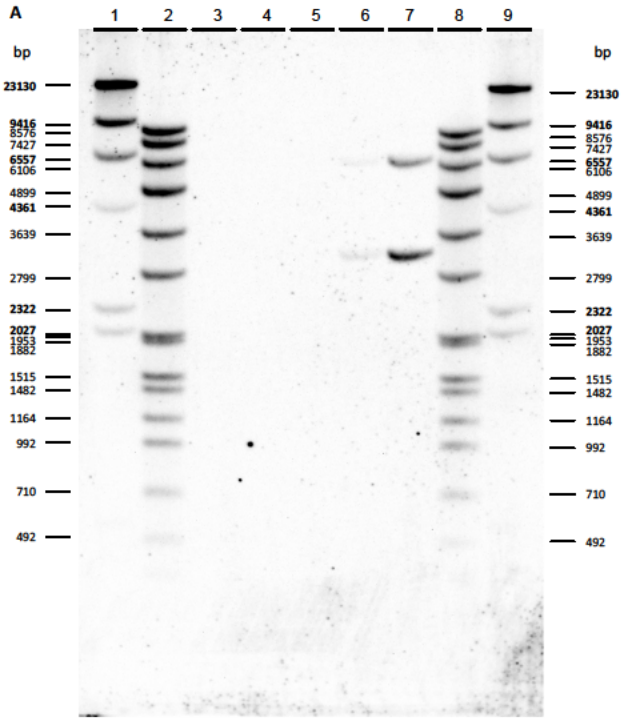


H1/THT063B/23-F1: Vector backbone probe P010-1

Figure 19: Hybridization performed with a vector backbone probe covering the *aadA* sequence (P010) to assess the vector backbone presence in the BC2F3 generation of cotton GHB811

gDNA was isolated from cotton GHB811 plants (BC2F3 generation) and from the non-GM counterpart plants. The gDNA samples were digested with restriction enzymes *Bsp*HI and *Psi*I/*Sap*I and hybridized with a vector backbone probe (P010-1) and with the T-DNA probe (P009-5) (data not shown).

Lane 1: 4 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 4 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 4 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 4: 4 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 5: 4 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 6: 4 µg gDNA from non-GM counterpart - *Hin*cII digested + 1/10th of an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 7: 4 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 8: 4 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 9: 4 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

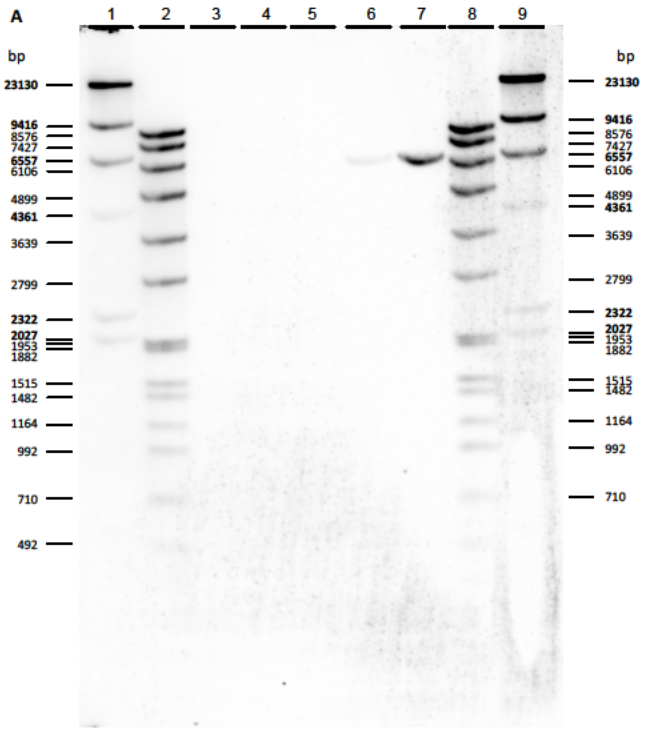


H2/THT063B/24-F4: Vector backbone probe P011-2

Figure 20: Hybridization performed with a vector backbone probe covering the ORI pVS1 region (P011) to assess the vector backbone presence in the BC2F3 generation of cotton GHB811

gDNA was isolated from cotton GHB811 plants (BC2F3 generation) and from the non-GM counterpart plants. The gDNA samples were digested with restriction enzymes *Bsp*HI and *Psi*I/*Sap*I and hybridized with a vector backbone probe (P011-2) and with the T-DNA probe (P009-2) (data not shown).

Lane 1: 4 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 4 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 4 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 4: 4 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 5: 4 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 6: 4 µg gDNA from non-GM counterpart - *Hin*cII digested + 1/10th of an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 7: 4 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 8: 4 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 9: 4 µg gDNA from non-GM counterpart - *Hin*dIII digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)



H1/THT063B/25-F4: Vector backbone probe P012-3

Figure 21: Hybridization performed with a vector backbone probe covering the ORI ColE1 region (P012) to assess the vector backbone presence in the BC2F3 generation of cotton GHB811

gDNA was isolated from cotton GHB811 plants (BC2F3 generation) and from the non-GM counterpart plants. The gDNA samples were digested with restriction enzymes *Bsp*HI and *Psi*I/*Sap*I and hybridized with a vector backbone probe (P012-3) and with the T-DNA probe (P009-5) (data not shown).

Lane 1: 4 µg gDNA from non-GM counterpart - *Hind*III digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)
Lane 2: 4 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 3: 4 µg gDNA from cotton GHB811 - *Bsp*HI digested
Lane 4: 4 µg gDNA from cotton GHB811 - *Psi*I/*Sap*I digested
Lane 5: 4 µg gDNA from non-GM counterpart - *Bsp*HI digested
Lane 6: 4 µg gDNA from non-GM counterpart - *Hin*cII digested + 1/10th of an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 7: 4 µg gDNA from non-GM counterpart - *Hin*cII digested + an equimolar amount of pTSIH09 - *Hin*cII digested
Lane 8: 4 µg gDNA from non-GM counterpart - *Eco*RI digested + 10 ng DNA Molecular Weight Marker VII, DIG-labeled (Roche)
Lane 9: 4 µg gDNA from non-GM counterpart - *Hind*III digested + 10 ng DNA Molecular Weight Marker II, DIG-labeled (Roche)

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

As detailed above in Section A.3 (c)(i) above and below in Section A.3 (c)(iii), Southern blot analysis and full DNA sequencing of the GHB811 cotton transgenic locus revealed that the inserted genetic material consists of one complete copy of the T-DNA that corresponds to the transforming plasmid. The arrangement of the GHB811 cotton transgenic locus is shown in Figure 22 below. The Southern blot analysis is detailed in Back (2016; M-572036-01; Node A.3 (c), (i), (CCI)), and sequencing of the transgenic locus and corresponding insertion locus is detailed in , M-533573-01, Node A.3 (c), (iii) (CCI)).

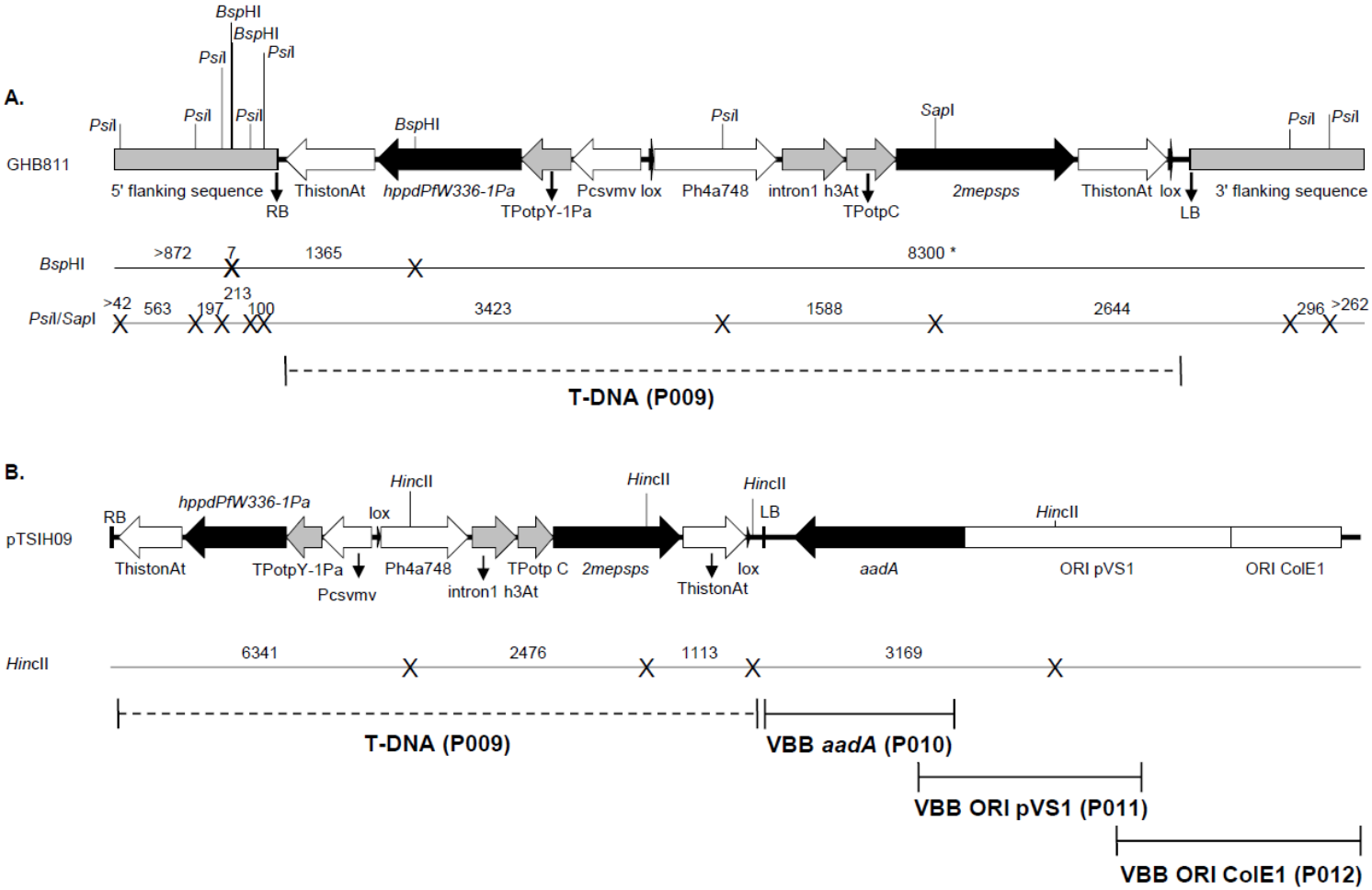


Figure 22 Schematic drawing of cotton GHB811 and pTSIH09 with indication of the restriction sites and position of the probes used for vector backbone assessment. Upper panel A: transgenic locus cotton GHB811. Lower panel B: pTSIH09

(iii) Full DNA sequence of each insertion site, including junction regions with the host DNA;

The DNA sequence of the cotton GHB811 transgenic locus and the corresponding insertion locus was determined [redacted] 2015; M-533573-01; Node A.3 (c), (i) (CCI)).

Three overlapping PCR fragments were prepared for the determination of the GHB811 transgenic locus, using GHB811 gDNA as a template. To determine the GHB811 insertion locus, one fragment was amplified from gDNA extracted from the non-GM counterpart (Table 8). Sanger sequencing was performed using the ABI PRISM® BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems).

Table 8: Overview of the sequencing fragments prepared

	Fragment ID	Template DNA	Primer pair	Length of final consensus sequence (bp)
GHB811 transgenic locus	FR-THT073-B-01	GHB811 gDNA	GLPA486 GLPA311	9320
	FR-THT073-B-01b(**)		GLPA696 GLPA311	
	FR-THT073-B-02		GLPA457 GLPA076	
	FR-THT073-B-03		GLPA485 GLPA487	
GHB811 insertion locus	FR-THT073-B-04	Coker 312 gDNA	GLPA486 GLPA487	2526
	FR-THT073-B-04b(**)		GLPA696 GLPA487	

(**) This fragment was amplified for a second time under different PCR conditions

The obtained consensus sequences of the transgenic and insertion loci were annotated by pairwise alignments using the Clone Manager software (Sci-Ed Central). An alignment between the GHB811 transgenic locus and the GHB811 insertion locus sequence was made to identify sequence regions of cotton origin within the GHB811 transgenic locus as well as the target site deletion (TSD) within the GHB811 insertion locus. The consensus sequence of the GHB811 transgenic locus was compared with the pTSIH09 sequence to identify the T-DNA region. Further sequence annotation within the T-DNA was performed by comparing the GHB811 transgenic locus sequence with each feature of the pTSIH09 T-DNA region.

Pairwise alignment between the GHB811 transgenic sequence and the GHB811 insertion locus sequence identified two regions sharing 100% pairwise sequence identity. These two regions are listed in the overview below.

Region of homology	% matches	Length (bp)	GHB811 transgenic locus		GHB811 insertion locus	
			start	end	Start	end
Region A: 5' flanking sequence	100	1217	bp 1	bp 1217	bp 1	bp 1217
Region B: 3' flanking sequence	100	1296	bp 8033	bp 9328	bp 1231	bp 2526

Homology region A was identified as 5' flanking sequence on the GHB811 transgenic locus sequence and the GHB811 insertion locus sequence. Homology region B was identified as 3' flanking sequence on the GHB811 transgenic locus sequence and the GHB811 insertion locus sequence.

In the GHB811 insertion locus sequence, 13 bp were observed which are not present in the GHB811 transgenic locus. These base pairs were deleted during the transformation process and are referred to as TSD.

Pairwise alignment between the GHB811 transgenic sequence and the pTSIH09 plasmid sequence identified three regions sharing 100% pairwise sequence identity which are listed below.

Region of homology:	% matches	Length (bp)	GHB811 transgenic locus		pTSIH09	
			start	end	start	end
Region A: T-DNA	100	6817	bp 1218	bp 8034	bp 24	bp 6840
Region B: ThistonAt	100	667	bp 7193	bp 7859	bp 749	bp 83
Region C: ThistonAt	100	667	bp 1277	bp 1943	bp 6665	bp 5999
Region D lox	100	34	bp 7864	bp 7897	bp 2796	bp 2829
Region E lox	100	34	bp 3390	bp 4023	bp 6670	bp 6703

Homology region A on the transgenic sequence which is 100 % identical to the T-DNA region of pTSIH09 was identified as T-DNA. The different features of the T-DNA were annotated as well. The four additional homologies result from the presence of two "ThistonAt" and two "lox" features within the T-DNA region.

Two base pairs at the 3' end of the T-DNA region (bp 8033 to bp 8034) were identical to both the plasmid sequence pTSIH09 and the insertion locus. These base pairs were annotated as 3' flanking sequence.

A schematic representation of the GHB811 transgenic locus in relation to the pTSIH09 plasmid is provided in Figure 23.

The results demonstrated that upon transformation, 13 bp from the GHB811 insertion locus were replaced by 6815 bp of inserted sequences. The flanking sequences obtained at the transgenic locus were identical to the homologous sequences obtained from the insertion locus. This demonstrates that the cotton GHB811 flanking sequences are of cotton origin within its original genomic organization. Annotation of the inserted sequences in the GHB811 transgenic locus sequence demonstrated that it corresponds to the complete T-DNA region of pTSIH09 and did not indicate any T-DNA rearrangements.

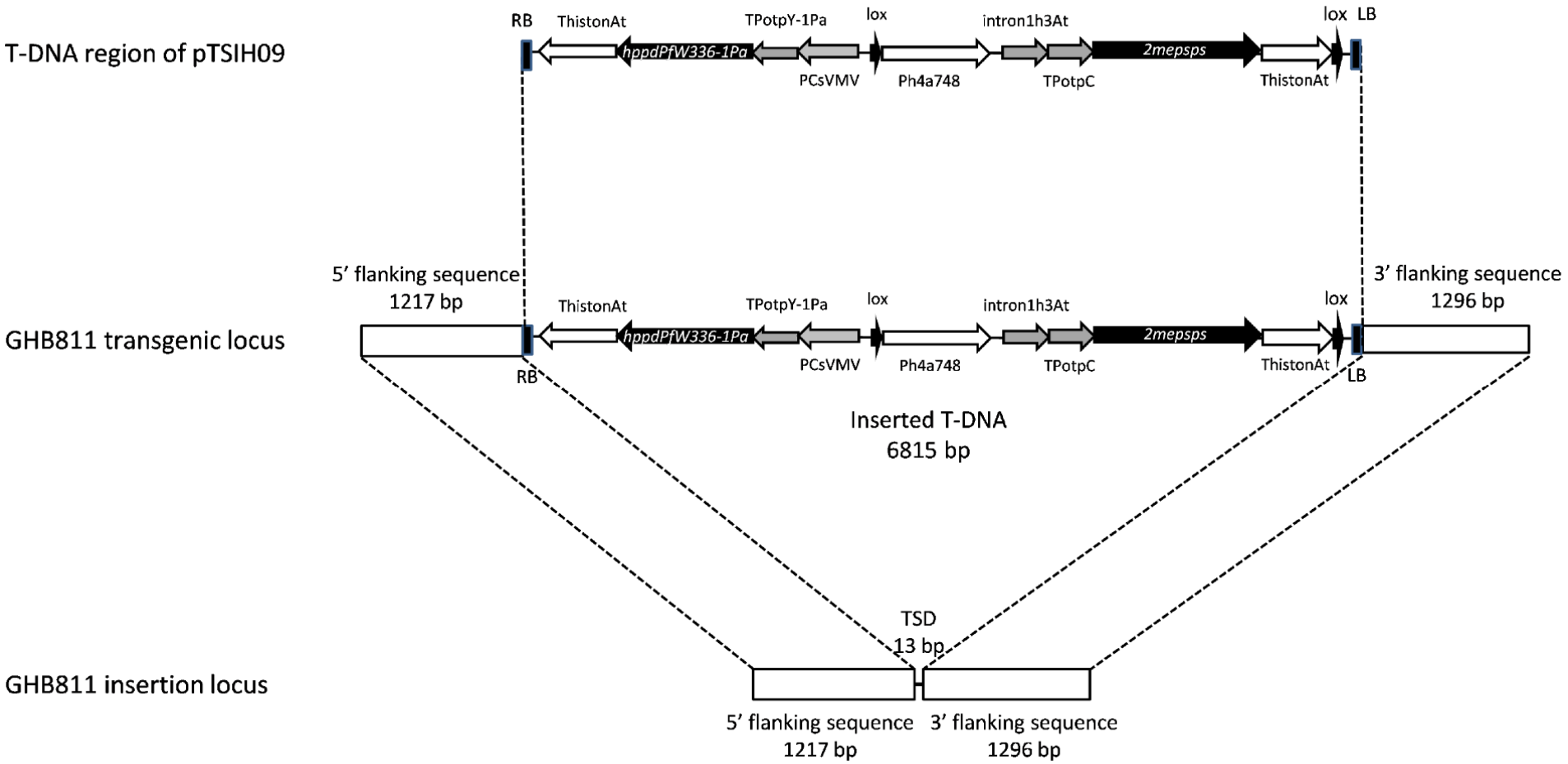


Figure 23 Schematic drawing of the GHB811 transgenic locus in relation to the GHB811 insertion locus and the T-DNA of transforming plasmid pTSIH09

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

The organisation of the GHB811 transgenic locus within the *G. hirsutum* genome, as confirmed by Southern blot M-572036-01; Node A.3 (c), (i), (CCI)) and DNA sequence analysis ; M-533573-01; Node A.3 (c), (iii) (CCI)), and described above in Sections A.3(c)(i-iii), is shown in Figure 24 below.

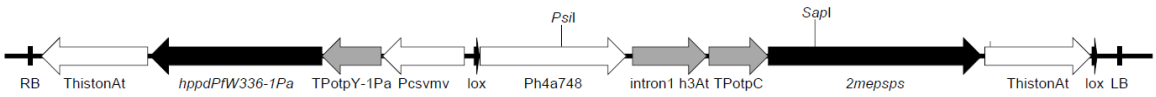


Figure 24 Organisation of the GHB811 *G. hirsutum* transformation event in the *G. hirsutum* genome demonstrated by Southern blot and DNA sequence analyses

(v) Details of an analysis of insert and junction regions for the occurrence of any open reading frames (ORFs).

A bioinformatics analysis was performed on the GHB811 cotton insertion locus sequence, to identify the position of the insertion locus in the genome and to determine whether regulatory sequences or endogenous cotton genes were interrupted upon the insertion of T-DNA sequences (M-581222-01; Node A.3 (c), (v) (CCI)).

The GHB811 transgenic locus, containing the inserted DNA together with the 5' and 3' flanking sequences, was used as query sequence. The Basic Local Alignment Search Tool (BLAST) searches were performed to find the location of the GHB811 cotton insertion locus in the cotton genome and to search for sequence similarities with known genes and proteins. BLAST analysis demonstrated that the insertion locus sequence originates from cotton chromosome A05.

The similarities between the GHB811 cotton insertion locus and sequences within the nucleotide collection and the Expressed Sequence Tag (EST) databases were identified using the BLASTn tool available on the NCBI website. In addition, A BLASTx search of the GHB811 cotton insertion locus sequence against the NCBI non-redundant protein database was performed. The results indicate that it is unlikely that the insertion of T-DNA sequences in the GHB811 cotton insertion locus interrupted or altered the transcriptional or translational activity of endogenous cotton genes.

A bioinformatics analysis was performed on the transgenic locus sequence of the GHB811 cotton to identify open reading frames (ORF) ; M-575144-01; Node A.3 (c), (v) (CCI)).

The GHB811 transgenic locus, containing the inserted DNA together with the 5' and 3' flanking sequences, was used as query sequence. The GetORF search program was used to identify all ORF crossing a junction or overlapping the inserted DNA, between two translation stop codons, with a minimum size coding for 3 amino acids. This search identified 549 ORF.

In the next step, the translated amino acid sequences from the identified ORF with a minimum size of 30 amino acids were used as query sequences in homology searches to

known allergens and toxins. After elimination of duplicates, they represented 126 unique sequences.

Two *in silico* approaches were used to evaluate the potential amino acid sequence identity with known allergens contained in the public allergen database AllergenOnline (www.allergenonline.org):

- An 8-mer search was carried out to identify any short sequences of 8 amino acids or longer that share 100% identity to an allergenic protein. This search was performed using SeqMatchAll from the EMBOSS suite, which compared each ORF sequence with all known allergens present in the allergen database.
- An overall identity search was carried out by using FASTA algorithm, which compared each complete query sequence with all protein sequences present in the AllergenOnline database. The scoring matrix was BLOSUM50. An E-value threshold of 1 was used. The criterion indicating potential relevant identity to an allergen was ≥35% identity over at least 80 amino acids for sequences of ≥80 amino acids, or ≥35% recalculated over a hypothetical 80 amino acid window for sequences of <80 amino acids.

In addition, each query sequence was evaluated for potential identity with known toxins. An overall identity search was carried out by using FASTA algorithm with all protein sequences present in the NCBI non-redundant database, using the BLOSUM50 scoring matrix. An E-value threshold of 0.1 was used for pre-selecting the most identical proteins. The biological relevance of the matches was further assessed. Biologically relevant matches provide insight on the familiarity and potential toxic properties of the potential polypeptide.

The 8-mer search showed no 100% identity with known allergenic proteins. The overall search showed no biologically relevant identity between the query sequences and any known allergenic proteins.

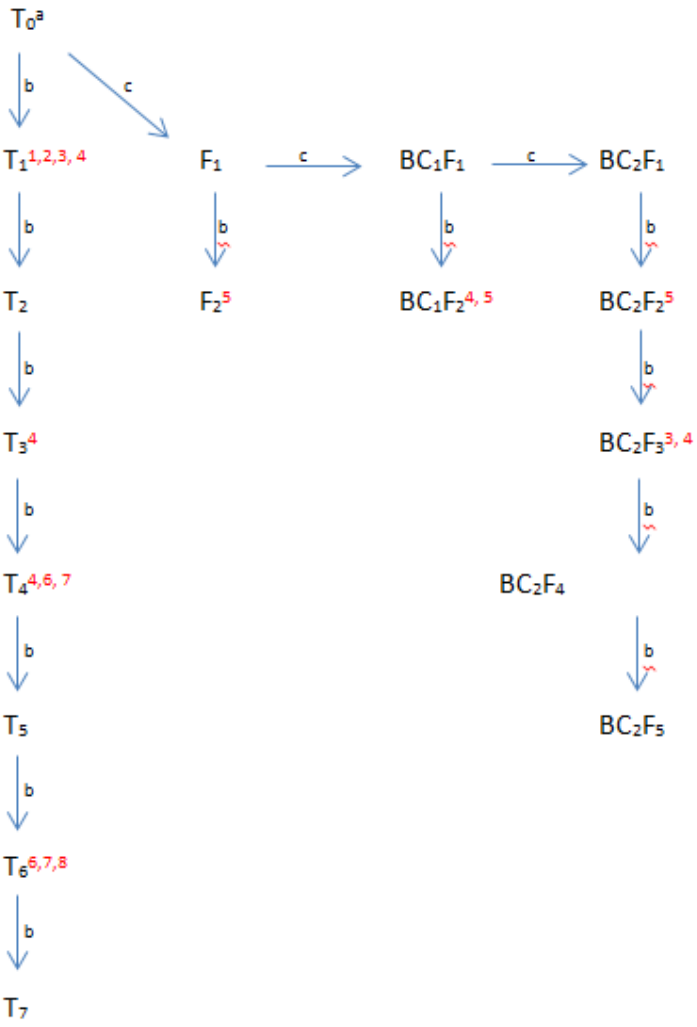
In addition, no biologically relevant identities were found with any toxic protein from the NCBI non-redundant database.

In conclusion, there are neither allergenic nor toxicological *in silico* findings associated with the potential ORF polypeptides.

(d) A description of how the line or strain from which food is derived was obtained from the original transformant (i.e. provide a family tree or describe the breeding process) including which generations have been used for each study.

Following *Agrobacterium*-mediated transformation of the conventional breeding line, Coker312 resulting in event GHB811, T_0 plants were treated with tembotrione (HPPD-inhibitor herbicides) to select for the expression of the *hppdPfw336-1Pa* genes. The surviving plants were then self-pollinated to generate T_1 seed. All subsequent T_2 to T_7 generations were produced through self-pollination. A subsample of the T_1 and T_2 plants were sprayed with glyphosate to ensure expression of the *2mepsps* gene at those generations. In the T_3 through T_7 generations which were grown in the field, each selfed generation was sprayed with glyphosate to ensure the expression of the *2mepsps* gene. In the development of GHB811 cotton varieties, T_0 plants were back-crossed into a conventional commercial cotton line.

The breeding program for the development of event GHB811 and its introgression into commercial cotton germplasm is demonstrated in Figure 25 below. Table 9 describes the GHB811 generations used for analysis and the associated reports describing these studies.



a: Coker312 was used for transformation
b: selfing
c: crossing with Stoneville 457 variety

Figure 25. Pedigree of GHB811

Table 9. GHB811 cotton generations used for analysis

No. in Tree	Experiment	Generation(s)	Comparator	Dart No.
1	DNA sequencing of insert and flanking region	T1	Coker312	M-533573-01
2	Insert Characterization by Southern Analysis	T1	Coker312	M-572036-01
3	Absence of Vector Backbone by Southern Analysis	T1, BC2F3	Coker312	M-572036-01
4	Structural Stability by Southern Analysis	T1, T3, T4, BC1F2, BC2F3	Coker312	M-548778-01
5	Inheritance of the Insert	F2, BC1F2, BC2F2	None	M-547925-01
6	Agronomic and phenotypic Analysis	T4, T6	Coker312	N/A
7	Composition Analysis	T4, T6	Coker312	M-566678-01
8	Protein Expression Analysis	T6	None	M-574232-01

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

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

The structural stability of GHB811 cotton was investigated by performing a Southern blot analysis on individual plants from five different generations (T1, T3, T4, BC1F2 and BC2F3 generations) M-548778-01; Node A.3 (e), (i) (CCI)).

Seeds from five different seed lots, each corresponding to a different generation, were used to produce cotton GHB811 leaf material. The identity of the produced plant material was confirmed.

Non-genetically modified (non-GM) cotton variety Coker 312 (non-GM counterpart) was used as a negative control. The transforming plasmid of cotton GHB811, pTSIH09 was used as a positive control.

The non-GM counterpart-derived gDNA and a set of individual gDNA samples from GHB811 cotton were digested using the restriction enzyme combination *Psi*I/*Sap*I. Additional restriction digests of the non-GM counterpart gDNA were prepared using respectively the *Hinc*II and the *Eco*RI restriction enzymes. Plasmid DNA of pTSIH09 was digested using the *Hinc*II restriction enzyme, as recommended by the manufacturer (New England BioLabs).

Table 10: Information on the used probe

Probe ID	Description	Primer pair	Primer sequence (5' → 3')	Primer position in pTSIH09 (bp)	Size probe (bp)
P009	T-DNA probe	GLPA467		79 → 98	6700
		GLPA468		6778 → 6759	

Hybridization was performed with the T-DNA probe P009 (Table 10). A schematic overview of the GHB811 transgenic locus, with indication of the restriction enzymes, the T-DNA probe used and the expected fragments is presented in Figure 26. The hybridization results are presented in Figure 27 to Figure 31. A summary of the results obtained is presented in Table 11 to Table 15.

Each membrane used for the analysis contained one negative control which was never shown to hybridize with the T-DNA probe. This confirmed the absence of any background hybridization. Similarly, each reported membrane contained one positive control. For all hybridizations, the expected fragments were detected for the positive control, indicating that the conditions of the Southern blot experiments allowed specific hybridization of the used probes with the target sequences.

Genomic DNA from individual GHB811 cotton plants was digested with restriction enzyme combination *Psi*I/*Sap*I and hybridized to the T-DNA probe. For all individual plants from the T1, T3, T4, BC1F2 and BC2F3 generation, all expected three fragments (3300 bp, 1588 bp, and 2600 bp) were obtained (Figure 27 to Figure 31 and Table 11 to Table 15).

Taken together, all obtained results demonstrate the structural stability of GHB811 cotton in the T1, T3, T4, BC1F2 and BC2F3 generations.

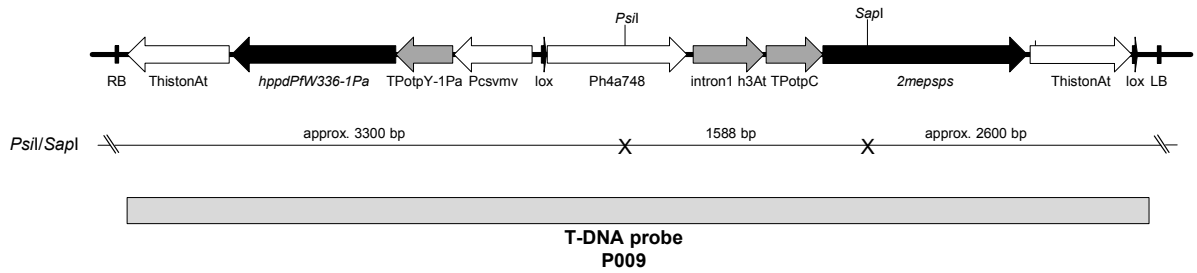


Figure 26: Schematic overview of the GHB811 cotton transgenic locus with indication of the restriction sites, the probe used and the expected fragment sizes in bp (based on the corresponding detailed insert characterization study)

Table 11: Stability of cotton GHB811 in the individual plants of the T1 generation - Expected and obtained hybridization fragments (Figure 27)

Sample	Reference to lane N° in Figure 27	Fragment size (bp)	Fragment description	Probe P009 T-DNA	
				Exp.	Obt.
18 samples Cotton GHB811, T1 generation – *Psi*I/*Sap*I	Lane 2 to 19	approx. 3300 *	5' integration fragment	Yes	Yes
		1588*	Internal fragment	Yes	Yes
		approx. 2600 *	3' integration fragment	Yes	Yes
non-GM counterpart - *Psi*I/*Sap*I	Lane 20	/	Negative control	/	/
non-GM counterpart – *Hinc*II + equimolar amount pTSIH09 - *Hinc*II	Lane 21	1113	Positive control	Yes	Yes
		2476		Yes	Yes
		3169		Yes**	No
		6341		Yes	Yes

* Results determined in the detailed insert characterization and confirmation of absence of vector backbone study of cotton GHB811
** Due to the small overlap of this fragment with the T-DNA probe, the likelihood to visualize this fragment is very low

Table 12: Stability of cotton GHB811 in the individual plants of the T3 generation - Expected and obtained hybridization fragments (28)

Sample	Reference to lane N° in	Fragment size (bp)	Fragment description	Probe P009 T-DNA	
				Exp.	Obt.
16 samples Cotton GHB811, T3 generation – *Psi*I/*Sap*I	Lane 2 to 17	approx. 3300 *	5' integration fragment	Yes	Yes
		1588*	Internal fragment	Yes	Yes
		approx. 2600 *	3' integration fragment	Yes	Yes
non-GM counterpart - *Psi*I/*Sap*I	Lane 18	/	Negative control	/	/
non-GM counterpart – *Hinc*II + equimolar amount pTSIH09 - *Hinc*II	Lane 19	1113	Positive control	Yes	Yes
		2476		Yes	Yes
		3169		Yes**	No
		6341		Yes	Yes

* Results determined in the detailed insert characterization and confirmation of absence of vector backbone study of cotton GHB811
** Due to the small overlap of this fragment with the T-DNA probe, the likelihood to visualize this fragment is very low

Table 13: Stability of cotton GHB811 in the individual plants of the T4 generation - Expected and obtained hybridization fragments (Figure 29)

Sample	Reference to Lane N° in Figure 29	Fragment size (bp)	Fragment description	Probe P009 T-DNA	
				Exp.	Obt.
15 samples Cotton GHB811, T4 generation – *Psi*I/*Sap*I	Lane 2 to 16	approx. 3300 *	5' integration fragment	Yes	Yes
		1588*	Internal fragment	Yes	Yes
		approx. 2600 *	3' integration fragment	Yes	Yes
non-GM counterpart - *Psi*I/*Sap*I	Lane 17	/	Negative control	/	/
non-GM counterpart – *Hinc*II + equimolar amount pTSIH09 - *Hinc*II	Lane 18	1113	Positive control	Yes	Yes
		2476		Yes	Yes
		3169		Yes**	Very faint
		6341		Yes	Yes

* Results determined in the detailed insert characterization and confirmation of absence of vector backbone study of cotton GHB811
** Due to the small overlap of this fragment with the T-DNA probe, the likelihood to visualize this fragment is very low

Table 14: Stability of cotton GHB811 in the individual plants of the BC1F2 generation - Expected and obtained hybridization fragments (Figure 30)

Sample	Reference to lane N° in Figure 30	Fragment size (bp)	Fragment description	Probe P009 T-DNA	
				Exp.	Obt.
14 samples Cotton GHB811, BC1F2 generation – *Psi*I/*Sap*I	Lane 2 to 15	approx. 3300 *	5' integration fragment	Yes	Yes
		1588*	Internal fragment	Yes	Yes
		approx. 2600 *	3' integration fragment	Yes	Yes
non-GM counterpart - *Psi*I/*Sap*I	Lane 16	/	Negative control	/	/
non-GM counterpart – *Hinc*II + equimolar amount pTSIH09 - *Hinc*II	Lane 17	1113	Positive control	Yes	Yes
		2476		Yes	Yes
		3169		Yes**	No
		6341		Yes	Yes

* Results determined in the detailed insert characterization and confirmation of absence of vector backbone study of cotton GHB811
** Due to the small overlap of this fragment with the T-DNA probe, the likelihood to visualize this fragment is very low

Table 15: Stability of cotton GHB811 in the individual plants of the BC2F3 generation - Expected and obtained hybridization fragments (Figure 31)

Sample	Reference to lane N° in Figure 31	Fragment size (bp)	Fragment description	Probe P009 T-DNA	
				Exp.	Obt.
15 samples Cotton GHB811, BC2F3 generation – *Psi*I/*Sap*I	Lane 2 to 16	approx. 3300 *	5' integration fragment	Yes	Yes
		1588*	Internal fragment	Yes	Yes
		approx. 2600 *	3' integration fragment	Yes	Yes
non-GM counterpart - *Psi*I/*Sap*I	Lane 17	/	Negative control	/	/
non-GM counterpart – *Hinc*II + equimolar amount pTSIH09 - *Hinc*II	Lane 18	1113	Positive control	Yes	Yes
		2476		Yes	Yes
		3169		Yes**	No
		6341		Yes	Yes

* Results determined in the detailed insert characterization and confirmation of absence of vector backbone study of cotton GHB811
** Due to the small overlap of this fragment with the T-DNA probe, the likelihood to visualize this fragment is very low

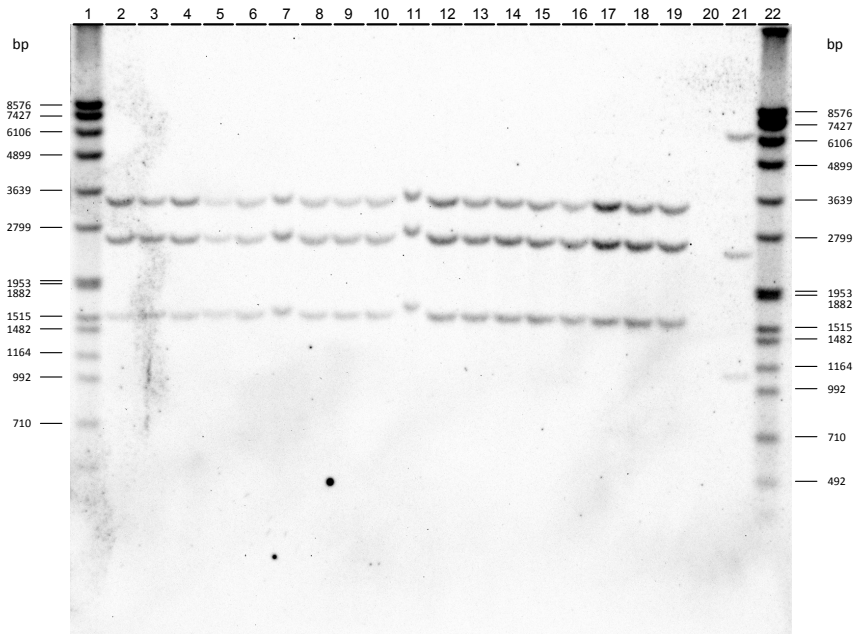


Figure 27: Southern blot analysis of cotton GHB811 – Hybridization performed with the T-DNA probe to assess structural stability of the individual plants of the T1 generation

Digital image ID: H1/THT068A/05-F3

Genomic DNA was isolated from individual cotton GHB811 plants of the T1 generation and from the non-GM counterpart. The gDNA samples were digested with the restriction enzyme combination *Psi*I/*Sap*I and hybridized with a probe corresponding to the cotton GHB811 T-DNA region (P009-5).

Lane 1: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested
Lane 2 to 11: 5 µg gDNA of individual hemizygous samples of cotton GHB811 of the T1 generation (one copy of GHB811) – *Psi*I/*Sap*I digested
Lane 12 to 19: 5 µg gDNA of individual homozygous samples of cotton GHB811 of the T1 generation (two copies of GHB811) – *Psi*I/*Sap*I digested
Lane 20: 5 µg gDNA of the non-GM counterpart – *Psi*I/*Sap*I digested (negative control)
Lane 21: 5 µg gDNA of the non-GM counterpart – *Hin*cII digested + an equimolar amount of plasmid pTSIH09 – *Hin*cII digested (positive control)
Lane 22: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested

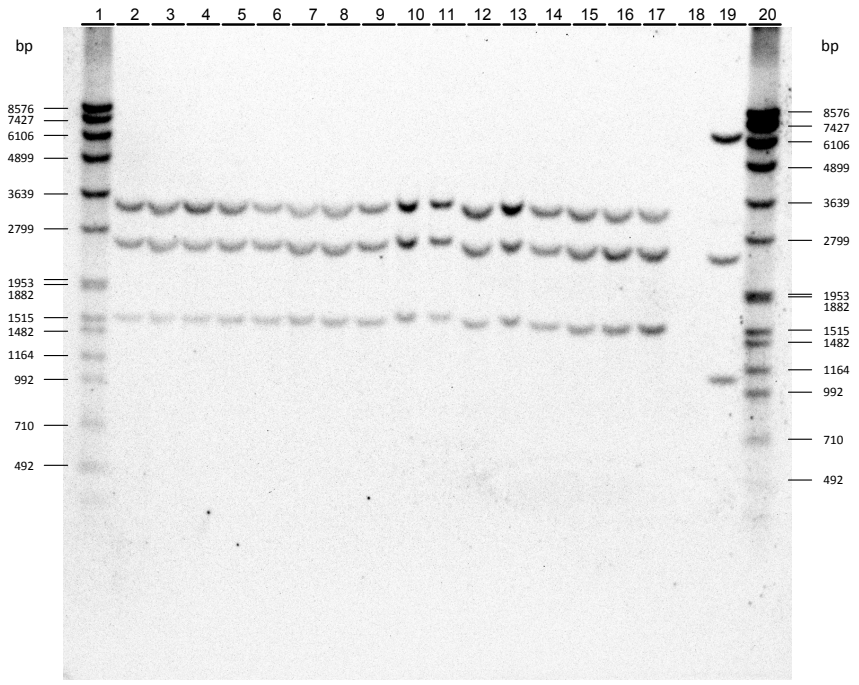


Figure 28: Southern blot analysis of cotton GHB811 – Hybridization performed with the T-DNA probe to assess structural stability of the individual plants of the T3 generation

Digital image ID: H1/THT068A/04-F3

Genomic DNA was isolated from individual cotton GHB811 plants of the T3 generation and from the non-GM counterpart. The gDNA samples were digested with the restriction enzyme combination *Psi*I/*Sap*I and hybridized with a probe corresponding to the cotton GHB811 T-DNA region (P009-5).

Lane 1: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 2.5 µg of gDNA of the non-GM counterpart – *Eco*RI digested
Lane 2 to 17: 2.5 µg gDNA of individual homozygous samples of cotton GHB811 of the T3 generation – *Psi*I/*Sap*I digested
Lane 18: 2.5 µg gDNA of the non-GM counterpart – *Psi*I/*Sap*I digested (negative control)
Lane 19: 2.5 µg gDNA of the non-GM counterpart – *Hin*cII digested + an equimolar amount of plasmid pTSIH09 – *Hin*cII digested (positive control)
Lane 20: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 2.5 µg of gDNA of the non-GM counterpart – *Eco*RI digested

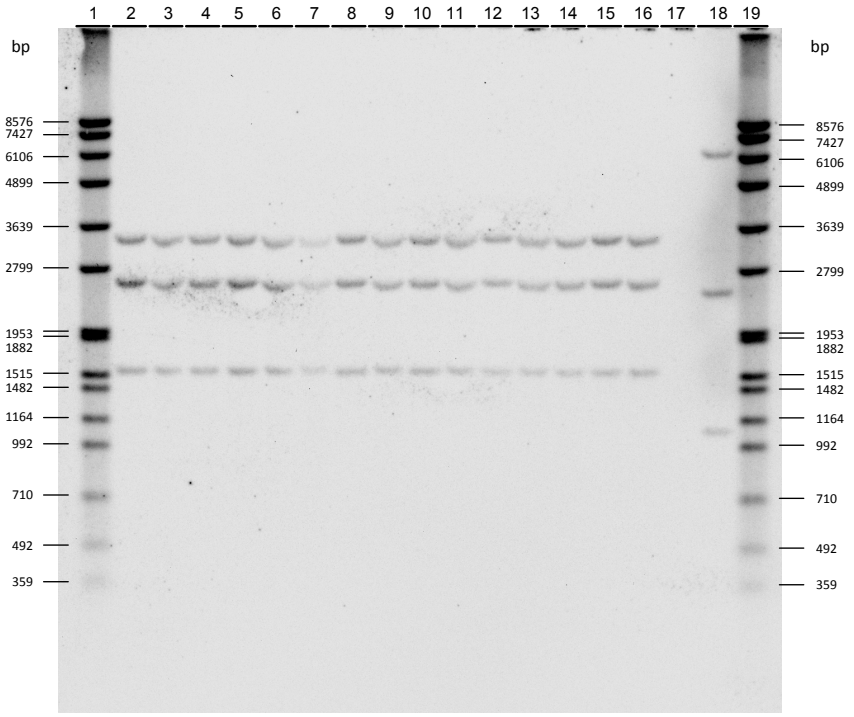


Figure 29: Southern blot analysis of cotton GHB811 – Hybridization performed with the T-DNA probe to assess structural stability of the individual plants of the T4 generation

Digital image ID: H1/THT068A/01-F3

Genomic DNA was isolated from individual cotton GHB811 plants of the T4 generation and from the non-GM counterpart. The gDNA samples were digested with the restriction enzyme combination *Psi*I/*Sap*I and hybridized with a probe corresponding to the cotton GHB811 T-DNA region (P009-5).

Lane 1: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested
Lane 2 to 16: 5 µg gDNA of individual homozygous samples of cotton GHB811 of the T4 generation – *Psi*I/*Sap*I digested
Lane 17: 5 µg gDNA of the non-GM counterpart – *Psi*I/*Sap*I digested (negative control)
Lane 18: 5 µg gDNA of the non-GM counterpart – *Hin*cII digested + an equimolar amount of plasmid pTSIH09 – *Hin*cII digested (positive control)
Lane 19: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested

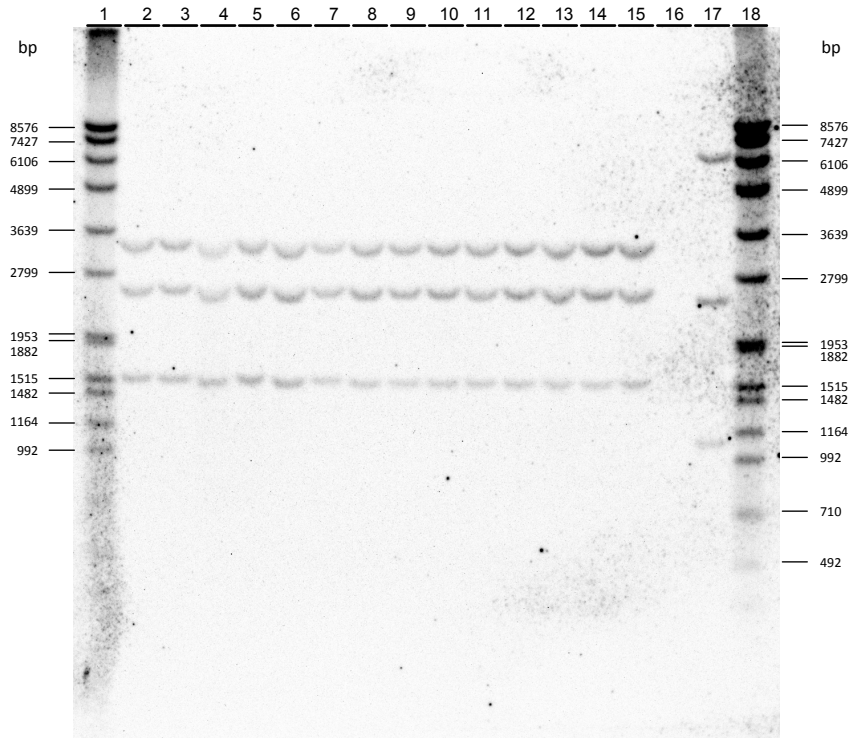


Figure 30: Southern blot analysis of cotton GHB811 – Hybridization performed with the T-DNA probe to assess structural stability of the individual plants of the BC1F2 generation

Digital image ID: H2/THT068A/03-F4

Genomic DNA was isolated from individual cotton GHB811 plants of the BC1F2 generation and from the non-GM counterpart. The gDNA samples were digested with the restriction enzyme combination *Psi*I/*Sap*I and hybridized with a probe corresponding to the cotton GHB811 T-DNA region (P009-5).

Lane 1: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested
Lane 2 to 15: 5 µg gDNA of individual hemizygous samples of cotton GHB811 of the BC1F2 generation – *Psi*I/*Sap*I digested
Lane 16: 5 µg gDNA of the non-GM counterpart – *Psi*I/*Sap*I digested (negative control)
Lane 17: 5 µg gDNA of the non-GM counterpart – *Hin*cII digested + an equimolar amount of plasmid pTSIH09 – *Hin*cII digested (positive control)
Lane 18: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested

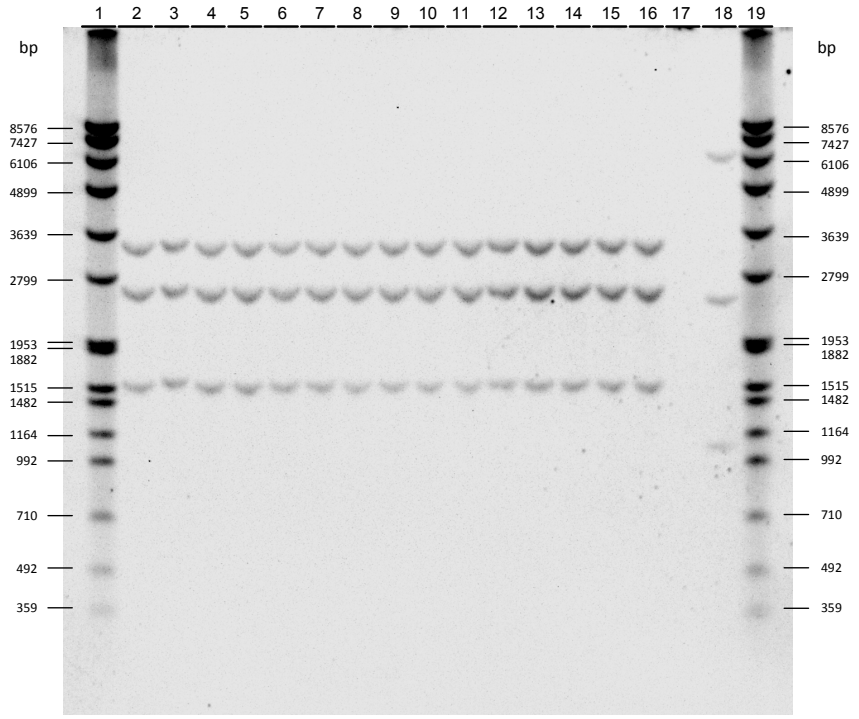


Figure 31: Southern blot analysis of cotton GHB811 – Hybridization performed with the T-DNA probe to assess structural stability of the individual plants of the BC2F3 generation

Digital image ID: H1/THT068A/02-F1

Genomic DNA was isolated from individual cotton GHB811 plants of the BC2F3 generation and from the non-GM counterpart. The gDNA samples were digested with the restriction enzyme combination *Psi*I/*Sap*I and hybridized with a probe corresponding to the cotton GHB811 T-DNA region (P009-5).

Lane 1: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested
Lane 2 to 16: 5 µg gDNA of individual homozygous samples of cotton GHB811 of the BC2F3 generation – *Psi*I/*Sap*I digested
Lane 17: 5 µg gDNA of the non-GM counterpart – *Psi*I/*Sap*I digested (negative control)
Lane 18: 5 µg gDNA of the non-GM counterpart – *Hin*cII digested + an equimolar amount of plasmid pTSIH09 – *Hin*cII digested (positive control)
Lane 19: 10 ng DNA Molecular Weight marker VII, DIG-labeled + 5 µg of gDNA of the non-GM counterpart – *Eco*RI digested

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

Section A.3(e) above details the GHB811 cotton breeding program, and section A.3(f)(i) (directly above) details experiments undertaken to test the structural stability of the GHB811 cotton transgenic locus over multiple generations. The inheritance pattern of the GHB811 cotton transgenic locus was tested over multiple generations. In addition, expression of the phenotype was tested across different environments.

Inheritance pattern of the GHB811 cotton transgenic locus

Genomic DNA from individual plants of three GHB811 cotton generations (F2, BC1F2, and BC2F2) was tested for the genotype of *hppdPfW336-1Pa* and *2mepsps* genes by polymerase chain reaction (PCR) analysis M-547925-01; Node A.3 (e), (ii)). The results from PCR analysis were used to calculate the segregation ratios of the genes

contained within the GHB811 insert. Chi-square analysis of the segregation data for three generations was performed to test the hypothesis that the GHB811 cotton insert is inherited in a manner that is predictable according to Mendelian principles and is consistent with insertion into a single chromosomal locus within the cotton nuclear genome.

Plant samples were analyzed using gene-specific quantitative real-time PCR to determine the zygosity status of the *hppdPfW336-1Pa* and *2mepsps* genes. For each sample, two distinct sets of primer pairs amplified the target gene (*hppdPfW336-1Pa* or *2mepsps* gene) together with the endogenous reference gene (*adhC*) from cotton. For each sample, the copy number of the *hppdPfW336-1Pa* or *2mepsps* gene was determined relative to the one copy reference gene.

Chi-square analysis was performed for three generations of GHB811 cotton to confirm the segregation and stability of the GHB811 insert. The Chi-square analysis is based on testing the observed segregation ratio relative to the expected segregation ratio from Mendelian inheritance principles. For the F2, BC1F2 and BC2F2 generations of GHB811 cotton, the expected segregation ratio of homozygous, hemizygous and null segregate was 1:2:1. The χ^2 values were calculated with Microsoft Excel 2010 using the following equation.

$$\chi^2 = \sum \frac{|(\text{Observed} - \text{Expected})|^2}{\text{Expected}}$$

The results are summarized in Table 16 and Table 17.

Table 16. Observed Versus Expected Genotype for the *2mepsps* gene in F2, BC1F2 and BC2F2 of GHB811 Cotton as Determined by PCR Analysis.

	F2		BC1F2		BC2F2	
	Observed	**Expected**	**Observed**	**Expected**	**Observed**	**Expected**
Homozygous	19	21.25	24	23	45	57.25
Hemizygous	49	42.50	50	46	116	114.5
Null	17	21.25	18	23	68	57.25
X^2 Value	2.08		1.48		4.66	

* The critical value to reject the null hypothesis at the 5% confidence level is < 5.99 with two degrees of freedom.

Table 17. Observed Versus Expected Genotype for the *hppdPfW336-1Pa* gene in F2, BC1F2 and BC2F2 of GHB811 Cotton as Determined by PCR Analysis.

	F2		BC1F2		BC2F2	
	Observed	**Expected**	**Observed**	**Expected**	**Observed**	**Expected**
Homozygous	19	21.25	24	23	45	57.25
Hemizygous	49	42.50	50	46	116	114.5
Null	17	21.25	18	23	68	57.25
X^2 Value	2.08		1.48		4.66	

* The critical value to reject the null hypothesis at the 5% confidence level is < 5.99 with two degrees of freedom.

Segregation ratios determined for three generations of GHB811 cotton confirmed that the *hppdPfW336-1Pa* and *2mepsps* genes contained within the GHB811 insert are inherited in a predictable manner and as expected for a single insertion. These data are consistent with Mendelian principles and support the conclusion that the GHB811 event consists of a single insert integrated into a single chromosomal locus within the cotton nuclear genome.

Expression of the phenotype across different environments

Protein expression levels of HPPD W336 and 2mEPSPS were determined by Enzyme-Linked immunosorbent assay (ELISA) in field-grown cotton matrices from GHB811 plants treated and not treated with trait-specific herbicides grown at three field trials in the USA in 2015 ; M-574232-01; Node A.3 (e), (ii)).

Protein expression analysis was conducted on tissue samples harvested from plants grown in the USA in 2015. Six field sites were located in areas representative of the commercial production of cotton in the USA and sampled throughout the growing season for different tissues. Three sites were selected for expression analysis based on their geographical distribution (Mississippi, North Carolina and Texas), diversity and representativeness of field site management, and proximity to intensive cotton production areas.

There were two plots of GHB811 included at each site. One plot was treated with trait-specific herbicide while the other plot was not treated. The isoxaflutole application to the treated GHB811 entry was made at a rate of 104.9 to 106.6 g ai/ha before emergence (BBCH 00). The glyphosate application was made at a rate of 1104 to 1123 g ai/ha at the

seven to eight leaf growth stage (BBCH 17 - 18). All entries were of the Coker 312 background.

The matrices analyzed are summarized below:

Growth Stage	Matrix	Sample Description
4-6 leaf (BBCH14-16)	Leaf	All true leaves from 1 plant.
	Root	All roots from 1 plant.
Square initiation (BBCH 51-55)	Leaf	All true leaves from 1 plant.
Flowering (BBCH 60-69)	Pollen	Composite pollen from multiple plants.
2 weeks after first flower (BBCH 60-67)	Leaf	All true leaves from 1 plant.
	Square	Composite of 6 pre-candle squares from 1 plant.
	Boll	Composite 4 - 6 immature bolls from 1 plant.
	Whole Plant	All above ground portion of 1 plant.
Maturity (BBCH 83-97)	Fuzzy Seed	All commercially acceptable open bolls from 1 plant.

The BBCH-scale is a system for a uniform coding of phenologically similar growth stages of mono- and dicotyledonous plant species.

The quantitation of 2mEPSPS protein in leaf, root, pollen square, boll, whole plant and fuzzy seed samples was conducted with a validated 2mEPSPS-specific ELISA method using the EnviroLogix QualiPlate™ Kit for 2mEPSPS. The quantitation of HPPD W336 protein in leaf, root, pollen square, boll, whole plant and fuzzy seed samples was conducted with a validated HPPD W336-specific ELISA method using the EnviroLogix QuantiPlate™ Kit.

Expression of 2mEPSPS in Cotton Matrices

The level of 2mEPSPS protein in not treated and treated GHB811 cotton leaf, root, square, boll, whole plant and fuzzy seed matrices ranged from 76.36 to 1762.54 µg/g DW and 86.67 to 1685.85 µg/g DW, respectively (Table 18). The 2mEPSPS protein concentrations in not treated and treated GHB811 cotton pollen ranged from 12.86 to 33.47 µg/g FW and 21.42 to 33.15 µg/g FW, respectively (Table 18).

Leaf at BBCH 60-67 and BBCH 51-55 growth stages demonstrated the highest mean 2mEPSPS protein expression levels (Table 18). Mean (±SD) 2mEPSPS expression levels in not treated and treated leaf at BBCH 60-67 of GHB811 cotton was 1422.12 ± 206.41 µg/g DW and 1267.95 ± 247.75 µg/g DW, respectively. Mean (±SD) 2mEPSPS expression levels in not treated and treated leaf at BBCH 51-55 of GHB811 cotton was 1344.37 ± 224.96 µg/g DW and 1269.39 ± 175.42 µg/g DW, respectively.

Fuzzy seed demonstrated the lowest mean 2mEPSPS protein expression in all matrices reported on a DW basis (Table 18). Mean (± SD) 2mEPSPS expression levels in not treated and treated fuzzy seed of GHB811 cotton was 145.11 ± 37.86 µg/g DW and 150.88 ± 27.87 µg/g DW, respectively.

The mean 2mEPSPS concentrations for not treated and treated pollen were 24.69 ± 6.60 µg/g FW and 27.68 µg/g ± 3.47 µg/g FW respectively.

Expression of HPPD W336 in Cotton Matrices

The level of HPPD W336 expression in not treated and treated GHB811 cotton leaf, root, square, boll, whole plant and fuzzy seed matrices ranged from 10.91 to 1673.89 µg/g DW and 11.01 to 1402.82 µg/g DW, respectively (Table 19). The HPPD W336 protein

concentrations in not treated and treated GHB811 cotton pollen ranged from <LLOQ to 0.69 µg/g FW and <LLOQ to 0.68 µg/g FW, respectively (Table 19).

Leaf at BBCH 51-55 growth stage demonstrated the highest mean HPPD W336 protein expression levels (Table 19). Mean (± SD) HPPD W336 expression levels in not treated and treated leaf at BBCH 51-55 of GHB811 cotton was 1043.64 ± 322.96 µg/g DW and 956.75 ± 204.79 µg/g DW, respectively.

Root demonstrated the lowest mean HPPD W336 protein expression levels in all matrices reported on a DW basis (Table 19). Mean (± SD) HPPD W336 expression levels in not treated and treated root of GHB811 cotton was 22.12 ± 8.37 µg/g DW and 25.42 ± 10.98 µg/g DW, respectively.

Mean (± SD) HPPD W336 expression levels in not treated and treated fuzzy seed of GHB811 cotton was 29.61 ± 14.96 µg/g DW and 27.01 ± 9.78 µg/g DW, respectively.

The HPPD W336 concentrations for majority of the not treated and treated pollen samples were below LLOQ (Table 19).

Table 18. Expression of 2mEPSPS in Cotton Matrices Harvested from Treated and Not Treated GHB811 Grown at Three Sites

Matrix	BBCH Growth Stage	Entry	2mEPSPS (µg/g DW)				2mEPSPS (µg/g FW)			
			Mean	SD	Min	Max	Mean	SD	Min	Max
Leaf	14-16	B	968.03	520.32	211.22	1423.94	161.89	83.89	36.55	236.36
		C	874.63	353.68	347.43	1418.35	144.25	43.41	64.31	193.40
Root	14-16	B	169.25	46.94	97.44	249.81	24.56	4.22	16.16	30.28
		C	163.76	33.16	118.30	218.48	23.08	3.71	15.02	28.60
Leaf	51-55	B	1344.37	224.96	840.89	1651.21	255.91	62.43	174.16	380.38
		C	1269.39	175.42	1029.95	1685.85	252.37	48.53	176.49	342.72
Leaf	60-67	B	1422.12	206.41	1117.38	1762.54	307.05	42.53	238.76	383.67
		C	1267.95	247.75	756.49	1600.36	273.05	59.69	156.17	351.98
Pollen	60-69	B	NA	NA	NA	NA	24.69	6.60	12.86	33.47
		C	NA	NA	NA	NA	27.68	3.47	21.42	33.15
Square	60-67	B	591.00	53.34	485.64	689.39	126.23	14.92	105.64	151.38
		C	506.64	84.21	381.81	659.10	106.20	14.36	78.39	124.97
Bolls	60-67	B	474.77	65.86	360.31	575.47	80.29	18.08	58.02	122.56
		C	437.00	61.85	318.89	522.18	71.26	9.67	55.93	86.08
Whole Plant	60-67	B	788.13	128.73	595.67	1080.81	182.13	47.46	127.19	276.04
		C	795.81	132.72	611.73	1065.80	176.28	48.95	116.95	290.24
Fuzzy Seed	83-97	B	145.11	37.86	76.36	221.42	129.79	38.41	65.07	205.88
		C	150.88	27.87	86.67	198.93	132.94	20.38	80.83	162.76

Entry B = GHB811 (not treated); Entry C = GHB811 (treated).
Mean and standard deviation (SD) for each entry was based on the total sample population (N=12).
NA = Not Applicable. Pollen samples were analyzed on fresh tissue only.

Table 19. Expression of HPPD W336 in Cotton Matrices Harvested from Treated and Not Treated GHB811 Grown at Three Sites

Matrix	BBCH Growth Stage	Entry	HPPD W336 (µg/g DW)				HPPD W336 (µg/g FW)			
			Mean	SD	Min	Max	Mean	SD	Min	Max
Leaf	14-16	B	668.06	478.48	136.50	1337.49	116.29	87.24	23.11	244.29
		C	808.10	403.34	376.72	1402.82	142.10	72.18	58.71	231.78
Root	14-16	B	22.12	8.37	12.76	43.01	3.40	1.52	1.66	6.62
		C	25.42	10.98	11.10	46.06	3.59	1.48	1.55	5.89
Leaf	51-55	B	1043.64	322.96	717.12	1673.89	198.51	69.61	124.77	328.42
		C	956.75	204.79	722.63	1232.40	188.78	40.22	124.84	245.22
Leaf	60-67	B	862.75	208.08	515.57	1225.60	184.73	36.97	119.55	250.74
		C	781.28	164.18	563.26	1013.20	166.51	31.37	116.28	221.29
Pollen	60-69	B	NA	NA	NA	NA	<LLOQ	ND	<LLOQ	0.69
		C	NA	NA	NA	NA	<LLOQ	ND	<LLOQ	0.68
Square	60-67	B	304.97	24.47	269.07	337.78	65.30	8.67	51.51	83.16
		C	284.52	34.38	235.73	365.42	60.88	14.69	44.51	95.87
Bolls	60-67	B	181.03	37.20	116.55	241.68	30.59	8.34	20.77	46.27
		C	125.62	34.32	70.03	193.68	20.29	4.50	12.28	27.37
Whole Plant	60-67	B	308.52	93.13	159.44	433.89	68.00	13.34	43.84	87.13
		C	297.03	73.31	182.34	399.34	63.21	9.37	44.81	77.78
Fuzzy Seed	83-97	B	29.61	14.96	10.91	62.33	26.45	13.65	9.30	55.96
		C	27.01	9.78	11.01	43.85	23.82	8.46	10.27	39.46

Entry B = GHB811 (not treated); Entry C = GHB811 (treated).
Mean and standard deviation (SD) for each entry was based on the total sample population (N=12).
NA = Not Applicable. Pollen samples were analyzed on fresh tissue only.
ND = Not Determined. SD for HPPD W336 expression levels were not determined, since only 1 sample from each entry had a quantifiable value (> Lower Limit of Quantitation (LLOQ)).

(g) an analysis of the expressed RNA transcripts, where RNA interference has been used.

RNA interference has not been used to develop this food product.

B.1 Characterisation and safety assessment of new substances

(a) A full description of the biochemical function and phenotypic effects of all new substances (e.g. a protein or an untranslated RNA) that are expressed in the new GM organism, including their levels and site of accumulation, particularly in edible portions

HPPD W336 protein

The coding sequence of the 4-hydroxyphenylpyruvate dioxygenase (HPPD) protein was isolated from the *Pseudomonas fluorescens* strain A32. One amino acid was substituted (glycine at position 336 with tryptophan) to improve the tolerance to the class of herbicides known as HPPD inhibitors. The modified protein is designated as HPPD W336 (Boudec, P.; et al.; 2001; M-229534-01; Node A.1 (a)).

The *hppd* gene was isolated from the bacterium *Pseudomonas fluorescens*, strain A32. *P. fluorescens* is a Gram-negative, rod-shaped, motile, asporogenous, aerobic bacterium. *P. fluorescens*, is ubiquitous in the environment, including soil, water and food (OECD; 1997; M-357528-01; Node A.2 (a), (i)). It has many beneficial uses in agriculture, human health and bioremediation. It is not described as allergenic, toxic or pathogenic to healthy humans and animals and has an overall history of safe use.

HPPD proteins are ubiquitous in nature across all kingdoms: bacteria, fungi, plants and animals including mammals. HPPD amino acid sequences have been determined in bacteria such as *Streptomyces avermitilis* (Accession number Q53586), in fungi such as *Aspergillus fumigatus* (Accession number Q4WPV8), in plants such as *Arabidopsis thaliana* (Accession number P93836), and in animals such as *Caenorhabditis elegans* (Accession number Q22633), mouse (*Mus musculus*, Accession number P49429), and human (*Homo sapiens*, Accession number P32754).

In particular, HPPD proteins have been characterized in organisms present in human food, such as carrot (*Daucus carota*, Accession number O23920), barley (*Hordeum vulgare* Accession number O48604; Falk *et al*., 2002), pork (*Sus scrofa*, Accession number Q02110)) and beef (*Bos Taurus*, Accession number Q5EA20).

HPPD proteins are present in food from plant, fungal or animal origin, with good safety records. Therefore, HPPD proteins have a history of safe use.

The biochemical pathways in which HPPD is involved differ between plants and non-photosynthetic organisms. In bacteria and animals, it merely serves catabolic purposes by catalyzing the first committed step in tyrosine degradation that in the end yields energetically exploitable glucogenic and ketogenic products (Brownlee, J. M.; et al.; 2004; M-358228-01; Node B.1 (a)). In plants, however, it is also involved in several anabolic pathways; its reaction product homogentisate (2,5-dihydroxyohenylacetate) being the aromatic precursor of tocopherol, tocotrienols and plastoquinone, which are essential to the photosynthetic transport chain and antioxidative systems (Fritze, I. M.; et al.; 2004; M-359884-01; Node B.1 (a)).; Figure 32 shows a diagram of the different metabolic pathways in which HPPD is involved in plants and non-photosynthetic organisms.

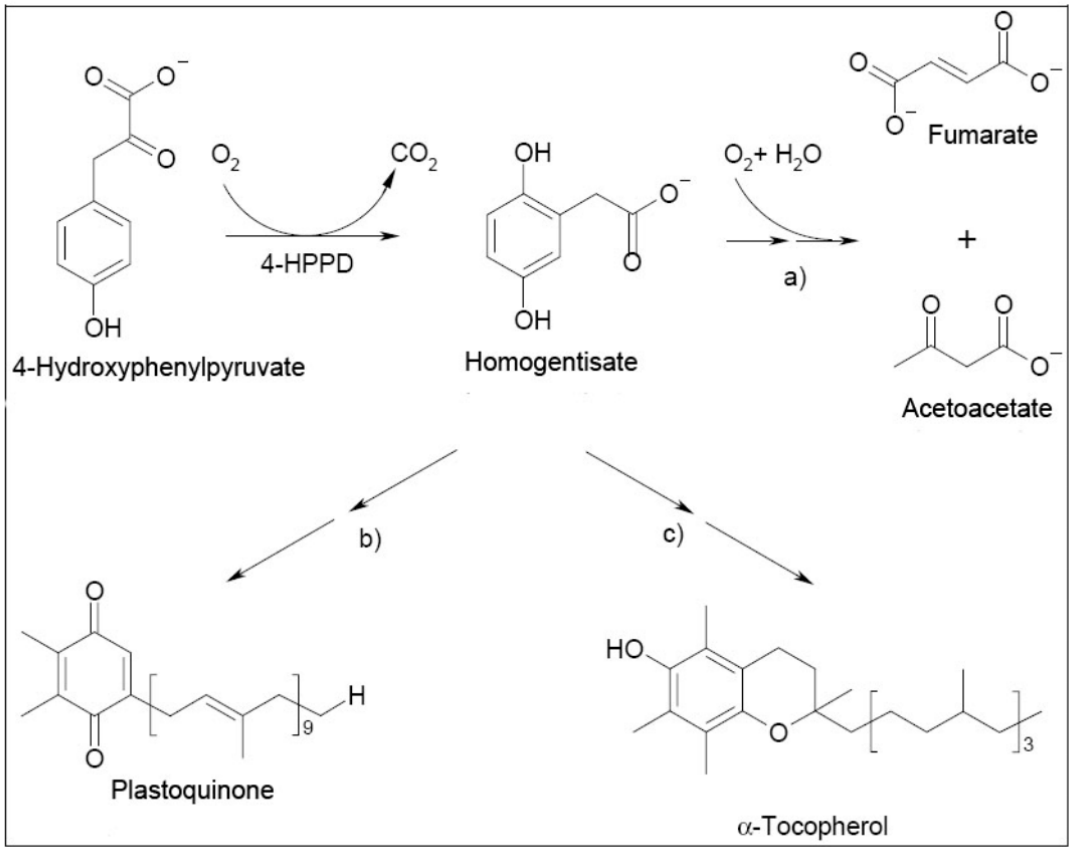


Figure 32. Biochemical pathways of HPPD proteins

a) catabolism of tyrosine, b) biosynthesis of plastoquinone (plants)
c) biosynthesis of tocopherol and tocotrienols (plants)

HPPD enzymes require an α-keto acid and molecular oxygen to oxidize or oxygenate a third molecule. The activity of HPPD is suppressed by benzoylisoxazoles bleaching herbicides, such as isoxaflutole (IFT), and by β-triketones such as sulcotrione and mesotrione (Pallett, K. E.; et al.; 2001; M-357534-01; Node B.1 (a)) ; Dayan, F. E.; et al.; 2007; M-357499-01; Node B.1 (a)).

The inhibitor of HPPD is the diketonitrile (DKN) derivative of Isoxaflutole (IFT) formed by the opening of the isoxazole ring. DKN is formed rapidly in plants following uptake of IFT by roots and shoots. HPPD enzyme inhibition results in the disruption of the biosynthesis of carotenoids, which destabilizes photosynthesis and leads to bleaching of the foliage and death of the plant (Figure 33).

In order to create a form of the HPPD enzyme with tolerance to IFT herbicide, a single amino acid substitution, glycine (G) to tryptophan (W) at position 336, was introduced to the native HPPD protein from *Pseudomonas fluorescens* (Boudec, P.; et al.; 2001) resulting in the modified IFT-tolerant HPPD W336 protein.

Several different HPPD variants, including the wild type HPPD and modified HPPD W336 enzymes were tested for their activity in the presence or absence of the inhibitor IFT. When compared to the wild type HPPD enzyme, HPPD W336 enzyme was significantly less inhibited by IFT (Fischer, K.; 2008; M-359847-01; Node B.1 (a)).

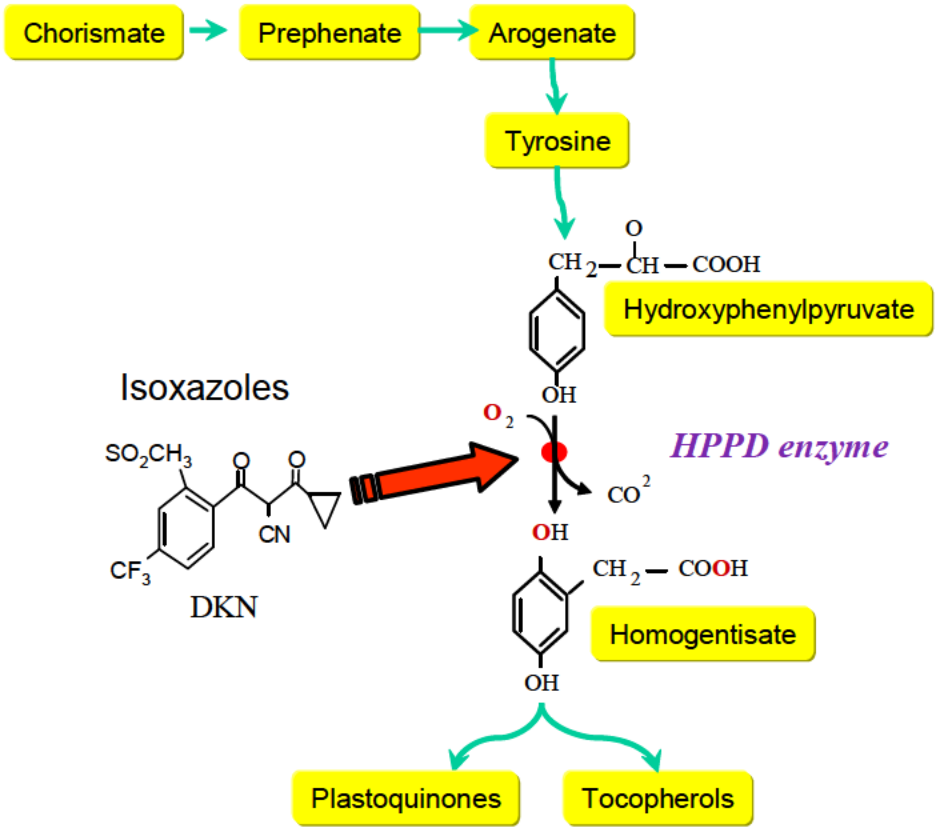


Figure 33. Interaction of HPPD and isoxazole herbicides

2mEPSPS protein

The coding sequence of 5-enol pyruvylshikimate-3-phosphate synthase (*epsps*) gene was isolated from maize (*Zea Mays* L.). Two amino acids were substituted (threonine by isoleucine at position 102 and proline by serine at position 106) (Lebrun, M. *et al*., 1997; M-216526-01; Node A.1 (a)). These modifications confer to the protein a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of the herbicide. Therefore, the plants expressing this modified protein become tolerant to glyphosate herbicides (Lebrun, M. *et al*., 1997; M-216526-01; Node A.1 (a)). The modified protein is designated as 2mEPSPS.

5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) catalyzes the penultimate step of the shikimate pathway, which is responsible for the synthesis of aromatic amino acids and other aromatic compounds in plants, fungi and microorganisms including apicomplexan parasites (Herrmann, K. M.; 1995; M-269843-01; Node B.1 (a)). As such, it has been shown that EPSPS enzymes are ubiquitous in nature and are present in food and feed derived from plant and microbial sources. No health-related adverse effects have been associated with these proteins.

The *2mepsps* gene was generated by introducing mutations into the *epsps* gene from maize (*Z. mays* L.) that result in two amino acid substitutions. The modified EPSPS (2mEPSPS) enzyme has a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of glyphosate herbicides (Lebrun, M. *et al*., 1997; M-216526-01; Node A.1 (a)). Since the 2mEPSPS protein is derived from maize and has only

two amino acid modifications, the safety profile of the novel protein is expected to remain unchanged relative to its wild-type counterpart.

EPSPS proteins are present in food and feed from plant and microbial sources with good safety records. Therefore, EPSPS proteins have a history of safe use.

5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) is an enzyme that catalyzes the condensation of shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) to 5-enolpyruvylshikimate-3-phosphate and phosphate in the shikimate pathway, which is responsible for the synthesis of aromatic amino acids and other aromatic compounds in plants, fungi and microorganisms including apicomplexan parasites (Figure 34).

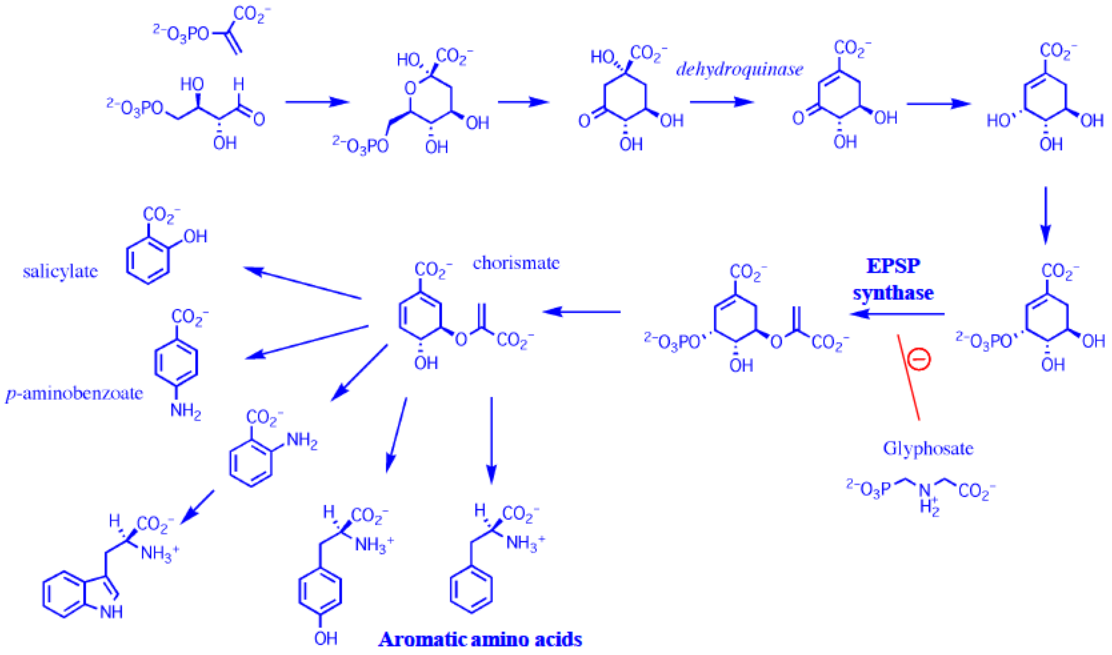


Figure 34. The shikimate pathway

Glyphosate inhibits the EPSPS enzyme and shuts down the shikimate pathway, leading to plant death. The modified EPSPS (2mEPSPS) enzyme has a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of glyphosate herbicides (Lebrun, M. *et al.*, 1997; M-216526-01; Node A.1 (a); Healy, M. L.; Schoenbrunn, E.; 2006; M-267370-01; Node B.1 (a)).

For the safety assessment of GM crops, certain safety tests and studies require large amounts of protein. The expression levels of the 2mEPSPS and HPPD W336 in GHB811 were too low to allow for purification of sufficient quantities of the two proteins directly from GHB811 for use in the safety assessment studies. Therefore, the 2mEPSPS and HPPD W336 proteins were produced in a high-expressing recombinant host organism, *E.coli*. and the proteins produced by *E.coli* were engineered to match the amino acid sequences of their counterparts expressed in GHB811. The equivalence of GHB811 cotton-produced and bacterially-produced proteins were examined to ensure that the proteins from the two host sources were equivalent.

The Equivalence of GHB811-purified and microbially-produced 2mEPSPS proteins.

A purification of 2mEPSPS protein was performed from the GHB811 cotton leaf matrix using affinity chromatography. GHB811 cotton-purified 2mEPSPS protein was characterized and evaluated for equivalence with bacterially-produced 2mEPSPS protein based on a panel of analytical tests and assays, including densitometry analysis of a Coomassie-stained SDS-PAGE; western blot analysis; glycostaining analysis; *Mass Spectroscopy*; N-terminal sequence analysis; and EPSPS enzymatic activity assay 2016; M-568145-01; Node B.1 (a) ; M-497839-01; Node B.1 (a)).

Assessment and comparison of the apparent molecular mass

The GHB811 cotton-purified 2mEPSPS protein and the bacterially-produced 2mEPSPS protein were compared side by side by means of an SDS-PAGE analysis (Figure 35). Additionally, the bacterially-produced 2mEPSPS protein was spiked into the protein extract sample resulting from treatment of non-GM cotton variety Coker 312, which was subjected to the same affinity purification procedure as the plant-purified 2mEPSPS protein sample (*i.e.* treated non-GM counterpart) to allow comparison in a similar cotton plant matrix.

A specific, predominant band was observed for both samples, which corresponds to the expected molecular mass of the 2mEPSPS protein (47.4 kDa). This demonstrated that the apparent molecular mass of the GHB811 cotton-purified and the bacterially-produced 2mEPSPS protein are comparable. The treated non-GM counterpart negative control showed some non-specific background staining derived from the plant matrix.

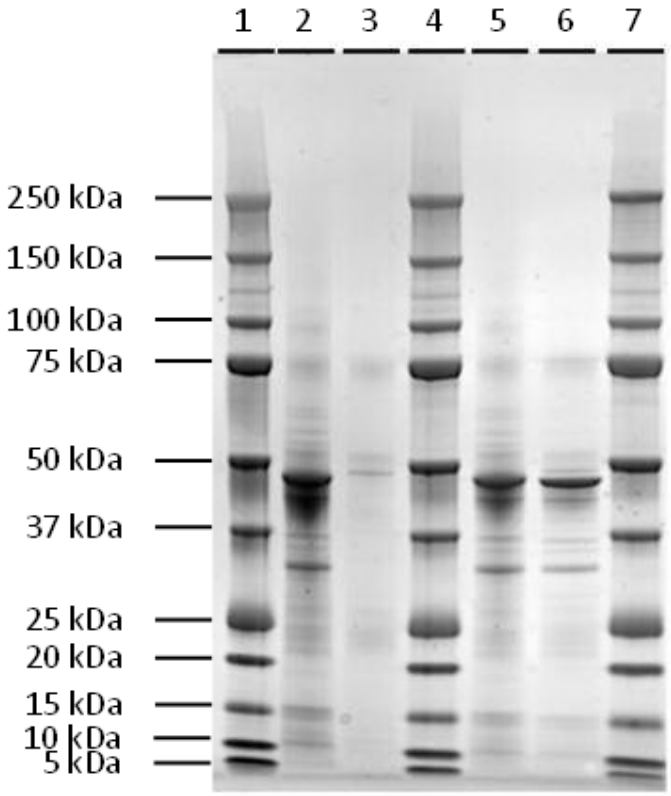


Figure 35: Apparent molecular mass assessment and comparison of GHB811 cotton-purified 2mEPSPS protein and the bacterially-produced 2mEPSPS protein

Both GHB811 cotton-purified and bacterially-produced 2mEPSPS protein samples were loaded on a Criterion XT Bis-Tris 4-12 % gel and SDS-PAGE gel electrophoresis was performed in 1x MOPS gel running buffer, followed by Coomassie staining.

Loading order :
Lane 1: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
Lane 2: 1 µg of 2mEPSPS protein of the GHB811 cotton-purified sample 16-RSTHN035-A-02
Lane 3: 1.78 µL of the non-GM counterpart derived negative control sample 16-RSTHN035-A-05
Lane 4: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
Lane 5: 1 µg of 2mEPSPS protein of the GHB811 cotton-purified sample 16-RSTHN035-A-02
Lane 6: 1 µg of bacterially-produced 2mEPSPS protein (batch 1417_2mEPSPS) spiked in 1.78 µL of non-GM counterpart derived negative control sample 16-RSTHN035-A-05 (16-RSTHN035-A-08)
Lane 7: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)

Assessment and comparison of the immuno-reactivity

The GHB811-purified 2mEPSPS protein and the bacterially-produced 2mEPSPS protein were compared side by side by means of western blot analysis (Figure 36).

Using a 2mEPSPS-specific polyclonal antibody, a signal corresponding to the expected molecular mass of the 2mEPSPS protein was detected for both samples. A very weak band of a comparable size was observed for the crude extract of the non-GM counterpart, which most likely corresponds to the cotton endogenous EPSPS protein.

The obtained results confirmed the immuno-reactivity of the GHB811-purified 2mEPSPS protein and the comparability to the bacterially-produced 2mEPSPS protein.

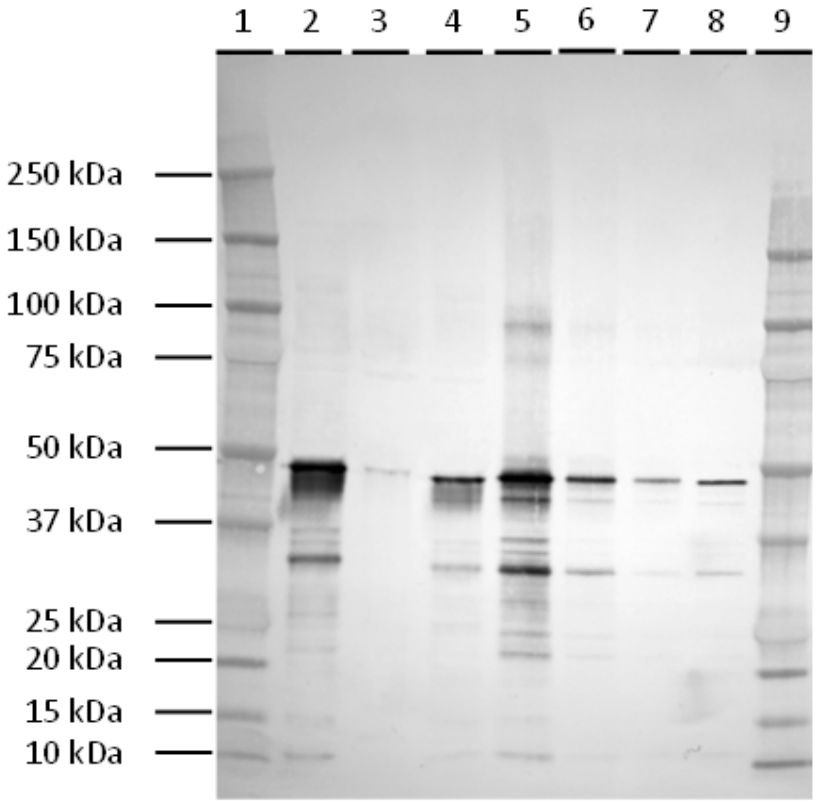


Figure 36: Assessment and comparison of immuno-reactivity of GHB811 cotton-purified 2mEPSPS protein and bacterially-produced 2mEPSPS protein

Both plant-purified and bacterially-produced 2mEPSPS protein samples were loaded on a Criterion XT Bis-Tris 4-12% gel and SDS-PAGE gel electrophoresis was performed in 1x MOPS gel running buffer. After semi-dry blotting, the proteins were visualized by colorimetric detection using a 1:5,000 dilution of the rabbit anti-2mEPSPS antibody (batch 1323_2mEPSPS_Ab) as primary antibody and a 1:7,000 dilution of the goat anti-rabbit antibody conjugated to Alkaline Phosphatase as the secondary antibody, followed by NBT BCIP substrate addition.

Loading order :
Lane 1: 5 µL of Precision Plus Protein™ Dual Xtra Standards
Lane 2: 10 µL of crude protein extract from GHB811 cotton (16-RSTHN035-A-00)
Lane 3: 10 µL of crude protein extract from the non-GM counterpart (16-RSTHN035-A-03)
Lane 4: 40 ng 2mEPSPS protein spiked into 10 µL of crude extract from the non-GM counterpart (16-RSTHN035-A-06)
Lane 5: 40 ng of plant-purified 2mEPSPS protein from GHB811 cotton (16-RSTHN035-A-02)
Lane 6: 10 ng of plant-purified 2mEPSPS protein from GHB811 cotton (16-RSTHN035-A-02)
Lane 7: 4 ng of plant-purified 2mEPSPS protein from GHB811 cotton (16-RSTHN035-A-02)
Lane 8: 10 ng of bacterially-produced 2mEPSPS protein (batch 1417_2mEPSPS)
Lane 9: 5 µL of the Precision Plus Protein™ Dual Xtra Standards

Assessment and comparison of the glycosylation status

The results of the glycostaining analysis are shown in Figure 37.

The glycosylated proteins of the horseradish peroxidase positive control and the alpha-one acidic glycoprotein of the glycoprotein mix were visualized as bright bands on the gel, while for the 2mEPSPS protein samples, no signal was observed (Figure 37, panel A).

The presence of sufficient 2mEPSPS protein on the gel was demonstrated by staining the gels with Coomassie after the glyco-staining procedure (Figure 37, panel B).

The absence of glycosylation was demonstrated for both the GHB811-purified 2mEPSPS protein and the bacterially-produced 2mEPSPS protein. Consequently, both 2mEPSPS protein samples have a comparable glycosylation status.

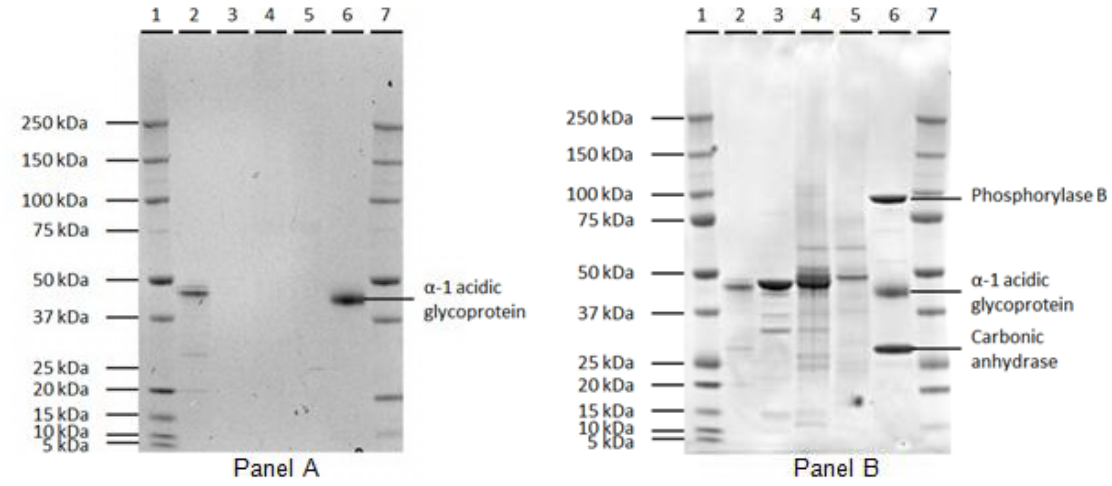


Figure 37: **Assessment and comparison of the of the glycosylation status of the GHB811 cotton-purified 2mEPSPS protein and the bacterially-produced 2mEPSPS**

One Criterion XT Bis-Tris 4-12 % SDS-PAGE gel was prepared and cut into two pieces, each part containing 3 µg 2mEPSPS protein of both the bacterially-produced 2mEPSPS protein batch 1417_2mEPSPS and the GHB811 cotton-purified 2mEPSPS protein together with the appropriate positive and negative controls to assess the glycosylation status.

Panel A shows the result of the staining using the Glycoprotein Detection Kit (Sigma) to demonstrate the absence of glycosylation of both the 2mEPSPS protein samples. For the second set of samples, a glycostaining was performed in which the oxidation step in the procedure was omitted to demonstrate the absence of any non-specific binding (data not shown).
Panel B shows a Coomassie staining of the SDS-PAGE gel to demonstrate the presence of the proteins on the gel.

Loading order of Panel A and B:
Lane 1: 5 µL of the Precision Plus Protein™ Dual Xtra Standards
Lane 2: 45.45 pmoles of Horseradish Peroxidase (positive control)
Lane 3: 3 µg of bacterially-produced 2mEPSPS protein batch 1417_2mEPSPS
Lane 4: 3 µg of GHB811 cotton-purified 2mEPSPS protein of sample 16-RSTHN035-A-01
Lane 5: 7.83 µL of the non-GM counterpart derived negative control sample (16-RSTHN035-A-04)
Lane 6: 45.45 pmoles of Glycoprotein mix
Lane 7: 5 µL of the Precision Plus Protein™ Dual Xtra Standards

Assessment of the intact molecular mass and peptide mapping

The intact molecular mass was determined using the UPLC-UV-MS and the peptide mapping was established using UPLC-UV-MS^E analysis for the GHB811-purified 2mEPSPS protein.

The determined intact molecular mass allowed the identification of two intact molecular masses. The first and major molecular mass of 47551.5 Da corresponds to the theoretical molecular mass of a 2mEPSPS protein with an N-terminal cysteinic sulfinic acid, derived from the transit peptide (47551.0 Da). The second, minor molecular mass of 47284.9 Da corresponds to the theoretical molecular mass of an N-terminal des-Methionine (mature form of the 2mEPSPS protein minus the initial methionine residue, desMet) (47284.7 Da).

Peptides resulting from a trypsin digest of the GHB811-purified 2mEPSPS protein were analysed. Figure 38 provides an overview of the mapped peptides against the theoretical amino acid sequence of the 2mEPSPS protein. A coverage of 89 % was determined, which confirms the identity of the 2mEPSPS protein.

The intact molecular mass was determined and the peptide mapping was established using LC-UV-MS analysis for bacterially-produced 2mEPSPS protein.

The determined intact molecular mass of 47288 Da confirmed the theoretical molecular mass of 47284 Da corresponding with the mass of the 2mEPSPS protein minus the methionine residue (desMet).

Peptides resulting from a trypsin digest of bacterially-produced 2mEPSPS protein were analysed using LC-UV-MS. Figure 39 provides an overview of the mapped peptides against the theoretical amino acid sequence of the 2mEPSPS protein. A coverage of 95.5 % was determined, which confirms the identity of 2mEPSPS protein.

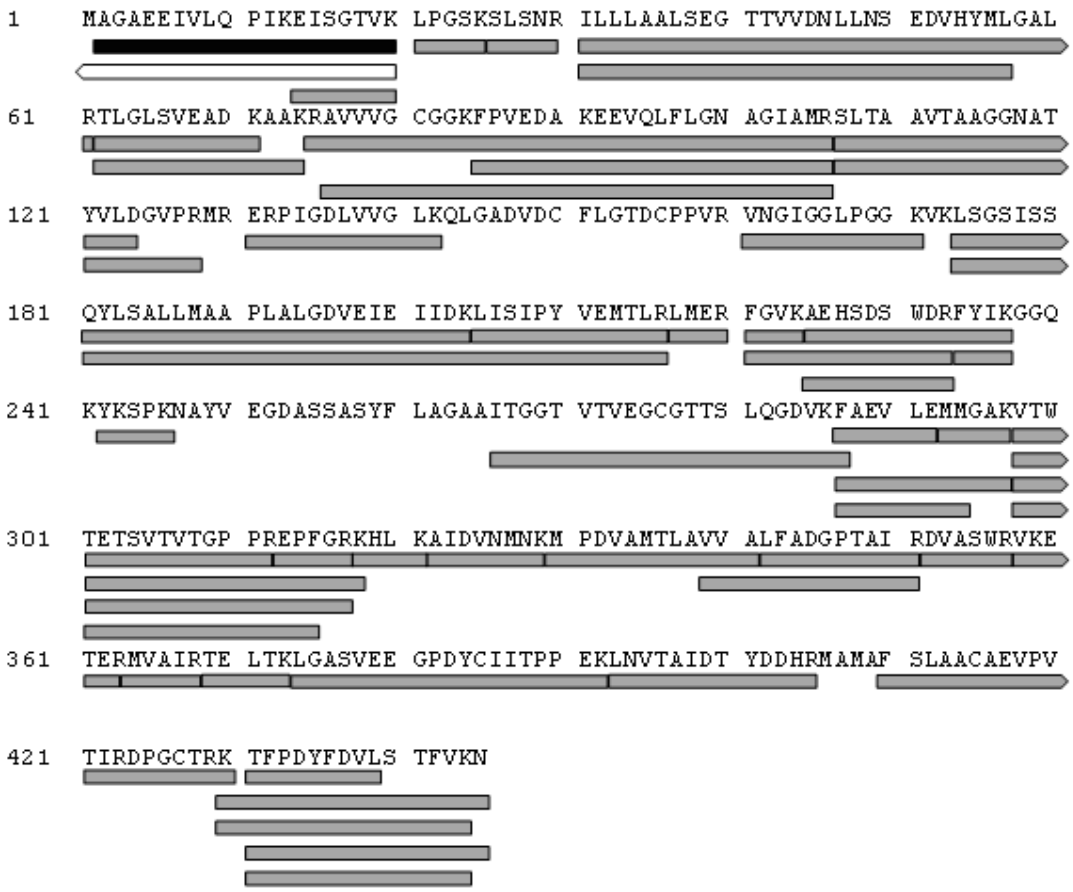


Figure 38: **Schematic overview of the coverage of the theoretical 2mEPSPS sequence by the tryptic peptides from the cotton GHB811-purified 2mEPSPS protein**

Black and white bars represent the N-terminal peptides corresponding respectively to the desMet 2mEPSPS protein and the 2mEPSPS protein with an N-terminal cysteinic sulfinic acid, derived from the transit peptide.

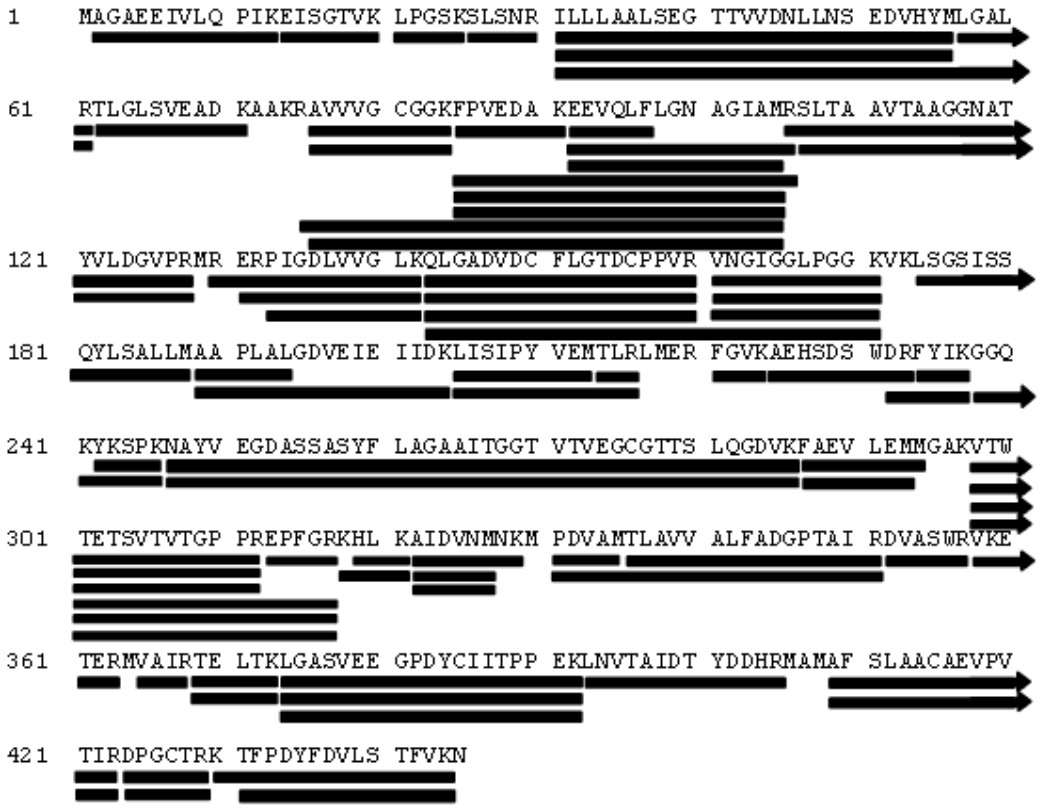


Figure 39: Schematic overview of the detected peptides derived from tryptic digestion of bacterially-produced 2mEPSPS

Assessment of the N-terminal sequence

The N-terminal sequence of the GHB811 cotton-purified 2mEPSPS protein was determined by Edman degradation. The obtained data for the GHB811 cotton-purified sample suggested the presence of two N-termini. The obtained sequence result could be resolved as AGAEEIVLQP, corresponding to the desMet N-terminus (*i.e.* N-terminus without methionine) of the 2mEPSPS protein, and sequence XMAGAEEIVL, potentially corresponding to incomplete cleavage of the transit peptide of the 2mEPSPS protein. This confirms the results observed with the intact molecular mass determination.

The N-terminal sequence of bacterially-produced 2mEPSPS protein was determined by Edman degradation. The obtained sequence result could be resolved as AGAEEIVLQP, corresponding to the desMet N-terminus (*i.e.* N-terminus without methionine) of the 2mEPSPS protein. This confirms the results observed with the intact molecular mass determination.

Assessment and comparison of EPSPS specific enzymatic activity

By means of a quantitative activity assay, the GHB811 cotton-purified 2mEPSPS protein was assessed for EPSPS functional activity. The result was compared to specific activity observed for the bacterially-produced 2mEPSPS protein spiked in the treated non-GM counterpart sample.

The observed mean specific EPSPS activity for the bacterially-produced 2mEPSPS protein was 5.230 U/mg 2mEPSPS protein, which demonstrated the validity of the observed results. A mean specific EPSPS activity was observed for the GHB811 cotton-purified 2mEPSPS

protein sample which was 2.451 U/mg 2mEPSPS protein. The mean specific EPSPS activity of the bacterially-produced 2mEPSPS protein spiked into the treated non-GM counterpart sample was 5.333 U/mg 2mEPSPS protein.

Since the observed EPSPS enzymatic activities of the GHB811 cotton-purified 2mEPSPS protein was within the same order of magnitude of the bacterially-produced 2mEPSPS protein, both 2mEPSPS proteins were considered as functionally equivalent.

Conclusion

The equivalence of the GHB811 cotton-purified 2mEPSPS protein with bacterially-produced 2mEPSPS protein was demonstrated based on a panel of analytical tests and assays, including densitometry analysis of a Coomassie-stained SDS-PAGE; western blot analysis; glycostaining analysis; *Mass Spectroscopy*; N-terminal sequence analysis; and EPSPS enzymatic activity assay. In addition to the des-Methionine 2mEPSPS protein (N-terminus without methionine), the GHB811 cotton-purified sample contained a second 2mEPSPS-derived structure with two additional amino acid residues at the N-terminus, potentially corresponding to incomplete cleavage of the transit peptide of the 2mEPSPS protein. This form did not have any impact on the observed characteristics for the GHB811 cotton-purified 2mEPSPS protein.

The Equivalence of GHB811-purified and microbially-produced HPPD W336 proteins.

A purification of HPPD W336 protein was performed from the GHB811 cotton leaf matrix using affinity chromatography. GHB811 cotton-purified HPPD W336 protein was characterized and the equivalence evaluated with bacterially-produced HPPD W336 protein based on a panel of analytical tests and assays, including densitometry analysis of a Coomassie-stained SDS-PAGE; western blot analysis; glycostaining analysis; *Mass Spectroscopy*; and N-terminal sequence analysis 2016; M-576569-01; Node B.1 (a); ; M-497842-01; Node B.1 (a)).

Assessment and comparison of the apparent molecular mass

The GHB811 cotton-purified HPPD W336 protein and the bacterially-produced HPPD W336 protein were compared side by side by means of an SDS-PAGE analysis (Figure 40). Additionally, the bacterially-produced HPPD W336 protein was spiked into the protein extract sample resulting from treatment of non-GM cotton variety Coker 312, which was subjected to the same affinity purification procedure as the plant-purified HPPD W336 protein sample (*i.e.* treated non-GM counterpart) to allow comparison in a similar cotton plant matrix.

A specific, predominant band was observed for both samples, which corresponds to the expected molecular mass of the HPPD W336 protein (40.3 kDa). This demonstrated that the apparent molecular mass of the GHB811 cotton-purified and the bacterially-produced HPPD W336 protein are comparable. The treated non-GM counterpart negative control showed some non-specific background staining derived from the plant matrix.

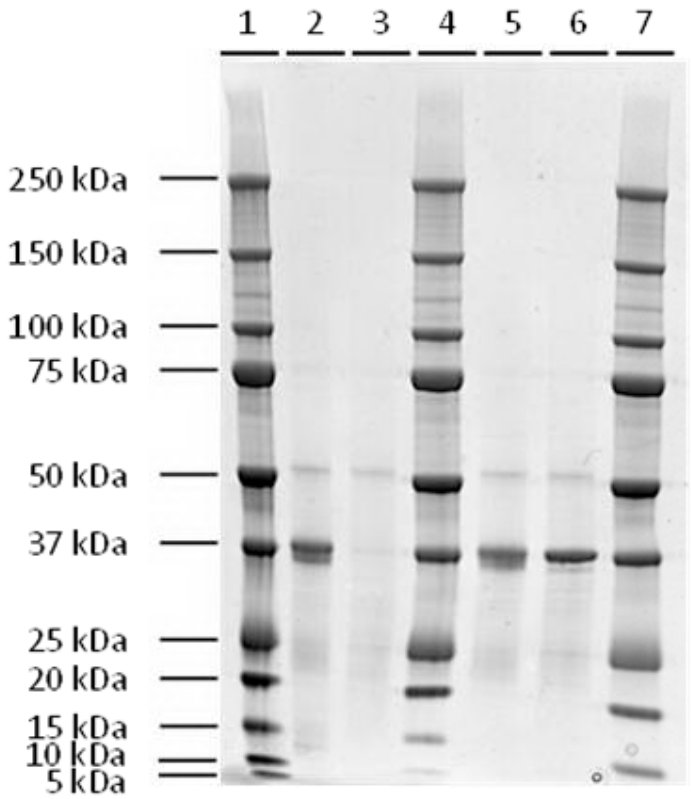


Figure 40: Apparent molecular mass assessment and comparison of GHB811 cotton-purified HPPD W336 protein and the bacterially-produced HPPD W336 protein

Both GHB811 cotton-purified and bacterially-produced HPPD W336 protein samples were loaded on a Criterion XT Bis-Tris 4-12 % gel and SDS-PAGE gel electrophoresis was performed in 1x MOPS gel running buffer, followed by Coomassie staining.

Loading order :
Lane 1: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
Lane 2: 1 µg of HPPD W336 protein of the GHB811 cotton-purified sample 16-RSTHN035-B-01
Lane 3: 5 µL of the non-GM counterpart derived negative control sample 16-RSTHN035-B-03
Lane 4: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
Lane 5: 1 µg of HPPD W336 protein of the GHB811 cotton-purified sample 16-RSTHN035-B-01
Lane 6: 1 µg of bacterially-produced HPPD W336 protein (batch 1411_HPPD W336) spiked in 5 µL of non-GM counterpart derived negative control sample 16-RSTHN035-B-03 (16-RSTHN035-B-05)
Lane 7: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)

Assessment and comparison of the immuno-reactivity

The GHB811 cotton-purified HPPD W336 protein and the bacterially-produced HPPD W336 protein were compared side by side by means of a western blot (Figure 41).

Using a HPPD W336-specific polyclonal antibody, a signal corresponding to the expected molecular mass of the HPPD W336 protein was detected for both samples. A very weak band of a slightly lower size was observed for the crude extract of the non-GM counterpart, which is due to cross-reactivity of the used polyclonal antibody batch to the plant matrix.

The obtained results confirmed the immuno-reactivity of the GHB811 cotton-purified HPPD W336 protein and the comparability to the bacterially-produced HPPD W336.

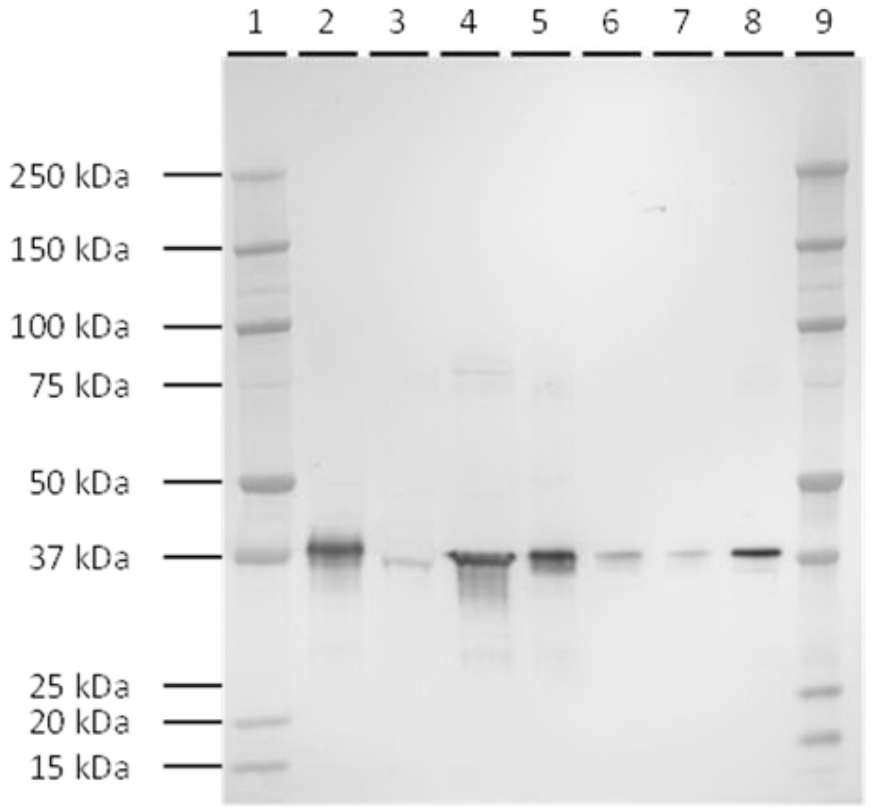


Figure 41: Assessment and comparison of immuno-reactivity of GHB811 cotton-purified HPPD W336 protein and bacterially-produced HPPD W336 protein

Both plant-purified and bacterially-produced HPPD W336 protein samples were loaded on a Criterion XT Bis-Tris 4-12% gel and SDS-PAGE gel electrophoresis was performed in 1x MOPS gel running buffer. After semi-dry blotting, the proteins were visualized by colorimetric detection using a 1:10,000 dilution of the rabbit anti-HPPD W336 antibody (batch 1227_HPPD W336_Ab) as primary antibody and a 1:7,000 dilution of the goat anti-rabbit antibody conjugated to Alkaline Phosphatase as the secondary antibody, followed by NBT and BCIP substrate addition.

Loading order :
Lane 1: 5 µL of Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
Lane 2: 10 µL of crude protein extract from GHB811 cotton (16-RSTHN035-B-00)
Lane 3: 10 µL of crude protein extract from the non-GM counterpart (16-RSTHN035-B-02)
Lane 4: 40 ng HPPD W336 protein spiked into 10 µL of crude extract from the non-GM counterpart (16-RSTHN035-B-04)
Lane 5: 40 ng of plant-purified HPPD W336 protein from GHB811 cotton (16-RSTHN035-B-01)
Lane 6: 10 ng of plant-purified HPPD W336 protein from GHB811 cotton (16-RSTHN035-B-01)
Lane 7: 4 ng of plant-purified HPPD W336 protein from GHB811 cotton (16-RSTHN035-B-01)
Lane 8: 10 ng of bacterially-produced HPPD W336 protein (batch 1411_HPPD W336)
Lane 9: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)

Assessment and comparison of the glycosylation status

The results of the glycostaining analysis are shown in Figure 42.

The glycosylated proteins of the horseradish peroxidase positive control and the alpha-one acidic glycoprotein of the glycoprotein mix were visualized as bright bands on the gel, while for the HPPD W336 protein samples, no signal was observed (Figure 42, panel A).

The presence of sufficient HPPD W336 protein on the gel was demonstrated by staining the gels with Coomassie after the glyco-staining procedure (Figure 42, panel B).

The absence of glycosylation was demonstrated for both the GHB811 cotton-purified HPPD W336 protein and the bacterially-produced HPPD W336 protein. Consequently, both HPPD W336 protein samples have a comparable glycosylation status.

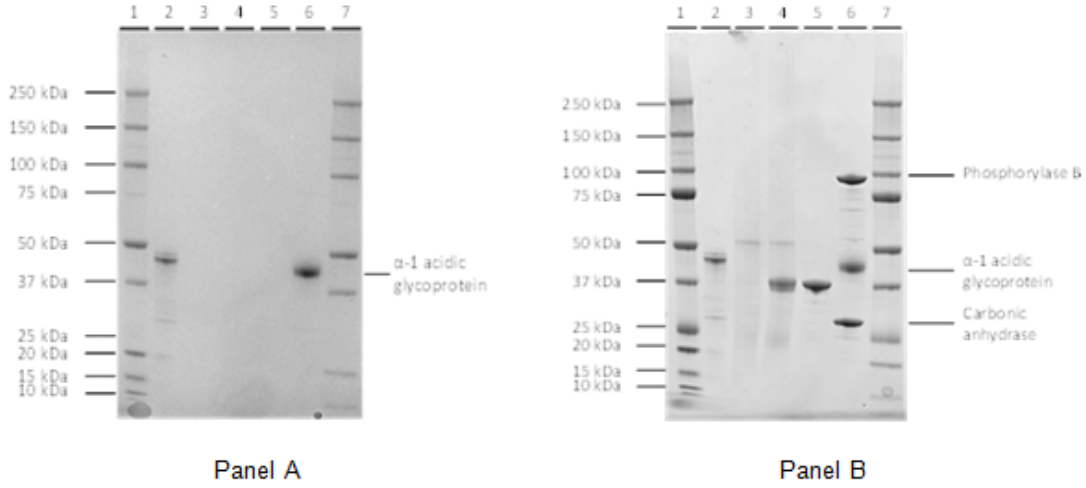


Figure 42: **Assessment and comparison of the of the glycosylation status of the GHB811 cotton-purified HPPD W336 protein and the bacterially-produced HPPD W336 protein**

One Criterion XT Bis-Tris 4-12 % SDS-PAGE gel was prepared and cut into two pieces, each part containing 2 µg HPPD W336 protein of both the bacterially-produced HPPD W336 protein batch 1411_HPPD W336 and the GHB811 cotton-purified HPPD W336 protein together with the appropriate positive and negative controls to assess the glycosylation status.

Panel A shows the result of the staining using the Glycoprotein Detection Kit (Sigma) to demonstrate the absence of glycosylation of both the HPPD W336 protein samples. For the second set of samples, a glycostaining was performed in which the oxidation step in the procedure was omitted to demonstrate the absence of any non-specific binding (data not shown).
Panel B shows a Coomassie staining of the SDS-PAGE gel to demonstrate the presence of the proteins on the gel.

Loading order of Panel A and B:
Lane 1: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)
Lane 2: 45.45 pmoles of Horseradish Peroxidase (positive control)
Lane 3: 10 µL of the non-GM counterpart derived negative control sample (16-RSTHN035-B-03)
Lane 4: 2 µg of GHB811 cotton-purified HPPD W336 protein of sample 16-RSTHN035-B-01
Lane 5: 2 µg of bacterially-produced HPPD W336 protein batch 1411_HPPD W336
Lane 6: 45.45 pmoles of Glycoprotein mix
Lane 7: 5 µL of the Precision Plus Protein™ Dual Xtra Standards (Bio-Rad)

Assessment of the intact molecular mass and peptide mapping

The intact molecular mass was determined using the UPLC-UV-MS and the peptide mapping was established using UPLC-UV-MS^E analysis for the GHB811-purified HPPD W336 protein.

The determined intact molecular mass allowed the identification of two intact molecular masses. The first and major molecular mass (40,446.8 Da) corresponds to a HPPD W336 protein with an N-terminal cysteinic sulfinic acid, derived from the transit peptide (40,447.2 da). The second, minor molecular mass (40,179.9 Da) corresponds to an N-terminal des-Methionine (mature form of the protein minus the initial methionine residue, desMet) HPPD W336 protein (40,180.8 Da).

Peptides resulting from a trypsin digest of the GHB811 cotton-purified HPPD W336 protein were analysed. Figure 43 provides an overview of the mapped peptides against the theoretical amino acid sequence of the HPPD W336 protein. A coverage of 98.6 % was determined, which confirmed the identity of the HPPD W336 protein.

The intact molecular mass was determined and the peptide mapping was established using LC-UV-MS analysis for bacterially-produced HPPD W336 protein.

The determined intact molecular mass (40181.0 Da) confirms the theoretical molecular mass of the protein of 40180.8 Da corresponding with the mass of the HPPD W336 protein minus the methionine residue (desMet).

Peptides resulting from a trypsin digest of bacterially-produced HPPD W336 protein were analysed using LC-UV-MS. Figure 44 provides an overview of the mapped peptides against the theoretical amino acid sequence of the HPPD W336 protein. A coverage of 96.1 % was determined, which confirms the identity of HPPD W336 protein.

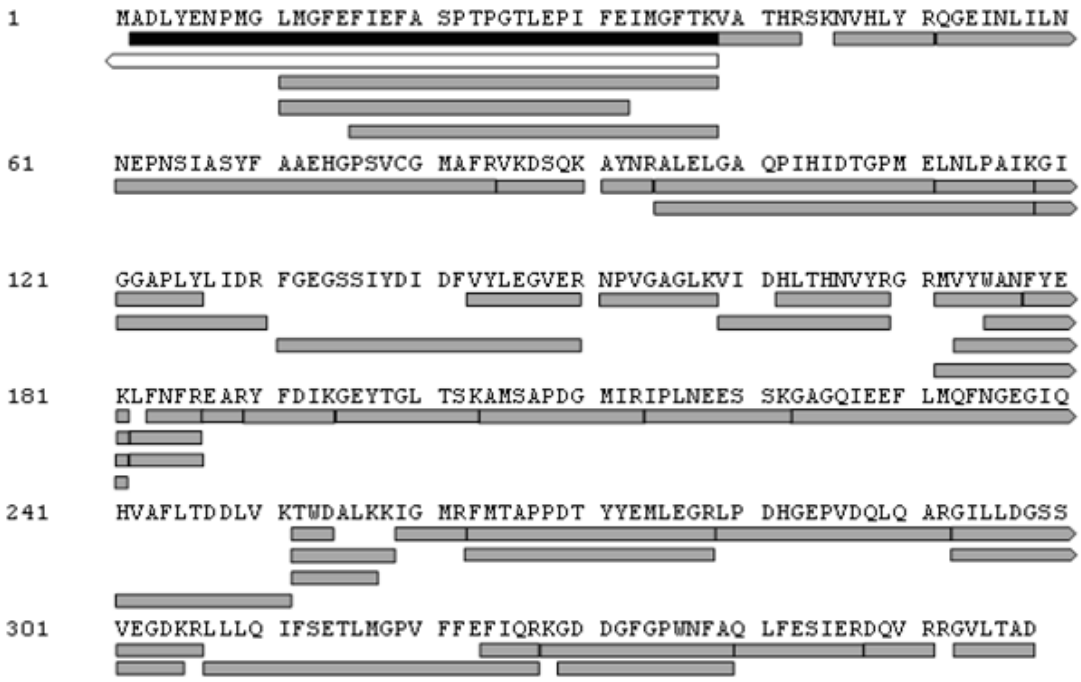


Figure 43: **Schematic overview of the coverage of the theoretical HPPD W336 sequence by the tryptic peptides from the cotton GHB811-purified HPPD W336 protein detected by UPLC-UV-MS^E**

Black and white bars represent the N-terminal peptides corresponding respectively to the desMet HPPD W336 protein and the HPPD W336 protein with an N-terminal cysteinic sulfinic acid, derived from the transit peptide.

Figure 44: Schematic overview of the tryptic peptides detected.

Mapping of the peptides derived from the DTT reduced and trypsin digested HPPD W336 protein sample batch 1411_HPPD W336 against the theoretical amino acid sequence of the HPPD W336 protein

Assessment of the N-terminal sequence

The N-terminal sequence of the GHB811 cotton-purified HPPD W336 protein was determined by Edman degradation. The obtained data for the GHB811 cotton-purified sample suggested the presence of two N-termini. The obtained sequence result could be resolved as ADLYENPMGL, corresponding to the desMet N-terminus (*i.e.* N-terminus without methionine) of the HPPD W336 protein, and sequence XMADLYENPM, potentially corresponding to incomplete cleavage of the transit peptide of the HPPD W336 protein. These results support the data obtained within the intact molecular mass determination.

The N-terminal sequence of bacterially-produced HPPD W336 protein was determined by Edman degradation. The obtained N-terminal sequence (ADLYENPMGL) corresponds to the desMet N-terminus (*i.e.* N-terminus without methionine) of the HPPD W336 protein. This confirms the results observed with the intact molecular mass determination.

Conclusion

The equivalence of the GHB811 cotton-purified HPPD W336 protein with bacterially-produced HPPD W336 protein was demonstrated based on a panel of analytical tests and assays, including densitometry analysis of a Coomassie-stained SDS-PAGE; western blot analysis; glycostaining analysis; *Mass Spectroscopy*; and N-terminal sequence analysis. In addition to the des-Methionine HPPD W336 protein, the GHB811 cotton-purified sample contained a second HPPD W336-derived structure with two additional amino acid residues at the N-terminus, potentially corresponding to incomplete cleavage of the transit peptide of the HPPD W336 protein. This form did not have any impact on the observed characteristics for the GHB811 cotton-purified HPPD W336 protein.

Expression of 2mEPSPS and HPPD W336 in Plant Tissues

The concentration of 2mEPSPS and HPPD W336 proteins in GHB811 fuzzy seed and processed fractions were determined by Enzyme-Linked ImmunoSorbent Assay (ELISA) M-574125-01; Node B.5 (c)).

The quantitation of 2mEPSPS protein in fuzzy seed, linters, delinted seed, untoasted meal, hull, crude oil and refined, bleached, and deodorized (RBD) oil samples was conducted with validated 2mEPSPS ELISA methods using the EnviroLogix QualiPlate™ Kit for 2mEPSPS (Catalog number: AP 084). The quantitation of HPPD W336 protein in fuzzy seed, linter, delinted seed, untoasted meal, hull, crude oil and RBD oil samples was conducted with validated HPPD W336 specific ELISA methods using the EnviroLogix QuantiPlate™ Kit (Catalog Number: AP 128 NW).

The concentrations of the 2mEPSPS protein ranged from 15.53 to 209.98 µg/g DW in GHB811 cotton fuzzy seed, linters, delinted seed, untoasted meal and hull samples (Table 20). The 2mEPSPS protein concentrations in toasted meal, crude oil and RBD oil were below the lower limit of quantification (LLOQ). The concentrations of 2mEPSPS in not treated and treated fuzzy seed were 150.66 and 123.48 µg/g DW, respectively.

The concentrations of the HPPD W336 protein ranged from 4.54 to 42.50 µg/g DW in GHB811 cotton fuzzy seed, linters, delinted seed, untoasted meal and hull samples (Table 21). The HPPD W336 protein concentrations in toasted meal, crude oil and RBD oil were below LLOQ. The concentrations of HPPD W336 protein in not treated and treated fuzzy seed were 42.50 and 28.58 µg/g DW, respectively.

Table 20. Concentrations of 2mEPSPS in GHB811 Cotton Fuzzy Seed and Processed Fractions

Matrix	Trait-Specific Herbicide Treatment	2mEPSPS (µg/g FW)[1]	2mEPSPS (µg/g DW)[1]
Fuzzy seed	Not Treated	138.12	150.66
	Treated	112.99	123.48
Linters	Not Treated	14.73	15.69
	Treated	14.64	15.53
Delinted seed	Not Treated	193.22	209.98
	Treated	193.60	209.87
Untoasted meal	Not Treated	25.56	28.49
	Treated	52.58	58.88
Toasted meal	Not Treated	<LLOQ	<LLOQ
	Treated	<LLOQ	<LLOQ
Hull	Not Treated	60.65	67.01
	Treated	56.67	62.66
Crude oil	Not Treated	<LLOQ	NA
	Treated	<LLOQ	NA
RBD oil	Not Treated	<LLOQ	NA
	Treated	<LLOQ	NA

[1] Values were calculated to full precision and rounded to 2 decimal places for reporting consistency.
LLOQ = Lower Limit of Quantitation
NA = Not Applicable. DW calculations for oil samples were not performed since oil samples were analyzed as received and moisture data for these samples were not obtained.

Table 21. Concentrations of HPPD W336 in GHB811 Cotton Fuzzy Seed and Processed Fractions

Matrix	Trait-Specific Herbicide Treatment	HPPD W336 (µg/g FW)[1]	HPPD W336 (µg/g DW)[1]
Fuzzy seed	Not Treated	38.97	42.50
	Treated	26.15	28.58
Linters	Not Treated	7.87	8.38
	Treated	8.32	8.82
Delinted seed	Not Treated	34.66	37.67
	Treated	28.69	31.10
Untoasted meal	Not Treated	4.07	4.54
	Treated	10.38	11.62
Toasted meal	Not Treated	<LLOQ	<LLOQ
	Treated	<LLOQ	<LLOQ
Hull	Not Treated	15.04	16.62
	Treated	13.07	14.46
Crude oil	Not Treated	<LLOQ	NA
	Treated	<LLOQ	NA
RBD oil	Not Treated	<LLOQ	NA
	Treated	<LLOQ	NA

[1] Values were calculated to full precision and rounded to 2 decimal places for reporting consistency.
LLOQ = Lower Limit of Quantitation
NA = Not Applicable. DW calculations for oil samples were not performed since oil samples were analyzed as received and moisture data for these samples were not obtained.

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

See the relevant parts of Section B.1(a) above on history of safe use and refer to the relevant studies.

HPPD W336 protein

The *hppd* gene was isolated from the bacterium *Pseudomonas fluorescens*, strain A32. *P. fluorescens* is a Gram-negative, rod-shaped, motile, asporogenous, aerobic bacterium. *P. fluorescens*, is ubiquitous in the environment, including soil, water and food (OECD; 1997; M-357528-01; Node A.2 (a), (i)). It has many beneficial uses in agriculture, human health and bioremediation. , It is not described as allergenic, toxic or pathogenic to healthy humans and animals and has an overall history of safe use.

HPPD proteins are ubiquitous in nature across all kingdoms: bacteria, fungi, plants and animals including mammals. HPPD amino acid sequences have been determined in bacteria such as *Streptomyces avermitilis* (Accession number Q53586), in fungi such as *Aspergillus fumigatus* (Accession number Q4WPV8), in plants such as *Arabidopsis thaliana* (Accession number P93836), and in animals such as *Caenorhabditis elegans* (Accession number Q22633), mouse (*Mus musculus*, Accession number P49429), and human (*Homo sapiens*, Accession number P32754).

In particular, HPPD proteins have been characterized in organisms present in human food, such as carrot (*Daucus carota*, Accession number O23920), barley (*Hordeum vulgare* Accession number O48604; Falk *et al*., 2002), pork (*Sus scrofa*, Accession number Q02110)) and beef (*Bos Taurus*, Accession number Q5EA20).

In conclusion, HPPD proteins are present in food from plant, fungal or animal origin, with good safety records. Therefore, HPPD proteins have a history of safe use.

EPSPS proteins

5-enolpyruvylshiklmate-3-phosphate synthase (EPSPS) catalyzes the penultimate step of the shikimate pathway, which is responsible for the synthesis of aromatic amino acids and other aromatic compounds in plants, fungi and microorganisms including apicomplexan parasites (Herrmann, K. M.; 1995; M-269843-01; Node B.1 (a)). As such, it has been shown that EPSPS enzymes are ubiquitous in nature and are present in food and feed derived from plant and microbial sources. No health-related adverse effects have been associated with these proteins.

The *2mepsps* gene was generated by introducing mutations into the *epsps* gene from maize (*Z mays* L.) that result in two amino acid substitutions. The modified EPSPS (2mEPSPS) enzyme has a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of glyphosate herbicides (Lebrun, M. *et al*., 1997). Since the 2mEPSPS protein is derived from maize and has only two amino acid modifications, the safety profile of the novel protein is expected to remain unchanged relative to its wild-type counterpart.

In conclusion, EPSPS proteins are present in food and feed from plant and microbial sources with good safety records. Therefore, EPSPS proteins have a history of safe use.

(c) Information on whether any new protein has undergone any unexpected post-translational modification in the new host

Post-translational modification is determined by glycosylation analysis. For 2mEPSPS and HPPD W336, glycosylation testing was performed on both proteins purified from GHB811 cotton and the microbially-produced proteins. No glycosylation was determined in either the GHB811 plant-purified proteins or the microbially-produced proteins. Therefore, it is deduced that there is no unexpected post-translational modification via glycosylation for 2mEPSPS or HPPD W336 protein.

(d) Where any ORFs have been identified (in subparagraph A.3 (c)(v) of this Guideline (3.5.1)), bioinformatics analyses to indicate the potential for allergenicity and toxicity of the ORFs.

A bioinformatics analysis was performed on the transgenic locus sequence of the GHB811 cotton to identify open reading frames (ORF) M-575144-01; Node B.1 (d)).

The GHB811 transgenic locus, containing the inserted DNA together with the 5' and 3' flanking sequences, was used as query sequence. The GetORF search program was used to identify all ORF crossing a junction or overlapping the inserted DNA, between two translation stop codons, with a minimum size coding for 3 amino acids. This search identified 549 ORF.

In the next step, the translated amino acid sequences from the identified ORF with a minimum size of 30 amino acids were used as query sequences in homology searches to known allergens and toxins. After elimination of duplicates, they represented 126 unique sequences.

Two *in silico* approaches were used to evaluate the potential amino acid sequence identity with known allergens contained in the public allergen database AllergenOnline (www.allergenonline.org):

- An 8-mer search was carried out to identify any short sequences of 8 amino acids or longer that share 100% identity to an allergenic protein. This search was performed using SeqMatchAll from the EMBOSS suite, which compared each ORF sequence with all known allergens present in the allergen database.
- An overall identity search was carried out by using FASTA algorithm, which compared each complete query sequence with all protein sequences present in the AllergenOnline database. The scoring matrix was BLOSUM50. An E-value threshold of 1 was used. The criterion indicating potential relevant identity to an allergen was ≥35% identity over at least 80 amino acids for sequences of ≥80 amino acids, or ≥35% recalculated over a hypothetical 80 amino acid window for sequences of <80 amino acids.

In addition, each query sequence was evaluated for potential identity with known toxins. An overall identity search was carried out by using FASTA algorithm with all protein sequences present in the NCBI non-redundant database, using the BLOSUM50 scoring matrix. An E-value threshold of 0.1 was used for pre-selecting the most identical proteins. The biological relevance of the matches was further assessed. Biologically relevant matches provide insight on the familiarity and potential toxic properties of the potential polypeptide.

The 8-mer search showed no 100% identity with known allergenic proteins. The overall search showed no biologically relevant identity between the query sequences and any known allergenic proteins.

In addition, no biologically relevant identities were found with any toxic protein from the NCBI non-redundant database.

In conclusion, there are neither allergenic nor toxicological *in silico* findings associated with the potential ORF polypeptides.

B.2 New proteins

If it can be shown the new protein(s) is identical to one previously assessed by FSANZ, the only other safety information that must be provided is an updated bioinformatics comparison of the amino acid sequence to known protein toxins, anti-nutrients and allergens.

Where the new protein is not identical to one previously assessed by FSANZ, the following must be provided:

(a) Information on the potential toxicity of any new proteins, including:

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

The 2mEPSPS and HPPD W336 proteins have been assessed for food safety previously by FSANZ as part of the assessments for GHB614 cotton (A614) and FG72 soybean (A1051) approval. Information has been provided above under Section B.1 (a) to show equivalence of 2mEPSPS and HPPD W336 proteins expressed by GHB811 cotton to the same proteins expressed by FG72 soybean and also microbially-produced proteins based on the known amino acid sequences of 2mEPSPS and HPPD W336. Therefore in this section, the only supplementary information on the proteins provided is updated bioinformatics analysis for these known proteins based on the amino acid sequences associated with the proteins.

2mEPSPS protein

The potential amino acid sequence homology of the double mutated maize 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) protein with known allergens and known toxins was evaluated by using several *in silico* approaches M-445681-04; Node B.2 (a), (i)).

This search evaluated the potential amino acid sequence identity of the query protein with known allergens by using two *in silico* approaches.

- An overall identity search was carried out to compare the complete query sequence with all protein sequences present in the public allergen database AllergenOnline (www.allergenonline.org). The FASTA algorithm was used, with the BLOSUM50 scoring matrix and an E-value threshold of 10. The criterion indicating potential allergenicity was ≥35% identity over at least 80 consecutive amino acids with an allergenic protein.
- An 8-mer search was carried out to identify any short sequences of 8 amino acids or longer that share 100% identity to an allergenic protein. This search was performed using SeqMatchAll from the EMBOSS suite, which compared the query sequence with all known allergens present in the allergen database.

Furthermore, this study considered the potential N-glycosylation sites by searching their known consensus sequences, potentially found in allergenic proteins.

In addition, two *in silico* approaches based on the FASTA algorithm associated with the BLOSUM50 scoring matrix were used to evaluate the potential amino acid sequence identity of the query protein with known toxins:

- An overall identity search with all protein sequences present in the NCBI non-redundant database. An E-value threshold of 0.1 was used for pre-selecting the most similar proteins. The biological relevance of the matches was further assessed. Biologically relevant matches provide insight on the familiarity and potential toxic properties of the query protein.
- An overall identity search with all protein sequences present in the in-house Bayer toxin database. An E-value threshold of 10 was used for pre-selecting the most identical toxins. The biological relevance of the matches was further assessed. Biologically relevant matches provide insight on the potential toxic properties of the query protein.

The overall identity search showed no biologically relevant identity between the query protein and any known allergenic proteins. In addition, the 8-mer search showed no 100% identity with known allergenic proteins.

Glycosylation is one of the principal co-translational and post-translational modifications of various membrane-bound and secreted proteins. The attachment of saccharides to target proteins is thought to enhance protein folding and stability. Some food allergens are N-glycosylated, therefore it is possible that glycosyl groups may contribute to allergenicity (Huby *et al*., 1995; Jenkins *et al*., 1996). However, many allergens are not glycosylated, and a large number of non-allergens are glycoproteins, indicating the glycosylation is neither necessary nor sufficient for allergenicity. It is therefore important to carefully interpret and confirm the results of glycosylation analyses in the safety assessment of a novel protein.
The best studied mode of glycosylation is the formation of an N-glycosidic linkage to Asparagin in the polypeptide chain. One criterion for protein N-glycosylation is the presence of the sequence N-X~(P)-S/T, where N = Asparagin, X~(P) = any amino acid except Proline (P), S = Serin and T = Threonine, in the query sequence. Another potential receptor site is N-X-C, where N = Asparagin, X = any amino acid and C = Cysteine. Therefore, the consensus sequences searched for in this analysis were N-X~(P)-[S/T] or N-X-C.

Two potential N-glycosylation sites were identified on the amino acid sequence of the query protein. However, the presence of these sites is neither necessarily predictive of a potential glycosylation of the protein *in planta* nor of a potential allergenicity.

As expected, the overall homology search against the general protein database showed that most of the matches corresponded to EPSPS sequences from various organisms. There is no record of potential toxicity associated with these proteins. Furthermore, no identities were found with any toxic proteins from the Bayer toxin database.

In conclusion, there are neither allergenic nor toxicological *in silico* findings associated with the 2mEPSPS protein.

HPPD W336 protein

The potential amino acid sequence homology of the single mutated 4-hydroxyphenylpyruvate dioxygenase (HPPD) protein with known allergens and known toxins was evaluated by using several *in silico* approaches M-445678-04; Node B.2 (a) (i)).

This search evaluated the potential amino acid sequence identity of the query protein with known allergens by using two *in silico* approaches.

- An overall identity search was carried out to compare the complete query sequence with all protein sequences present in the public allergen database AllergenOnline (www.allergenonline.org). The FASTA algorithm was used, with the BLOSUM50 scoring matrix and an E-value threshold of 1. The criterion indicating potential allergenicity was ≥35% identity over at least 80 consecutive amino acids with an allergenic protein.
- An 8-mer search was carried out to identify any short sequences of 8 amino acids or longer that share 100% identity to an allergenic protein. This search was performed using SeqMatchAll from the EMBOSS suite, which compared the query sequence with all known allergens present in the allergen database.

Furthermore, this study considered the potential N-glycosylation sites by searching their known consensus sequences, potentially found in allergenic proteins.

In addition, two *in silico* approaches based on the FASTA algorithm associated with the BLOSUM50 scoring matrix were used to evaluate the potential amino acid sequence identity of the query protein with known toxins:

- An overall identity search with all protein sequences present in the NCBI non-redundant database. An E-value threshold of 0.1 was used for pre-selecting the most similar proteins. The biological relevance of the matches was further assessed. Biologically relevant matches provide insight on the familiarity and potential toxic properties of the query protein.
- An overall identity search with all protein sequences present in the in-house Bayer toxin database. An E-value threshold of 10 was used for pre-selecting the most identical toxins. The biological relevance of the matches was further assessed. Biologically relevant matches provide insight on the potential toxic properties of the query protein.

The overall identity search showed no biologically relevant identity between the query protein and any known allergenic proteins. In addition, the 8-mer search showed no 100% identity with known allergenic proteins.

No potential N-glycosylation sites were identified on the amino acid sequence of the query protein.

As expected, the overall homology search against the general protein database showed that, in most cases, the HPPD W336 protein matched with other HPPD proteins from various origins, which have safety records. In addition, no significant similarities were found with any toxic protein from the Bayer toxin database.

In conclusion, there are neither allergenic nor toxicological *in silico* findings associated with the HPPD W336 protein.

(ii) Information on the stability of the protein to proteolysis in appropriate gastrointestinal model systems

Information to establish the stability of the 2mEPSPS and HPPD W336 proteins to proteolysis have been presented previously to FSANZ in association with the submissions for approval of GHB614 cotton (A614) and FG72 soybean (A1051). These data are not presented again in this submission.

(iii) An animal toxicity study if the bioinformatics comparison and biochemical studies indicate either a relationship with known protein toxins/anti-nutrients or resistance to proteolysis.

FSANZ concluded in their 2011 Safety Assessment Report for FG72 soybean the following for 2mEPSPS and HPPD W336 protein, “Both proteins exhibit a degree of heat stability however, given their digestive lability combined with their lack of similarity to known protein toxins or allergens and the loss of enzyme activity with heating, this does not raise any safety concerns.” (FSANZ, 2011) On this basis it is considered that no further data on animal toxicity is required when this assessment of the ability of 2mEPSPS and HPPD W336 protein to resist proteolysis is taken in combination with updated bioinformatics data for the two proteins presented in Section B.2 (a) (i) above. In addition to this information, acute toxicity studies for the two proteins were previously assessed by FSANZ (A1051), the data confirming absence of toxicity in animals (FSANZ, 2011).

(b) Information on the potential allergenicity of any new proteins, including:

(i) Source of the new protein

2mEPSPS protein

The 2mEPSPS protein is not a new protein. It has been assessed previously for food safety as part of the FSANZ approvals of GHB614 cotton (A614) and FG72 soybean (A1051).

The coding sequence of 5-enol pyruvylshikimate-3-phosphate synthase (*epsps*) gene was isolated from maize (*Zea Mays* L.) Two amino acids were substituted (threonine by isoleucine at position 102 and proline by serine at position 106) (Lebrun, M. *et al*., 1997; M-216526-01; Node A.1 (a)). These modifications confer to the protein a decreased binding affinity for glyphosate, allowing it to maintain sufficient enzymatic activity in the presence of the herbicide. Therefore, the plants expressing this modified protein become tolerant to glyphosate herbicides (Lebrun, M. *et al*., 1997; M-216526-01; Node A.1 (a)). The modified protein is designated as 2mEPSPS.

The coding sequence of 5-enol pyruvylshikimate-3-phosphate synthase (*epsps*) gene was isolated from maize (*Zea Mays* L.). Maize is grown in nearly all areas of the world over a wide range of climatic conditions. Because of its high levels of starch, protein, oil and other nutritionally valuable substances, maize is an important crop in human and animal nutrition.

As described in the food allergen labeling lists from various regulatory authorities, most immunoglobulin E (IgE)-mediated food allergies are attributable to a small group of 8-12 food groups. For example, it is estimated that these foods or food groups account for more than 90% of all food allergies in the USA.

Corn is not part of these food allergen labeling lists ; M-402872-01; Node B.2 (b) (i)). Despite its wide consumption, corn is considered by allergy experts as an uncommon allergenic food. Clinical symptoms of maize food allergy range from oral allergy syndrome to anaphylaxis. An extensive literature review revealed that only few maize allergenic proteins have been identified, sequenced and recorded in allergen databases.

Overall, maize is considered by allergy experts to be a very uncommon allergenic food M-402872-01; Node B.2 (b) (i); M-278169-01; Node B.2 (b) (i)).

HPPD W336 protein

The HPPD W336 protein is not a new protein. It has been assessed previously for food safety as part of the FSANZ approval of FG72 soybean (A1051).

The coding sequence of the 4-hydroxyphenylpyruvate dioxygenase (HPPD) protein was isolated from the *Pseudomonas fluorescens* strain A32. One amino acid was substituted (glycine at position 336 with tryptophan) to improve the tolerance against HPPD inhibitors. The modified protein is designated as HPPD W336 (Boudec, P.; et al.; 2001; M-229534-01; Node A.1 (a)).

The *hppd* gene was isolated from the bacterium *Pseudomonas fluorescens*, strain A32. *P. fluorescens* is a Gram-negative, rod-shaped, motile, asporogenous, aerobic bacterium. *P. fluorescens*, is ubiquitous in the environment, including soil, water and food (OECD; 1997; M-357528-01; Node A.2 (a) (i)). It has many beneficial uses in agriculture, human health and

bioremediation. It is not described as allergenic, toxic or pathogenic to healthy humans and animals and has an overall history of safe use.

HPPD proteins are ubiquitous in nature across all kingdoms: bacteria, fungi, plants and animals including mammals. HPPD amino acid sequences have been determined in bacteria such as *Streptomyces avermitilis* (Accession number Q53586), in fungi such as *Aspergillus fumigatus* (Accession number Q4WPV8), in plants such as *Arabidopsis thaliana* (Accession number P93836), and in animals such as *Caenorhabditis elegans* (Accession number Q22633), mouse (*Mus musculus*, Accession number P49429), and human (*Homo sapiens*, Accession number P32754).

In particular, HPPD proteins have been characterized in organisms present in human food, such as carrot (*Daucus carota*, Accession number O23920), barley (*Hordeum vulgare* Accession number O48604; Falk *et al*., 2002), pork (*Sus scrofa*, Accession number Q02110)) and beef (*Bos Taurus*, Accession number Q5EA20).

In conclusion, HPPD proteins are present in food from plant, fungal or animal origin, with good safety records. Therefore, HPPD proteins have a history of safe use.

(ii) A bioinformatics comparison of the amino acid sequence of the novel protein to known allergens

HPPD W336 protein

The potential amino acid sequence homology of the single mutated 4-hydroxyphenylpyruvate dioxygenase (HPPD) protein with known allergens and known toxins was evaluated by using several *in silico* approaches M-445678-04; Node B.2 (a) (i)). The nature of the bioinformatics comparison with known allergens was discussed in further detail above in Section B.2 (a) (i).

2mEPSPS protein

The potential amino acid sequence homology of the double mutated maize 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) protein with known allergens and known toxins was evaluated by using several *in silico* approaches M-445681-04; Node B.2 (a) (i)). The nature of the bioinformatics comparison with known allergens was discussed in further detail above in Section B.2 (a) (i).

(iii) The new protein's structural properties, including, but not limited to, its susceptibility to enzymatic degradation (e.g. proteolysis), heat and/or acid stability

Please refer to Section B.1 (d) (ii) above for information on enzymatic degradation, heat and acid stability for the 2mEPSPS and HPPD W336 proteins.

(iv) Specific serum screening where a new protein is derived from a source known to be allergenic or has sequence homology with a known allergen

The 2mEPSPS protein is derived from *Zea mays* L. (maize). Overall, maize is considered by allergy experts to be a very uncommon allergenic food M-402872-01; Node B.2 (b) (i); M-278169-01; Node B.2 (b) (i)). In the most recent bioinformatics analysis performed on the 2mEPSPS protein (M-445681-04; Node B.2 (a) (i)), it was not found to display amino acid sequence homology with known allergens.

The HPPD W336 protein is not from a source known to be allergenic nor does it display sequence homology with known allergens.

(v) Information on whether the new protein(s) have a role in the elicitation of gluten-sensitive enteropathy, in cases where the introduced genetic material is obtained from wheat, rye, barley, oats, or related cereal grains

Not applicable. The introduced genetic materail is not obtained from wheat, rye, barley, oats or related cereal grains.

Where the new protein has been produced from an alternative source (e.g. microbial expression system) in order to obtain sufficient quantities for analysis, information must be provided to demonstrate that the protein tested is biochemically, structurally and functionally equivalent to that expressed in the food produced using gene technology.

2mEPSPS and HPPD W336 proteins

Refer to Section B.1 (a), under equivalence of proteins expressed *in planta* with those produced microbially for evidence that proteins that have been tested are biochemically, structurally and functionally equivalent to that expressed in the food produced using gene technology.

Information on the potential toxicity and potential allergenicity of a newly expressed protein is also not required if:

(a) The protein is expressed from a transferred gene that is derived from the same species as the host or a species that is cross-compatible with the host, provided evidence is provided to demonstrate the following:

(i) The gene donor belongs to a species that is commonly used as food and has a history of safe use

In the case of the *2mepsps* gene is derived from *Zea ma*ys L (maize). The wild-type maize 5-enol pyruvylshikimate-3-phosphate synthase (*epsps*) gene was mutated using site directed mutagenesis. The wild-type maize *epsps* gene was mutated at positions 102 (substitution of threonine by isoleucine) and position 106 (substitution of proline by serine) (Lebrun, M. *et al*.,

1997; M-216526-01; Node A.1 (a); M-234184-01; Node A.3 (c) (iii) (CCI)). Maize, both wild type and genetically modified, has a history of safe use as a food.

The gene donor for the *hppdPf W336* gene, although not derived from a species commonly used as food, results in the expression of HPPD proteins that have been characterized in organisms present in human food, such as carrot (*Daucus carota*, Accession number O23920), barley (*Hordeum vulgare* Accession number O48604; Falk *et al*., 2002), pork (*Sus scrofa*, Accession number Q02110)) and beef (*Bos Taurus*, Accession number Q5EA20).

(ii) The protein is expressed at levels in the new food produced using gene technology that are consistent with the levels in the gene donor.

Not applicable.

(b) Evidence is provided to demonstrate the absence of the newly expressed protein from the parts of the host organism consumed as food.

The concentration of 2mEPSPS and HPPD W336 proteins in GHB811 fuzzy seed and processed fractions were determined by Enzyme-Linked ImmunoSorbent Assay (ELISA) ; M-574125-01; Node B.5 (c)).

The quantitation of 2mEPSPS protein in fuzzy seed, linters, delinted seed, untoasted meal, hull, crude oil and refined, bleached, and deodorized (RBD) oil samples was conducted with validated 2mEPSPS ELISA methods using the EnviroLogix QualiPlate™ Kit for 2mEPSPS (Catalog number: AP 084). The quantitation of HPPD W336 protein in fuzzy seed, linter, delinted seed, untoasted meal, hull, crude oil and RBD oil samples was conducted with validated HPPD W336 specific ELISA methods using the EnviroLogix QuantiPlate™ Kit (Catalog Number: AP 128 NW).

The concentrations of the 2mEPSPS protein ranged from 15.53 to 209.98 µg/g DW in GHB811 cotton fuzzy seed, linters, delinted seed, untoasted meal and hull samples (Table 20). The 2mEPSPS protein concentrations in toasted meal, crude oil and RBD oil were below the lower limit of quantification (LLOQ). The concentrations of 2mEPSPS in not treated and treated fuzzy seed were 150.66 and 123.48 µg/g DW, respectively.

The concentrations of the HPPD W336 protein ranged from 4.54 to 42.50 µg/g DW in GHB811 cotton fuzzy seed, linters, delinted seed, untoasted meal and hull samples (Table 21). The HPPD W336 protein concentrations in toasted meal, crude oil and RBD oil were below LLOQ. The concentrations of HPPD W336 protein in not treated and treated fuzzy seed were 42.50 and 28.58 µg/g DW, respectively.

B.3 Other (non-protein) new substances

If other (non-protein) substances are produced as a result of the introduced DNA, information must be provided on the following:

(a) The identity and biological function of the substance

Non-protein substances cannot be created from DNA. The central maxim of molecular biology is that DNA makes RNA and RNA makes protein. Therefore, no non-protein substances could be created from the introduction of the DNA insert.

(b) Whether the substance has previously been safely consumed in food

Not relevant.

(c) Potential dietary exposure to the substance

Not relevant.

(d) Where RNA interference has been used:

(i) The role of any endogenous target gene and any changes to the food as a result of silencing that gene

Not applicable. RNA interference has not been used.

(ii) The expression levels of the RNA transcript

Not applicable. RNA interference has not been used.

(iii) The specificity of the RNA interference

Not applicable. RNA interference has not been used.

B.4 Novel herbicide metabolites in GM herbicide-tolerant plants

Data must be provided on the identity and levels of herbicide and any novel metabolites that may be present in the food produced using gene technology.

If novel metabolites are present then the application should address the following, where appropriate:

(a) Toxicokinetics and metabolism

(b) Acute toxicity

(c) Short-term toxicity

(d) Long-term toxicity and carcinogenicity

(e) Reproductive and developmental toxicity

(f) Genotoxicity

Isoxaflutole (IFT), 5-cyclopropyl-4-isoxazoly [2-(methylsulfonyl)-4-(trifluoromethyl)phenyl] methanone, is being developed for pre-emergent and early post-emergent control of broadleaf and grass weeds. The mode of herbicidal action of IFT is the inhibition of p-hydroxyphenyl pyruvatedioxygenase (HPPD). The HPPD enzyme is involved in the catabolism of the aromatic amino acid tyrosine. It catalyzes the second step which consists of the formation of homogentisate (HG) and CO_2 out of 4-hydroxyphenylpyruvate (4-HPP) and O_2. Basically, this pathway yields the products acetoacetate and fumarate which have a direct energetic contribution for both prokaryotes and eukaryotes. The blocking of HG formation by use of HPPD-inhibiting herbicides such as isoxaflutole (IFT) depletes the available pools of end products leading to foliage bleaching through the destabilization of the photosynthetic apparatus (Matringe *et al.*, 2005). This highly effective herbicide target site has raised interest in the development of herbicide tolerant crops. A successful strategy to

achieve this is through the expression of altered HPPD enzymes, such as the HPPD enzyme from *Pseudomonas fluorescens* which has been modified by site-directed mutagenesis at position 336 (HPPD W336).

The nature of the IFT derived residues in HPPD tolerant soybean event FG72 expressing the HPPD W336 as the result of a pre- or a post-emergent application of IFT was well characterized (Document M-368555-01-1; A1051). It showed no novel metabolites were formed when IFT was applied to HPPD tolerant soybean, compared to those metabolites revealed in earlier studies performed with IFT in other crops. Therefore, it is expected no new metabolites to be formed in cotton varieties including GHB811 cotton expressing the HPPD W336 when IFT is applied, compared to those metabolites revealed in FG72 soybean.

Regarding metabolites associated with glyphosate application to GHB811 cotton, the 2mEPSPS protein is also expressed by the *2mepsps* gene within cotton event GHB614 (FSANZ, A614). The nature of the glyphosate derived residues in cotton event GHB614 expressing the 2mEPSPS as the result of glyphosate application was well characterized (Kowite, 2006a, Appendix 16 within dossier for FSANZ A614; Document M-279965-01). On the basis of the nature of glyphosate metabolism within GHB614 cotton, it is anticipated that GHB811 cotton will reveal the same metabolites when glyphosate is applied due to the *2mepsps* gene that is shared between the two events.

B.5 Compositional Analyses of the Food Produced Using Gene Technology

(a) The levels of key nutrients, toxicants and anti-nutrients in the food produced using gene technology compared with the levels in appropriate comparator (usually the non-GM counterpart). A statistical analysis of the data must be provided.

Composition analyses were conducted to determine levels of key nutrients and anti-nutrients of GHB811 cotton and compare those results to the non-GM counterpart and non-GM reference cotton varieties (M-566678-01; Node B.5 (a)).

Field Production

A total of 15 field trials were successfully completed during 2014 and 2015 using typical commercial agricultural production practices. Of the 15 sites that successfully completed field production, eight sites were selected to represent a distribution of trials conducted in 2014 and 2015, a wide geographical distribution clustered in primary US cotton production areas and representative of field trial management with respect to regional commercial cotton production standards. Composition analysis was conducted on samples collected from these eight field trial sites shown in the table below.

Field Trial Sites

Year	Site Code	Nearest Town or City	State	County or Parish
2014	03	Kerman	California	Fresno
	07	Chula	Georgia	Tift
	09	Cheneyville	Louisiana	Rapides
	10	Greenville	Mississippi	Washington
	11	Elko	South Carolina	Barnwell
2015	15	Wall	Texas	Tom Green
	17	Hertford	North Carolina	Perquimans
	21	Edmonson	Texas	Hale

In addition to the GHB811 cotton and its non-GM counterpart, seven reference varieties that represent the natural variability existing in cotton were included in this study to provide reference ranges for the composition assessment. Each field trial site planted three of the seven reference varieties. The entries included are presented in the table below.

Description of Entries

Entry ID	Description	Background	Trait-Specific Herbicide Treatment	Seed Lot Number (Year)	Sites
A	Non-GM Counterpart (Coker 312)	Coker 312	Not Treated	12PRGH050001 (2014) 14SHGH500001 (2015)	All
J	GHB811	Coker 312	Not Treated	13WAGH01252 (2014) 14SHGH000603 (2015)	All
K	GHB811	Coker 312	Treated	13WAGH01252 (2014) 14SHGH000603 (2015)	All
B	FM958	Non-GM	Not Treated	12LUGH000332 (2014)	03
				14SHGH500002 (2015)	15, 21
C	FM989	Non-GM	Not Treated	12LUGH000334 (2014)	03
				14SHGH500004 (2015)	15, 21
D	ST457	Non-GM	Not Treated	12LUGH000336 (2014)	07, 09, 10, 11
				14SHGH500005 (2015)	17
E	DP399	Non-GM	Not Treated	14LUGH000002 (2014 and 2015)	07, 09, 10, 11, 17
F	ST468	Non-GM	Not Treated	13WAGH03142 (2014)	07, 09, 10, 11
				14SHGH500007 (2015)	17
G	Acala Maxxa	Non-GM	Not Treated	13WAGH02234 (2014)	03
N	FM966	Non-GM	Not Treated	14SHGH500003 (2015)	15, 21

Entries were replicated four times in a randomized complete block design at field trial sites as shown in the table above. The entries were randomly assigned to plots at each field trial site independently by the eStudy™ electronic notebook software.

Conventional herbicide management (CHM) was applied to all entries. The GHB811 cotton plots treated with trait-specific herbicides (Entry K) received one application of isoxaflutole at a rate of 100.3 to 115.2 grams active ingredient per hectare (g ai/ha) at BBCH Growth Stage BBCH 00–13 and one application of glyphosate at a rate of 1067 to 1222 g ai/ha at BBCH Growth Stage 16–19.

Seed cotton samples were harvested, without bias, from all plots at crop maturity and ginned to produce fuzzy seed for composition analysis. Samples were shipped frozen to Bayer CropScience, Morrisville, North Carolina, USA where they were placed in frozen storage (-5º C or lower).

Composition Analysis

The composition analysis of the cotton fuzzy seed samples was conducted at EPL Bio Analytical Services (Niantic, IL, USA). The samples were pre-ground and then completely homogenized using an Ultra Centrifugal Mill homoge. They were maintained at a temperature of approximately -20 °C for the duration of the experimental phase, except when removed from the freezer for homogenization, sample preparation or analysis.

Composition analytes, units and EPL Bio Analytical Services method mnemonics are presented in the table below. The analytical methods and reference standards are detailed in the Appendix 1 of the composition assessment report (; M-566678-01; Node B.5 (a)).

Composition Analytes, Units and Methods for Cotton Fuzzy Seed

Analyte	Units	EPL Method
Proximates and Fiber		
Moisture	% FW	NC-4
Ash	% FW, DW	NC-2
Carbohydrates		Calculated (NC-494)
Crude Fat		NC-230
Crude Protein		NC-20
Acid Detergent Fiber		NC-3
Neutral Detergent Fiber		NC-9
Total Dietary Fiber		NC-359
Amino Acids		
Alanine, Arginine, Aspartic Acid, Glutamic Acid, Glycine, Histidine, Isoleucine, Leucine, Lysine, Phenylalanine, Proline, Serine, Threonine, Tyrosine, Valine	% FW, DW	NC-58
Cystine, Methionine		NC-279
Tryptophan		NC-22
Fatty Acids		
C8:0, C10:0, C12:0, C14:0, C14:1, C15:0, C15:1, C16:0, C16:1, C17:0, C17:1, C18:0, C18:1, C18:2, C18:3,C18:4, C20:0, C20:1, C20:2, C20:3, C20:4, C20:5, C22:0, C22:1, C22:5 n-3,C22:5 n-6, C22:6, C24:0	% FW, DW, % Total Fatty Acids	NC-319
Minerals and Alpha Tocopherol		
Calcium, Copper, Iron, Magnesium, Manganese, Phosphorus, Potassium, Sodium, Zinc	mg/kg FW, DW	NC-60
Alpha Tocopherol (Vitamin E)	mg/kg FW, DW	NC-346
Anti-nutrients		
Free Gossypol	% FW, DW	NC-37
Total Gossypol		NC-36
Cyclopropenoid Fatty Acids		
Dihydrosterculic Acid	% FW, DW % Total Fatty Acids	NC-231
Sterculic Acid		
Malvalic Acid		

Statistical Analysis

Composition data generated from eight sites were used for statistical analysis. The data were read into SAS and statistical analyses were performed using SAS version 9.3 (SAS, 2011).

Analytes, for which more than one third of the values were less than the limit of quantitation (< LOQ), were excluded from statistical evaluation (i.e., ANOVA and mean comparisons) and discussion. These analytes are presented in Table 22. Minimum and maximum values are presented in Table 23 – Table 27 for analytes where some values are above LOQ, but there is insufficient data for statistical analysis.

Combined-site Analysis

Data for the combined-site analysis for each of 54 quantifiable analytes measured for cotton fuzzy seed included mean and standard deviation for Entry A, Entry J and Entry K and the minimum and maximum values for the seven cotton reference varieties (Entries B–G, N). Also included are tolerance intervals calculated for each analyte based on the reference varieties over all combined sites. The tolerance intervals are specified to contain 99% of the population with 95% confidence. In addition, pairwise comparisons were made between the non-GM counterpart (Entry A) and the GM variety not treated with trait-specific herbicides (Entry J) and between the non-GM counterpart (Entry A) and the GM variety treated with trait-specific herbicides (Entry K). A significant difference was noted when the t-test p-value between the comparators was <0.05.

By-site Analysis

For each composition analyte, the by-site analysis was performed using a mixed model analysis of variance with the fixed entry effect and the random block effect, followed by pairwise t-tests comparing entries Entry A vs Entry J and Entry A vs Entry K. A summary of the analytes, by analyte category, indicating the number of sites with significant differences for each analyte is presented in the report M-566678-01; Node B.5 (a)).

.

Table 22. Parameters with Values Below the Limit of Quantitation (LOQ)

Parameter	Number of Values >= LOQ	Number of Values < LOQ	LOQ value	Unit	Excluded from analysis
Tyrosine	189	3	0.313	% DW	No
C8:0 Caprylic Acid	0	192	0.0197	% Total Fatty Acids	Yes
C10:0 Capric Acid	0	192	0.0200	% Total Fatty Acids	Yes
C12:0 Lauric Acid	1	191	0.00505	% Total Fatty Acids	Yes
C14:1 Myristoleic Acid	0	192	0.0203	% Total Fatty Acids	Yes
C15:0 Pentadecanoic Acid	0	192	0.0204	% Total Fatty Acids	Yes
C15:1 Pentadecenoic Acid	0	192	0.0204	% Total Fatty Acids	Yes
C17:1 Heptadecenoic Acid	175	17	0.00513	% Total Fatty Acids	No
C18:4 Stearidonic Acid	46	146	0.00514	% Total Fatty Acids	Yes
C20:1 Eicosenoic Acid	191	1	0.00517	% Total Fatty Acids	No
C20:2 Eicosadienoic Acid	2	190	0.00517	% Total Fatty Acids	Yes
C20:3 Eicosatrienoic Acid	0	192	0.0207	% Total Fatty Acids	Yes
C20:4 Arachidonic Acid	0	192	0.0207	% Total Fatty Acids	Yes
C20:5 Eicosapentaenoic Acid	2	190	0.00516	% Total Fatty Acids	Yes
C22:1 Erucic Acid	0	192	0.0207	% Total Fatty Acids	Yes
C22:5 N3 Docosapentaenoic Acid	111	81	0.00518	% Total Fatty Acids	Yes
C22:5 N6 Docosapentaenoic Acid	103	89	0.00518	% Total Fatty Acids	Yes
C22:6 Docosahexaenoic Acid	0	192	0.00518	% Total Fatty Acids	Yes
C24:0 Lignoceric Acid	142	50	0.0104	% Total Fatty Acids	No
Malvalic Acid	178	14	0.00476	% Total Fatty Acids	No
Sterculic Acid	176	16	0.00477	% Total Fatty Acids	No

Results

Proximates and Fiber in Cotton Fuzzy Seed (Table 23**).**
No significant differences were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K) for moisture, ash, carbohydrates, crude fat, acid detergent fiber, and total dietary fiber.

No significant differences were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated with trait-specific herbicides (Entry J) for neutral detergent fiber. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K) for crude protein. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 cotton treated with trait-specific herbicides (Entry K) for neutral detergent fiber.

However, the means for all entries, for all proximates and fiber, were within the range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Amino Acids in Cotton Fuzzy Seed (Table 24**).**
No statistically significant differences were observed between the non-GM counterpart (Entry A) and GHB811 COTTON not treated or treated with trait-specific herbicides (Entry J and Entry K) for alanine, arginine, aspartic acid, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine and valine.

No statistically significant differences were observed between the non-GM counterpart (Entry A) and GHB811 COTTON not treated with trait-specific herbicides (Entry J) for cystine and methionine. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 cotton treated with trait-specific herbicides (Entry K) for cystine and methionine.

However, the means for all entries, for all amino acids, were within the range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Fatty Acids in Cotton Fuzzy Seed (Table 25**).**
No statistically significant differences were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K) for myristic, palmitic, heptadecanoic, heptadecenoic, oleic, linoleic, linolenic, eicosenoic, behenic and lignoceric acids.

No significant differences were observed between the non-GM counterpart (Entry A) and GHB811 COTTON not treated with trait-specific herbicides (Entry J) for arachidic acid. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 COTTON not treated or treated with trait-specific herbicides (Entry J and Entry K) for palmitoleic and stearic acids. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 cotton treated with trait-specific herbicides (Entry K) for arachidic acid.

However, the means for all entries, for all fatty acids, were within the range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Minerals and Alpha Tocopherol in Cotton Fuzzy Seed (Table 26**).**
No statistically significant differences were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K) for calcium, copper, iron, magnesium, manganese, phosphorus, potassium, sodium and zinc.

Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K) for alpha tocopherol.

However, the means for all entries, for all minerals and alpha tocopherol, were within the range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Anti-nutrients in Cotton Fuzzy Seed (Table 27**).**
No statistically significant differences were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K) for malvalic acid and sterculic acid.

No statistically significant differences were observed between the non-GM counterpart (Entry A) and GHB811 cotton not treated with trait-specific herbicides (Entry J) for dihydrosterculic acid. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 COTTON not treated or treated with trait-specific herbicides (Entry J and Entry K) for free gossypol and total gossypol. Statistically significant differences ($p < 0.05$) were observed between the non-GM counterpart (Entry A) and GHB811 cotton treated with trait-specific herbicides (Entry K) for dihydrosterculic acid.

However, the means for all entries, for all anti-nutrients, were within the range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Of the 69 composition analytes, 54 had sufficient levels above LOQ for statistical analysis. Of the 54 analytes that were analyzed, statistically significant differences ($p < 0.05$) were observed for 11 analytes, six of which were statistically different between both the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K), and five of which were statistically different between the non-GM counterpart (Entry A) and GHB811 cotton treated with trait-specific herbicides (Entry K). However the means of all analytes were within the range of the reference varieties and the

tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant.

Table 23. Comparison of Proximates and Fiber in Fuzzy Seed of GHB811 Cotton with its Non-GM Counterpart[a]

Parameter	Non-GM Counterpart (Entry A) Mean ± SD	GHB811 Not Treated (Entry J) Mean ± SD	GHB811 Treated (Entry K) Mean ± SD	Non-GM Reference Varieties Range (Entries B-G,N) (Min-Max)	Tolerance Interval Non-GM Reference Varieties (Entries B-G,N) [b] (Lower-Upper)	Comparison t-test A vs J [c] p-value	Comparison t-test A vs K [c] p-value
Moisture (% FW)	10.8 ± 1.9	11.1 ± 2.8	11.5 ± 2.7	8.26 - 23.7	2.51 - 20.2	0.534	0.177
Ash (% DW)	3.80 ± 0.29	3.80 ± 0.29	3.79 ± 0.42	3.07 - 5.16	2.66 - 4.99	0.983	0.919
Carbohydrates (% DW)	52.3 ± 2.8	53.3 ± 2.6	53.4 ± 2.6	46.3 - 61.0	46.5 - 63.0	0.104	0.084
Crude Fat (% DW)	20.5 ± 2.8	20.4 ± 2.5	20.5 ± 2.2	13.7 - 24.4	13.6 - 25.6	0.730	0.939
Crude Protein (% DW)	23.3 ± 2.6	22.5 ± 2.4	22.3 ± 2.6	15.8 - 28.7	14.9 - 28.8	0.029	0.008
Acid Detergent Fiber (% DW)	43.0 ± 5.6	42.9 ± 4.2	43.3 ± 3.7	36.6 - 54.1	33.1 - 54.2	0.944	0.721
Neutral Detergent Fiber (% DW)	45.6 ± 3.2	46.7 ± 3.1	46.9 ± 3.0	41.2 - 58.6	39.2 - 57.7	0.094	0.048
Total Dietary Fiber (% DW)	43.7 ± 8.8	47.4 ± 5.3	47.6 ± 6.6	25.2 - 77.0	30.3 - 64.9	0.067	0.058

[a] Composition samples were derived from eight field trials conducted in the USA in 2014 and 2015.

[b] 99% Tolerance Interval: Range of reference lines based on tolerance intervals specified to contain 99% of the population with 95% confidence.

[c] t-Test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 24. Comparison of Amino Acids in Fuzzy Seed of GHB811 Cotton with its Non-GM Counterpart[a] (% DW)

Parameter	Non-GM Counterpart (Entry A) Mean ± SD	GHB811 Not Treated (Entry J) Mean ± SD	GHB811 Treated (Entry K) Mean ± SD	Non-GM Reference Varieties Range (Entries B-G,N) (Min-Max)	Tolerance Interval Non-GM Reference Varieties (Entries B-G,N) [b] (Lower-Upper)	Comparison t-test A vs J [c] p-value	Comparison t-test A vs K [c] p-value
Alanine	0.910 ± 0.088	0.904 ± 0.090	0.928 ± 0.088	0.634 - 1.07	0.607 - 1.14	0.729	0.265
Arginine	2.43 ± 0.39	2.40 ± 0.35	2.44 ± 0.35	1.45 - 3.18	1.33 - 3.29	0.597	0.969
Aspartic Acid	2.10 ± 0.30	2.12 ± 0.29	2.12 ± 0.31	1.32 - 2.58	1.22 - 2.78	0.684	0.650
Cystine	0.439 ± 0.077	0.421 ± 0.082	0.408 ± 0.067	0.271 - 0.574	0.224 - 0.607	0.257	0.047
Glutamic Acid	4.53 ± 0.63	4.52 ± 0.58	4.55 ± 0.64	2.97 - 5.70	2.73 - 6.06	0.918	0.886
Glycine	1.013 ± 0.103	0.988 ± 0.100	1.009 ± 0.106	0.667 - 1.26	0.650 - 1.27	0.256	0.864
Histidine	0.680 ± 0.083	0.666 ± 0.072	0.676 ± 0.074	0.455 - 0.969	0.417 - 0.887	0.376	0.824
Isoleucine	0.743 ± 0.082	0.737 ± 0.076	0.744 ± 0.075	0.518 - 0.866	0.490 - 0.935	0.663	0.951
Leucine	1.38 ± 0.15	1.37 ± 0.14	1.39 ± 0.13	0.942 - 1.64	0.910 - 1.74	0.616	0.700
Lysine	1.05 ± 0.10	1.05 ± 0.12	1.07 ± 0.12	0.746 - 1.30	0.693 - 1.35	0.949	0.461
Methionine	0.319 ± 0.047	0.308 ± 0.055	0.297 ± 0.047	0.188 - 0.384	0.191 - 0.412	0.303	0.047
Phenylalanine	1.29 ± 0.17	1.26 ± 0.16	1.29 ± 0.15	0.801 - 1.57	0.754 - 1.68	0.450	0.953
Proline	0.877 ± 0.101	0.866 ± 0.093	0.884 ± 0.089	0.595 - 1.08	0.569 - 1.12	0.519	0.665
Serine	1.045 ± 0.112	1.025 ± 0.109	1.035 ± 0.140	0.707 - 1.66	0.626 - 1.39	0.465	0.723
Threonine	0.777 ± 0.071	0.769 ± 0.067	0.776 ± 0.074	0.556 - 0.894	0.537 - 0.954	0.558	0.928
Tryptophan	0.241 ± 0.034	0.249 ± 0.036	0.246 ± 0.034	0.177 - 0.293	0.156 - 0.310	0.258	0.460
Tyrosine	0.499 ± 0.069	0.473 ± 0.083	0.487 ± 0.068	0.176 - 0.608	0.250 - 0.688	0.100	0.443
Valine	1.036 ± 0.121	1.029 ± 0.115	1.041 ± 0.111	0.720 - 1.24	0.671 - 1.32	0.741	0.808

[a] Composition samples were derived from eight field trials conducted in the USA in 2014 and 2015.

[b] 99% Tolerance Interval: Range of reference lines based on tolerance intervals specified to contain 99% of the population with 95% confidence.

[c] t-Test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 25. Comparison of Fatty Acids in Fuzzy Seed of GHB811 Cotton with its Non-GM Counterpart[a] (% Total Fatty Acids)

Parameter	Non-GM Counterpart (Entry A) Mean ± SD	GHB811 Not Treated (Entry J) Mean ± SD	GHB811 Treated (Entry K) Mean ± SD	Non-GM Reference Varieties Range (Entries B-G,N) (Min-Max)	Tolerance Interval Non-GM Reference Varieties (Entries B-G,N) [b] (Lower-Upper)	Comparison t-test A vs J [c] p-value	Comparison t-test A vs K [c] p-value
C12:0 Lauric Acid	<LOQ	<LOQ	<LOQ	<LOQ - 0.0377	NA	NA	NA
C14:0 Myristic Acid	0.651 ± 0.150	0.665 ± 0.145	0.665 ± 0.138	0.470 - 1.17	0.217 - 1.11	0.310	0.291
C16:0 Palmitic Acid	23.2 ± 1.8	23.2 ± 1.8	23.4 ± 1.6	19.7 - 27.6	17.7 - 27.7	0.837	0.342
C16:1 Palmitoleic Acid	0.486 ± 0.049	0.524 ± 0.052	0.525 ± 0.043	0.392 - 0.641	0.325 - 0.618	<.001	<.001
C17:0 Heptadecanoic Acid	0.0820 ± 0.0096	0.0827 ± 0.0116	0.0808 ± 0.0094	0.0653 - 0.1180	0.0517 - 0.1206	0.752	0.569
C17:1 Heptadecenoic Acid	0.0436 ± 0.0121	0.0446 ± 0.0116	0.0458 ± 0.0095	0.0175 - 0.0821	0.0050 - 0.0937	0.682	0.376
C18:0 Stearic Acid	2.37 ± 0.16	2.28 ± 0.17	2.26 ± 0.15	2.00 - 2.97	1.75 - 3.07	<.001	<.001
C18:1 Oleic Acid	14.2 ± 1.0	14.2 ± 1.0	14.2 ± 1.0	13.4 - 20.8	9.2 - 21.3	0.912	0.956
C18:2 Linoleic Acid	57.3 ± 2.9	57.5 ± 3.0	57.3 ± 2.8	46.3 - 60.7	45.2 - 67.8	0.682	0.877
C18:3 Linolenic Acid	0.229 ± 0.121	0.228 ± 0.113	0.210 ± 0.109	0.0301 - 0.547	0 - 0.562	0.984	0.481
C18:4 Stearidonic Acid	<LOQ - 0.0653	<LOQ - 0.0562	<LOQ - 0.0509	<LOQ - 0.0832	NA	NA	NA
C20:0 Arachidic Acid	0.254 ± 0.032	0.247 ± 0.035	0.244 ± 0.033	0.196 - 0.376	0.117 - 0.388	0.112	0.036
C20:1 Eicosenoic Acid	0.0666 ± 0.0154	0.0709 ± 0.0166	0.0663 ± 0.0136	0.0396 - 0.0980	0.0289 - 0.1129	0.268	0.931
C20:2 Eicosadienoic Acid	<LOQ	<LOQ - 0.0382	<LOQ - 0.0451	<LOQ	NA	NA	NA
C20:5 Eicosapentaenoic Acid	<LOQ	<LOQ	<LOQ	<LOQ - 0.0427	NA	NA	NA
C22:0 Behenic Acid	0.148 ± 0.021	0.149 ± 0.026	0.148 ± 0.023	0.110 - 0.212	0.074 - 0.224	0.829	0.903
C22:5 N3 Docosapentaenoic Acid	<LOQ - 0.0767	<LOQ - 0.128	<LOQ - 0.0712	<LOQ - 0.104	NA	NA	NA
C22:5 N6 Docosapentaenoic Acid	<LOQ - 0.22	<LOQ - 0.259	<LOQ - 0.185	<LOQ - 0.277	NA	NA	NA
C24:0 Lignoceric Acid	0.117 ± 0.054	0.099 ± 0.059	0.109 ± 0.071	0.0295 - 0.285	0 - 0.285	0.238	0.603

[a] Composition samples were derived from eight field trials conducted in the USA in 2014 and 2015.

[b] 99% Tolerance Interval: Range of reference lines based on tolerance intervals specified to contain 99% of the population with 95% confidence.

[c] t-Test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

NA=Not Applicable because more than 1/3 of the values are <LOQ. Minimum and maximum are reported instead of mean and standard deviation.

Table 26. Comparison of Minerals and Alpha Tocopherol in Fuzzy Seed of GHB811 Cotton with its Non-GM Counterpart[a] (mg/kg DW)

Parameter	Non-GM Counterpart (Entry A) Mean ± SD	GHB811 Not Treated (Entry J) Mean ± SD	GHB811 Treated (Entry K) Mean ± SD	Non-GM Reference Varieties Range (Entries B-G,N) (Min-Max)	Tolerance Interval Non-GM Reference Varieties (Entries B-G,N) [b] (Lower-Upper)	Comparison t-test A vs J [c] p-value	Comparison t-test A vs K [c] p-value
Calcium	983 ± 288	1022 ± 296	1039 ± 270	702 - 1960	456 - 1930	0.435	0.263
Copper	6.19 ± 2.35	6.21 ± 2.55	6.15 ± 2.27	1.99 - 11.5	0 - 13.3	0.937	0.835
Iron	37.2 ± 6.7	36.1 ± 6.1	37.1 ± 6.8	22.1 - 60.1	13.2 - 60.0	0.415	0.937
Magnesium	4013 ± 489	4002 ± 486	3917 ± 355	2848 - 5328	2380 - 5438	0.896	0.226
Manganese	12.9 ± 2.5	12.5 ± 2.0	12.4 ± 2.2	9.94 - 25.2	5.35 - 23.4	0.111	0.077
Phosphorus	6613 ± 1298	6749 ± 960	6449 ± 1277	3888 - 9333	3137 - 10251	0.466	0.381
Potassium	12287 ± 1158	12508 ± 1385	12193 ± 943	10263 - 15727	9269 - 15027	0.342	0.684
Sodium	30.3 ± 15.5	37.3 ± 29.8	32.5 ± 24.9	10.0 - 148	0 - 74.9	0.116	0.626
Zinc	33.7 ± 5.4	33.7 ± 5.8	33.2 ± 5.8	21.5 - 49.6	17.2 - 50.3	0.957	0.553
Alpha Tocopherol	121.6 ± 19.1	109.8 ± 18.1	109.1 ± 15.0	49.8 - 141	37.0 - 157	0.002	0.001

[a] Composition samples were derived from eight field trials conducted in the USA in 2014 and 2015.

[b] 99% Tolerance Interval: Range of reference lines based on tolerance intervals specified to contain 99% of the population with 95% confidence.

[c] t-Test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

Table 27. Comparison of Anti-nutrients in Fuzzy Seed of GHB811 Cotton with its Non-GM Counterpart[a]

Parameter	Non-GM Counterpart (Entry A) Mean ± SD	GHB811 Not Treated (Entry J) Mean ± SD	GHB811 Treated (Entry K) Mean ± SD	Non-GM Reference Varieties Range (Entries B-G,N) (Min-Max)	Tolerance Interval Non-GM Reference Varieties (Entries B-G,N) [b] (Lower-Upper)	Comparison t-test A vs J [c] p-value	Comparison t-test A vs K [c] p-value
Free Gossypol (% DW)	0.591 ± 0.145	0.534 ± 0.122	0.526 ± 0.118	0.273 - 0.941	0.106 - 1.15	0.007	0.002
Total Gossypol (% DW)	0.804 ± 0.167	0.700 ± 0.160	0.735 ± 0.145	0.346 - 1.34	0.153 - 1.53	0.002	0.033
Cyclopropenoic Fatty Acids (% Total Fatty Acids)							
Dihydrosterculic Acid	0.198 ± 0.052	0.188 ± 0.062	0.169 ± 0.048	0.126 - 0.407	0.0972 - 0.438	0.405	0.021
Malvalic Acid	0.380 ± 0.176	0.384 ± 0.147	0.355 ± 0.172	0.014 - 0.983	0 - 1.062	0.928	0.533
Sterculic Acid	0.146 ± 0.066	0.136 ± 0.056	0.127 ± 0.062	0.014 - 0.366	0 - 0.423	0.527	0.236

[a] Composition samples were derived from eight field trials conducted in the USA in 2014 and 2015.

[b] 99% Tolerance Interval: Range of reference lines based on tolerance intervals specified to contain 99% of the population with 95% confidence.

[c] t-Test p-value: Pairwise comparison to the non-GM counterpart (Entry A).

NA=Not Applicable because more than 1/3 of the values are <LOQ. Minimum and maximum are reported instead of mean and standard deviation.

By-site Analysis
For the six analytes (crude protein, palmitoleic acid, stearic acid, alpha tocopherol, free gossypol and total gossypol) for which there was a statistically significant combined-site difference between the non-GM counterpart (Entry A) and GHB811 not treated (Entry J) and the 11 analytes (same six analytes as Entry A vs Entry J, with the addition of neutral detergent fiber, cystine, methionine, arachidic acid, and dihydrosterculic acid) for which there was a significant combined-site difference between the non-GM counterpart (Entry A) and GHB811 treated (Entry K) by-site results were examined.

There were significant differences between the non-GM counterpart (Entry A) and GHB811 not treated or treated with trait-specific herbicides (Entry J and Entry K) at over half the sites for palmitoleic acid.

For the remaining analytes (crude protein, neutral detergent fiber, cystine, methionine, arachidic acid, stearic acid, alpha tocopherol, free gossypol, total gossypol and dihydrosterculic acid) where there were significant differences between the non-GM counterpart (Entry A) and GHB811 cotton not treated or treated with trait-specific herbicides (Entry J and Entry K), less than half of the sites showed significant differences.

Conclusions
Comparison of key nutrient and anti-nutrient levels of GHB811 cotton fuzzy seed (not treated and treated with trait-specific herbicides) to the non-GM counterpart revealed statistically significant differences for 11 of the 54 analytes examined. However, the means of all 11 analytes were within the range of the reference varieties and the tolerance intervals. Therefore, the statistically significant differences are not considered biologically relevant. Based on the comparative assessment, nutrient and anti-nutrient levels in GHB811 cotton fuzzy seed are comparable to that of the non-GM counterpart and reference varieties.

(b) Information on the range of natural variation for each constituent measured to allow for assessment of biological significance should any statistically significant differences be identified

The OECD "Consensus document on compositional considerations for new varieties of cotton (*Gossypium hirsutum* and *Gossypium barbadense*): Key food and feed nutrients and anti-nutrients" (2009) provides the required information on natural variation for each constituent measured within the compositional studies to allow assessment of biological significance should any statistically significant differences be identified in the above studies by FSANZ.

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

Other than the intended presence of the 2mEPSPS and HPPD W336 proteins in GHB811 cotton, food products derived from GHB811 cotton have been shown to be compositionally and nutritionally similar to products derived from commercial varieties of non-transgenic cotton (see Section B.5(a) directly above).

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

Cotton is not considered to produce allergenic proteins.

Part C Information Related to the Nutritional Impact of the Food Produced Using Gene Technology

The application must contain the following information if the compositional analysis indicates biologically significant changes to the levels of certain nutrients in the food produced using gene technology compared to the non-GM counterpart food:

(a) Data are required on the anticipated dietary intake of the GM food in relation to the overall diet, together with any information which may indicate a change to the bioavailability of the nutrients from the GM food

Based on the composition analysis, where nutrient and anti-nutrient levels were found to be similar between the non-GM conventional counterpart and the GHB811 cotton seed and processed fraction samples, no analysis of dietary intake in relation to the overall diet is required as bioavailability of the nutrients from GHB811 cotton derived foods is expected to be similar to bioavailability of nutrients from non-GM cotton derived foods.

(b) Where the GM food contains an intended nutritional change, information, such as clinical trial data, must be provided to determine the nutritional impact of the GM food.

Not applicable.

Part D Other Information

There is no requirement to conduct animal feeding or whole food toxicity studies on the food produced using gene technology. However, if a 90-day (or longer) whole food toxicity study in rodents has been provided to satisfy the data and information requirements of another jurisdiction, this should also be provided to FSANZ as additional supporting information.

A 90-day (or longer) whole food toxicity study in rodents is not available. Should this situation change the study results will be provided to FSANZ.

LITERATURE CITED

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

1. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25, 3389-3402.
2. Astwood, J.D. (1996) Stability of food allergens to digestion *in vitro*. *Nature Biotechnology* 14, 1269-1273.
3. Australia New Zealand Food Authority (ANZFA) (2002) Final assessment report (Full assessment – S.15) Application A436: Oil and linters derived from insect-protected cotton containing event 15985. Australia New Zealand Food Authority, Canberra, Australia.
4. Bailey, A. V., Harris, J. A., Skau, E.L. (1966) Cyclopropenoid fatty acid content and fatty acid composition of crude oils from twenty-five varieties of cottonseed. *Journal of the American Oil Chemists Society* 43, 107- 110.
5. Bernton, H. S., Coulson, E. J., Stevens, H. (1949) On allergy to cottonseed oil. *Journal of the American Medical Association* 140, 869-871.

6. Bernton, H. S., Spies, J. R., Stevens, H. (1940) Significance of cottonseed sensitiveness. *Journal of Allergy* 11, 138-146.
7. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W. (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene 2, 95-113.
8. Bossis, E., Lemanceau, P., Latour, X. and Gardan, L. (2000) The taxonomy of *Pseudomonas fluorescens* and *Pseudomonas putida*: current status and need for revision. *Agronomie* 20:51-63.
9. Boudec P., Rodgers M., Dumas F., Sailland A., Bourdon H. (2001) Mutated hydroxyphenypyruvate dioxygenase, DNA sequence and isolation of plants which contain such a gene and which are tolerant to herbicides. US Patent US6245968B1 (12-Jun-2001).
10. Brownlee, J. M., Johnson-Winters, K., Harrison, D. H. T., Moran G. R. (2004) Structure of the Ferrous Form of (4-Hydroxyphenyl)pyruvate Dioxygenase from Streptomyces avermitilis in Complex with the Therapeutic Herbicide. Biochemistry 43: 6370 – 6377.
11. Brubaker, C.L., Bourland, F.M., Wendel, J.E. (1999) Chapter 1.1: The origin and domestication of cotton. In: Smith, C.W., Cothren J.T., eds. *Cotton: Origin, History, Technology, and Production*. John Wiley and Sons, Inc. New York. 3-31.
12. Bundock, P., Hooykaas, P. (1998). Interactions between *Agrobacterium tumefaciens* and Plant Cells. Phytochemical Signals and Plant-Microbe Interaction. Phytochemical Signals and Plant-Microbe Interaction/Recent Advances in Phytochemistry. Chapter 11: 207-229. Document no. M-550150-01-1.
13. Chabouté M., Chaubet N., Philipps G., Ehling M. and Gigot C. (1987) Genomic organization and nucleotide sequences of two histone H3 and two histone H4 genes of *Arabidopsis thaliana*. *Plant Molecular Biology*, 8:179-191.
14. Chaubet, N., Clement, B. and Gigot C. (1992) Genes encoding a histone H3.3-like variant in *Arabidopsis* contain intervening sequences. *Journal of Molecular Biology* 225:569-574.
15. Cherry, J.P., Leffler, H.R. (1984) Seed. Chapter 13. In: R.J. Kohel, C.F. Lewis, eds. Cotton, Agronomy Monograph No. 24, Edition 24. ASA-CSSA-SSSA, Madison, WI. pp. 511-558.
16. *Codex Alimentarius* Commission (CAC) (2009) *Codex Alimentarius* Foods derived from modern biotechnology. *Codex Alimentarius* Commission, Rome. 85 pages. ftp://ftp.fao.org/codex/Publications/Booklets/Biotech/Biotech_2009e.pdf
17. Compant, S., Duffy, B., Nowak, J., Clement, C. and Barka, E.A. (2005) Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Applied Environmental Microbiology* 71:4951-4959.
18. Cornelissen, M.; Vandewiele, M. (1989) Nuclear transcriptional activity of the tobacco plasmid psbA promoter. Nucleic Acids Research, 17:1, pp. 19-29. Document no. M-147994-01-2.
19. Deblaere, R.; Reynaerts, A.; Hoefte, H.; Hernalsteens, J. P.; Leemans, J.; van Montagu, M. (1987) Vectors for cloning in plant cells. Document no. M-147996-01-1.
20. De Groot, M.J.A., Bundock, P., Hooykaas, P.J.J. and Beijersbergen, A.G.M. (1998) *Agrobacterium*-mediated transformation of filamentous fungi. *Nature Biotechology* 16:839–842.
21. De Kochko, A., Verdaguer, B., Taylor, N., Carcamo, R., Beachy, R.N., Fauquet, C. (1998) Cassava vein mosaic virus (CsVMV), type species for a new genus of plant double stranded DNA viruses? Arch Virol.143(5):945-62.
22. Depicker, A., Stachel, S., Dhaese, P., Zambryski, P., Goodman, H.M. (1982) Nopaline synthase: Transcript mapping and DNA sequence. *Journal of Molecular and Applied Genetics* 1, 561-573. Document no. M-131630-01-2.
23. Dhaese, P.; De Greve, H.; Gielen, J.; Seurinck, J.; van Montagu, M.; Schell, J. (1983) Identification of sequences involved in the polyadenylation of higher plant nuclear

transcripts using Agrobacterium T-DNA genes as models The EMBO Journal, 2:3, pp. 419-426. Document no. M-180190-01-1.
24. Downing, K.J, Leslie, G. and Thomson, J.A. (2000) Biocontrol of the sugarcane borer *Eldana saccharina* by expression of the *Bacillus thuringiensis cry1Ac7* and *Serratia marcescens chiA* genes in sugarcane-associated bacteria. *Applied Environmental Microbiology* 66:2804–2810.
25. Ensminger, A.H., Ensminger, M.E., Konlande, J.E., Robson, J.R.K. (1994) *Foods and Nutrition Encyclopaedia*, 2nd Ed. Ann Harbor, MI. p 497-507.
26. European Food Safety Agency (EFSA) (2011) Guidance on the risk assessment of genetically modified microorganisms and their products intended for food and feed use. EFSA Journal 2011;9(6):2193. 54 pages. http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/2193.pdf.
27. Fliessbach, A., Winkler, M., Lutz, M.P., Oberholzer, H.R. and Mader, P. (2009) Soil amendment with *Pseudomonas fluorescens* CHA0: lasting effects on soil biological properties in soils low in microbial biomass and activity. *Microbial Ecology* 1-13.
28. Fling, M.E., Kopf, J., Richards, C. (1985) Nucleotide sequence of the transposon Tn7 gene encoding an aminoglycoside-modifying enzyme, 3"(9)-*O*-nucleotidyltransferase. *Nucleic Acids Research* 13, 7095-7106.
29. Food and Agriculture Organization/World Health Organization (FAO/WHO) (2001) Evaluation of allergenicity of genetically modified foods. Report of a joint Food and Agriculture Organisation of the United Nations (FAO)/World Health Organisation (WHO) expert consultation on foods derived from biotechnology. Rome, Italy. 35 pages. ftp://ftp.fao.org/es/esn/food/allergygm.pdf
30. Food Standards Australia New Zealand (FSANZ) (2014) Safety assessment – Application A1094 (at approval). Food derived from herbicide-tolerant Cotton Line DAS – 81910-7.
31. Food Standards Australia New Zealand (FSANZ) (2011) Application A1051 – Food derived from Herbicide-tolerant Soybean line FG72. Safety Assessment Report (Approval). http://www.foodstandards.gov.au/code/applications/pages/applicationa1051food4902.aspx
32. Franck, A.W. (1987) Food uses of cottonseed protein. Developments in Food proteins - 5. New York. p 31-80.
33. Fritze, I. M., Linden, L., Freigang, J., Auerbach, G., Huber, R., Steinbach, S. (2004) The crystal structures of *Zea mays* and *Arabidopsis* 4-hydroxyphenylpyruvate dioxygenase. *Plant Physiology* 134: 1388 – 1400.
34. Fuchs, R.L., Astwood, J.D. (1996) Allergenicity assessment of foods derived from genetically modified plants. *Food Technology* 9, 83-88.
35. Gershman, M.D., Kennedy, D.J., Noble-Wang, J., Kim, C., Gullion, J., Kacica, M., Jensen, B., Pascoe, N., Saiman, L., McHale, J., Wilkins, M., Schoonmaker-Bopp, D., Clayton, J., Arduino, M. and Srinivasan, A. (2008) Multistate outbreak of *Pseudomonas fluorescens* bloodstream infection after exposure to contaminated heparinized saline flush prepared by a compounding pharmacy. *Clinical Infectious Diseases* 47:1372-1379.
36. Gibb, A.P., Martin, K.M., Davidson, G.A., Walker, B. and Murphy, W.G. (1995) Rate of growth of *Pseudomonas fluorescens* in donated blood. *Journal of Clinical Pathology* 48:717-718.
37. Gregory, S.R., Hernandez, E., Savoy, B.R. (1999) Cottonseed processing. Chapter 4.5. In: C.W. Smith, J.T. Cothren, eds. Cotton: Origin, History, Technology and Production. John Wiley & Sons, New York. Pp. 793-819.
38. Haas, D. and Defago, G. (2005) Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nature Review Microbiology* 3:307-319.

39. Hajdukiewicz, P., Svab, Z., Maliga, P. (1994) The small, versatile *pPZP* family of *Agrobacterium* binary vectors for plant transformation. *Plant Molecular Biology* 25, 989-994.
40. Hallauer, A.R., Russell, W.A. and Lamkey, K.R. (1988) Corn breeding. In: Sprague, G.F. and Dudley, J.W. (Eds). *Corn and corn improvement*, Third Edition. No 18 in the series Agronomy. American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc., Madison, Wisconsin, USA. pp: 463-564.
41. Hamid, M., Siddiqui, I.A. and Shahid Shaukat, S. (2003) Improvement of *Pseudomonas fluorescens* CHA0 biocontrol activity against root-knot nematode by the addition of ammonium molybdate. *Letters in Applied Microbiology* 36:239-244.
42. Harris, J.A., Magne, F.C., Skau, E.L. (1964) Methods for the determination of cyclopropenoid fatty acids. IV. Application of the step-wise HBr titration method to the analysis of refined and crude cottonseed oils. *Journal of the American Oil Chemists Society* 41, 309-311.
43. Hefle, S. L., Taylor, S. L. (1999) Allergenicity of edible oils. *Foodtechnology* 53, 62- 70.
44. Herrmann, K. M. (1995) The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. The Plant Cell 7: 907-919.
45. Herouet-Guicheney, C., Rouquie, D., Freyssinet, M., Currier, T., Martine, A., Zhou, J., Bates, E.E.M., Ferullo, J-M., Hendrickx, K. and Rouan, D. (2009) Safety evaluation of the double mutant 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) from maize that confers tolerance to glyphosate herbicide in transgenic plants. *Regulatory Toxicology and Pharmacology* 54:143-153.
46. Hoess R. H., Abremski K. (1985) Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. Journal of Molecular Biology, 181, 351-362.
47. Hothersall, J., Wu, J., Rahman, A.S., Shields, J.A., Haddock, J., Johnson, N., Cooper, S.M., Stephens, E.R., Cox, R.J., Crosby, J., Willis, C.L., Simpson, T.J. and Thomas, C.M. (2007) Mutational analysis reveals that all tailoring region genes are required for production of polyketide antibiotic mupirocin by *Pseudomonas fluorescens*: pseudomonic acid B biosynthesis precedes pseudomonic acid A. *Journal of Biological Chemistry* 282:15451-15461.
48. Huby, R.D., Dearman, R.J., Kimber, I. (1995) Why are some proteins allergens? *Toxicological Sciences* 55, 235-246.
49. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) (1984) Nomenclature and symbolism for amino acids and peptides, Recommendations 1983. *European Journal of Biochemistry* 138, 9-37.
50. Jenkins, J.N. (2003) Cotton. In: *Traditional crop breeding practices: an historical review to serve as a baseline for assessing the role of modern biotechnology*. OECD. 61-70.
51. Jenkins, N., Parekh, R.B., James, D.C. (1996) Getting the glycosylation right: implications for the biotechnology industry. *Nature Biotechnology* 14, 975-981.
52. Jones, L. A., King C. C. (1993) Cottonseed oil. National Cottonseed Products Association, Inc. and the Cotton Foundation, Memphis, TN.
53. Joshi C.P., Zhou, H., Huang X., Chiang V.L. (1997) Context sequences of translation initiation codon in plants. *Plant Molecular Biology* 35, 993-1001.
54. Keeler, K.H., Turner, C.E., Bolick, M.R. (1996) Chapter 20: Movement of crop transgenes into wild plants. In: Duke, S.O., ed. *Herbicide resistant plants*. CRC Press Inc. 303-330.
55. Konig, A., Cockburn, A., Crevel, R.W., Debruyne, E., Grafstroem, R., Hammerling, U., Kimber, I., Knudsen, I., Kuiper, H.A., Peijnenburg, A.A., Penninks, A.H., Poulsen, M., Schauzu, M., Wal, J.M. (2004) Assessment of the safety of foods derived from genetically modified (GM) crops. *Food Chemistry and Toxicology* 42, 1047-1088.
56. Lazarides, M., Cowley, K., Hohnen, P. (1997) *CSIRO Handbook of Australian Weeds.* CSIRO, Canberra, ACT.

57. Lebrun, M.; Leroux, B.; Sailland, A. (1996) Chimeric gene for the transformation of plants. US Patent US5510471.
58. Lebrun, M., Sailland, A. and Freyssinet, G. (1997) Mutated 5-enol-pyruvylshikimate-3-phosphate synthase, gene coding for said protein and transformed plants containing said gene. Patent Application: WO9704103-A 1. Rhone Poulenc Agrochimie, France.
59. Levi, R. S., Reilich, H. G., O'Neill, H. J., Cucullu, A. F., Skau, E. L. (1967) Quantitative determination of cyclopropenoid fatty acids in cottonseed meal. *Journal of the American Oil Chemists Society* 44, 249-252.
60. Liu, P.V. (1964) Pathogenicity of *Pseudomonas fluorescens* and related Pseudomonads to warm-blooded animals. *American Journal of Clinical Pathology* 41:150-153.
61. Lusas, E.W., Jividen, G.M. (1987) Glandless cottonseed: a review of the first 25 years of processing and utilization research. *Journal of American Oil Chemical Society* 64:839-854.
62. Matringe M., Sailland A., Pelissier B., Rolland A. and Zink O. (2005) p-hydroxyphenylpyruvate dioxygenase inhibitor-resistant plants. *Pest Management Science*, **61**, 269-273.
63. McKellar, R.C. (1982) Factors influencing the production of extracellular proteinase by *Pseudomonas fluorescens*. *Journal of Applied Bacteriology* 53:305-316.
64. Metcalfe, D.D., Astwood, J.D., Townsend, R., Sampson, H.A., Taylor, S.L., Fuchs, R.L. (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Critical Reviews in Food Science and Nutrition* 36 Suppl:S165-186.
65. Mühldorfer, I., Hacker, J. (1994) Genetic aspects of *Escherichia coli* virulence. *Microbial Pathogenesis* 16, 171-181.
66. National Cotton Council of America (NCCA) (1999) Cotton: from field to fabric. Educational Materials.
67. National Cottonseed Products Association (NCPA) (2000) Cottonseed Oil. CSO Bulletin. National Cottonseed Products Association, Inc., Memphis, TN.
68. National Cottonseed Products Association (NCPA) (1999) *Cottonseed and its Products*. CSIP 10th Edition, 1999. National Cottonseed Products Association, Inc, Memphis, TN.
69. Office of the Gene Technology Regulator (OGTR) (2016) The biology of *Gossypium hirsutum* L. and *Gossypium barbadense* L. (cotton). Australian Government Department of Health.
70. Office of the Gene Technology Regulator (OGTR) (2008) The biology of *Gossypium hirsutum* L. and *Gossypium barbadense* L. (cotton). Australian Government Department of Health.
71. Office of the Gene Technology Regulator (OGTR) (2002) The biology and ecology of cotton (*Gossypium hirsutum*) in Australia. Australian Government Department of Health.
72. Organisation for Economic Co-operation and Development (OECD) (1997) Series on harmonization of regulatory oversight Nº6. Consensus document on information used on the assessment of environmental applications involving Pseudomonas. March 4, 1997.
73. Organisation for Economic Co-operation and Development (OECD) (2008) Environment Directorate. Consensus document on the biology of cotton (*Gossypium* spp.). OECD Environment, Health and Safety Publications, Series on Harmonisation of Regulatory Oversight in Biotechnology, No. 45.
74. Organisation for Economic Co-operation and Development (OECD) (2009) 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. 11, OECD Environment, Health and Safety Publications, Paris.
75. Palleroni, N.J. (1981) Introduction to the family *Pseudomonadaceae*. In: Starr, M.P., Stolp, H., Truper, H.G., Balows, A. and Schegel, H.G. (Eds) *The Prokaryotes - A*

Handbook on Habitats, Isolation and Identification of Bacteria Chapter 59. Springer-Verlag, Berlin. pp. 655-665.

76. Peng, R., Xiong, A., Li, X., Fuan, H. and Yao, Q. (2003) A delta-endotoxin encoded in *Pseudomonas fluorescens* displays a high degree of insecticidal activity. *Applied Microbiology and Biotechnology* 63:300-306.
77. Phelps, R.A., Shenstone, F.S., Kemmerer, R. J., Evans, R.J. (1965) A review of cyclopropenoid compounds: biological effects of some derivatives. *Poultry Science* 44, 358-394.
78. Polovic, N., Blanusa, M., Gavrovic-Jankulovic, M., Atanaskovic-Markovic, M., Burazer, L., Jankov, R., Velickovic, T.C. (2007) A matrix effect in pectin-rich fruits hampers digestion of allergen by pepsin in vivo and in vitro. Clinical and Experimental Allergy 37, 764-771.
79. Pringle, C.R. Arch. Virol. (1999) 144: 421. doi:10.1007/s007050050515 Virus Taxonomy – 1999 The Universal System of Virus Taxonomy, updated to include the new proposals ratified by the International Committee on Taxonomy of Viruses during 1998.
80. Puckett, A., Davison, G., Entwistle, C.C. and Barbara, J.A. (1992) Post transfusion septicaemia 1980-1989: importance of donor arm cleansing. *Journal of Clinical Pathology* 45:155-157.
81. Raaijmakers, J.M., de Bruijn, I. and de Kock, M.J. (2006) Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: diversity, activity, biosynthesis, and regulation. *Molecular Plant-Microbe Interactactions* 19:699-710.
82. Raybould, A.; Kilby, P.; Graser, G. (2013) Characterising microbial protein test substances and establishing their equivalence with plant-produced proteins for use in risk assessments of transgenic crops. Transgenic Res. 22: 445-460.
83. Skerman, V.B.D., McGowan, V. And Sneath, P.H.A. (1980) Approved lists of bacterial names. *International Journal of Systematic Bacteriology* 30:225-420. Available at: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=bacname
84. Smith, C.W. (1995) Cotton (*Gossypium hirsutum* L.). Chapter 6. In: Crop Production: Evolution, History, and Technology. John Wiley and Sons, Inc., New York. Pp. 287-349.
85. Stachel S.E., Zambryski P.C. (1986) virA and virG control the plant-induced activation of the T-DNA transfer process of *A. tumefaciens*. *Cell* 46, 325-333.
86. Tacconelli, E., Carmeli, Y., Aizer, A., Ferreira, G., Foreman, M.G. and D'Agata, E.M. (2003) Mupirocin prophylaxis to prevent *Staphylococcus aureus* infection in patients undergoing dialysis: a meta-analysis. *Clinical Infectious Disease* 37:1629-1638.
87. Thomas, K., Bannon, G., Hefle, S., Herouet, C., Holsapple, M., Ladics, G., MacIntosh, S., Privalle, L. (2005) *In silico* methods for evaluating human allergenicity to novel proteins: International bioinformatics workshop meeting report, 23-24 February 2005. *Toxicological Sciences* 88, 307-310.
88. 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., 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.
89. Thrane, C., Nielsen, M.N., Sorensen, J. and Olsson, S. (2001) *Pseudomonas fluorescens* DR54 reduces sclerotia formation, biomass development, and disease incidence of *Rhizoctonia solani* causing damping-off in sugar beet. *Microbial Ecology* 42:438-445.
90. Tothill, J.C., Mott, J.J., Gillard, P. (1982) Chapter 35: Pasture weeds of the tropics and subtropics with special reference to Australia. In: Holzner, W., Numata M., eds. *Biology and Ecology of Weeds*. Dr W. Junk Publishers The Hague. 403-427.

91. United States Environmental Protection Agency (US-EPA) (1994) *Pseudomonas fluorescens* - Tolerance Exemption 8/94. Available at: http://pmep.cce.cornell.edu/profiles/biopestbiocont/biofung/pseudomo_fluor/pseuflu_exm_0894.html
92. United States Environmental Protection Agency (US-EPA) (2008a) Frost-preventing bacteria: *Pseudomonas fluorescens* A506 (006438); *Pseudomonas fluorescens* 1629RS (006439); *Pseudomonas syringae* 742RS (006411) Fact Sheet.
93. United States Environmental Protection Agency (US-EPA) (2008b) Frost-Preventing Bacteria: *Pseudomonas fluorescens* A506 (006438); *Pseudomonas fluorescens* 1629RS (006439); *Pseudomonas syringae* 742RS (006411) Product pages.
94. United States Food and Drug Administration (US-FDA) (2012) Guidance for industry. Recommendations for the early food safety evaluations of new non-pesticidal proteins produced by new plant varieties intended for food use. http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/Biotechnology/ucm096156.htm#paper
95. van Larebeke, N.; Engler, G.; Holsters, M.; van den Elsacker, S.; Zaenen, J.; Schilperoort, R. A.; Schell, J. (1974) Large plasmid in Agrobacterium tumefaciens essential for crown gall-inducing ability. Nature, 252: 169-170. Document no. M-147632-01-1.
96. Verdaguer, B.; de Kochko, A.; Beachy, R. N.; Fauquet, C. (1996) Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. Plant Molecular Biology 31: 1129-1139.
97. Voisard, C., Keel, C., Haas, D. and Defago, G. (1989) Cyanide production by *Pseudomonas fluorescens* helps suppress black root rot of tobacco under gnotobiotic conditions. *EMBO Journal* 8:351-358.
98. Wilson, M. and Lindow, S.E. (1993) Release of recombinant microorganisms. *Annual Review of Microbiology* 47:913-944.
99. Zambryski, P. (1988) Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. *Annual Review of Genetics* 22, 1-30.
100. Zhu J., Oger P.M., Schrammeijer B., Hooykaas P.J.J., Farrand S.K, Winans S.C. (2000) The Bases of Crown Gall Tumorigenesis. *Journal of Bacteriology* 182, 3885-3895.