

## **Supporting Document 1**

### **Application A1080 – Food derived from Herbicide-tolerant Cotton Line MON88701**

## **SAFETY ASSESSMENT**

### **SUMMARY AND CONCLUSIONS**

#### **Background**

Monsanto Company has developed genetically modified (GM) cotton line MON88701 (OECD Unique identifier MON-88701-3) tolerant to two broad spectrum herbicides, dicamba and glufosinate-ammonium. Neither of the enzymes conferring these traits is new to the food supply.

In conducting a safety assessment of food derived from cotton line MON88701, a number of criteria have been addressed including: characterisation of the transferred genes including their origin, function and stability in the cotton genome; the nature of the introduced proteins and their potential to be either allergenic or toxic in humans; compositional analyses and any resultant changes in the whole food. This approach evaluates the intended and any unintended changes in the plant.

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

- potential environmental risks related to the environmental release of GM plants used in food production, and
- the safety of animal feed, or food produced from animals that consumed GM feed.

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

The Applicant anticipates that MON88701 cotton will be commercially cultivated in major cotton-producing countries, including North America and Australia. Food products derived from MON88701 cotton would therefore be expected to enter the Australian and New Zealand food supply via domestic production in Australia, and via imported products.

#### **History of Use**

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

## **Molecular Characterisation**

Two novel gene expression cassettes are present in MON88701 cotton. One contains the *dmo* gene which was isolated from soil bacteria, and encodes the enzyme dicamba mono-oxygenase (DMO), an enzyme that inactivates the herbicide dicamba. The other contains the *bar* gene that encodes phosphinothricin N-acetyltransferase (PAT), a protein conferring tolerance to herbicides containing glufosinate ammonium (phosphinothricin).

Comprehensive molecular analyses of MON88701 indicate there is a single insertion site comprised of one complete copy of the two gene expression cassettes. Plasmid backbone analysis shows no extraneous sequences derived from the plasmid were incorporated into the cotton genome. The introduced genetic elements are stably inherited from one generation to the next. Bioinformatic analyses of ORFs in the T-DNA and junction regions in MON88701 demonstrate no novel polypeptides with relevant homology to proteins that are known to be toxic, allergenic or have other biologically adverse properties.

## **Characterisation of Novel Proteins**

The two proteins newly expressed in MON88701 cotton are DMO and PAT. Analysis of MON88701 plant tissues shows that DMO and PAT are detectable in leaves, roots, pollen and seed at low levels; DMO levels correspond to 0.008% of total protein (80 ppm) and PAT levels correspond to 0.002% of total protein (20 ppm) in MON88701 cottonseed.

The identity and physicochemical and functional properties of the newly expressed proteins were confirmed via a number of laboratory studies. These studies demonstrated that the DMO and PAT proteins conform in size and amino acid sequence to that expected, are not glycosylated in the plant, and exhibit the expected functional enzyme activity. In relation to their potential to be toxic or allergenic in humans, neither DMO nor PAT is derived from a source with known toxicity or allergenicity, and bioinformatic studies confirm the lack of any significant amino acid sequence similarity to known protein toxins or allergens. In addition, digestibility studies demonstrate that both proteins would be rapidly degraded in the gastric environment if ingested. Taken together, the evidence supports the conclusion that DMO and PAT are unlikely to be toxic or allergenic in humans.

## **Herbicide Metabolites**

Expression of the DMO protein in MON88701 results in the demethylation of dicamba and produces the metabolite 3,6-dichlorosalicylic acid (DCSA) which is herbicidally inactive. The results of field trials indicated the residues generated on MON88701 cottonseed as a result of spraying with dicamba are generally low, and that there is some concentration of the major metabolite DCSA in undelinted cottonseed and meal. In establishing herbicide residue limits for dicamba on GM cotton, the proposed definition will need to include DCSA in addition to dicamba and 5-hydroxydicamba, which is the existing analyte description. In the absence of any significant exposure to either parent herbicide or metabolites, the risk to public health and safety is negligible.

## **Compositional Analyses**

Detailed compositional analyses were conducted on acid-delinted seed from MON88701 plants, the conventional control line Coker 130, and eight other conventional cotton varieties grown in field trials across the United States. MON88701 plants were sprayed with dicamba and glufosinate ammonium herbicides at particular stages of growth and all lines were grown under normal agricultural conditions.

Analyses included proximates (crude protein, crude fat, ash and total carbohydrates), ADF, NDF, fatty acids, amino acids, micronutrients (minerals and  $\alpha$ -tocopherol) and anti-nutrients (gossypol and cyclopropenoid fatty acids). The levels of these key constituents in the GM line were compared with those in the non-GM control, as well as to the normal ranges found in conventional cotton varieties already grown commercially or reported in the literature.

A number of statistically significant differences were found in individual analytes between seeds from MON88701 and the control, however all differences were small in magnitude and did not occur at all trial sites. The composition of cotton can vary significantly with the site and the prevailing agricultural conditions, and the differences reported are attributable to normal biological variation. The mean analyte levels in MON88701 seed were within the range established for existing commercial cotton varieties. Overall, the compositional data support the conclusion that there are no biologically significant differences in the levels of key constituents in MON88701 cottonseed when compared with conventional cotton varieties available on the market.

### **Nutritional Impact**

The results of an eight week feeding study in channel catfish demonstrated that cottonseed meal from MON88701 was nutritionally equivalent to meal from the non-GM control and reference cotton varieties. Based on measurement of a set of growth parameters, MON88701 cottonseed meal adequately supported typical growth in the animals when incorporated into the diet.

### **Conclusion**

No potential public health and safety concerns have been identified in the assessment of MON88701 cotton. On the basis of the data required from the Applicant, and other available information, food derived from herbicide tolerant cotton line MON88701 is as safe for human consumption as food derived from conventional cotton varieties.

## TABLE OF CONTENTS

<b>SUMMARY AND CONCLUSIONS .....</b>	<b>I</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>5</b>
<b>1. INTRODUCTION.....</b>	<b>6</b>
<b>2. HISTORY OF USE.....</b>	<b>6</b>
2.1 Host organism .....	6
2.2 Donor organisms .....	8
<b>3. MOLECULAR CHARACTERISATION .....</b>	<b>9</b>
3.1 Method used in the transformation.....	9
3.2 Introduced gene construct .....	9
3.3 Breeding process and analyses.....	12
3.4 Characterisation of the genes in the plant .....	13
3.5 Stability of the genetic changes.....	15
3.6 Antibiotic resistance marker genes .....	16
3.7 Conclusion.....	16
<b>4. CHARACTERISATION OF NOVEL PROTEINS .....</b>	<b>17</b>
4.1 Function of the DMO protein in MON88701 .....	17
4.2 Function of the PAT protein in MON88701 .....	21
4.3 Novel protein expression in plant tissues.....	23
4.4 Potential toxicity of the newly expressed proteins.....	26
4.5 Potential allergenicity of the newly expressed proteins .....	29
4.6 Herbicide metabolites.....	32
4.7 Conclusion.....	35
<b>5. COMPOSITIONAL ANALYSIS .....</b>	<b>36</b>
5.1 Key components.....	36
5.2 Study design and conduct for key components.....	36
5.3 Analyses of key components.....	37
5.4 Conclusion from compositional analysis .....	43
<b>6. NUTRITIONAL IMPACT .....</b>	<b>44</b>
<b>REFERENCES.....</b>	<b>44</b>

## List of abbreviations

ADF	acid detergent fibre
ADI	Acceptable daily intake
ARfD	Acute Reference Dose
<i>bar</i>	<i>bialaphos resistance gene</i>
BLAST	Basic Local Alignment Search Tool
bp	base pairs
DCGA	2,5-dichloro-3,6-dihydroxybenzoic acid
DCSA	3,6-dichlorosalicylic acid
<i>dmo</i>	<i>Dicamba mono oxygenase gene</i>
DMO	dicamba mono-oxygenase
DNA	deoxyribonucleic acid
T-DNA	transfer DNA
dw	dry weight
EFSA	European Food Safety Authority
ELISA	enzyme linked immunosorbent assay
EPA	Environmental Protection Agency – United States of America
FAO	Food and Agriculture Organization of the United Nations
FARRP	Food Allergy Research and Resource Program
FSANZ	Food Standards Australia New Zealand
GM	genetically modified
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
ILSI	International Life Sciences Institute
kb	kilobase
kDa	kilo Dalton
LB	left border
LC/MS	high performance liquid chromatography/electrospray mass spectrometry
MALDI-TOF	Matrix-assisted laser desorption/ionisation-time of flight
NDF	neutral detergent fibre
OECD	Organisation for Economic Co-operation and Development
OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
PAT	phosphinothricin N-acetyltransferase
PCR	polymerase chain reaction
L-PPT	L-phosphinothricin
RB	right border
RBD	Refined, bleached, deodorised
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
U.S.	United States of America
WHO	World Health Organisation

## 1. Introduction

Cotton line MON88701 has been genetically modified for tolerance to two herbicides: dicamba and glufosinate ammonium. Dicamba is used to control broadleaf weeds from pre-emergence to seven days pre-harvest. Glufosinate ammonium-based herbicides are used for broad spectrum weed control from emergence through to early bloom growth stage. The Applicant claims both herbicides together provide effective weed management for this crop, including the control of weeds that are tolerant to glyphosate.

Tolerance to dicamba is achieved through expression of a dicamba mono-oxygenase (DMO) protein encoded by a gene from *Stenotrophomonas maltophilia*. The DMO protein rapidly demethylates dicamba to the herbicidally inactive metabolite 3,6-dichlorosalicylic acid (DCSA). Expression of the enzyme phosphinothricin N-acetyltransferase (PAT), encoded by the *bar* gene from *Streptomyces hygroscopicus*, confers tolerance to glufosinate herbicides. The PAT protein acetylates the free amino group of glufosinate to produce the herbicidally inactive metabolite, 2-acetamido-4-methylphosphinico-butanoic acid (N-acetyl glufosinate). Both newly expressed proteins have been assessed previously by FSANZ in a range of GM crops.

Using traditional plant breeding methods, the Applicant intends to combine MON88701 cotton with other approved herbicide-tolerant GM lines, and possibly also with previously approved insect-protected lines to create stacked events with multiple agronomic traits.

Initially MON88701 cotton will be grown in North America, but the Applicant intends to apply at some future date for a licence to grow the crop commercially in Australia. Therefore, if approved, food from this line may enter the Australian and New Zealand food supply both as imported and domestically-produced food products. There are no plans to grow this GM cotton variety in New Zealand. The main food products derived from cotton are edible cottonseed oil and linters.

## 2. History of use

### 2.1 Host organism

The host organism is cultivated cotton (*Gossypium hirsutum* L.). Cotton is one of the oldest cultivated crops and is grown primarily as a fibre crop, providing almost 50% of the textile fibre used in the world (OECD, 2004; OGTR 2008). Only the cotton boll, which develops from the plant ovary, is used for either textile fibre or food/feed. The cotton boll, once harvested, is processed ('ginned') to separate the fibre from the cottonseed.

Cottonseed is processed into four major by-products: oil, meal, hulls and linters (Figure 1), of which the oil and linters are typically used as human food. By weight, processing of cottonseed typically yields 16% oil, 45% meal, 26% hulls and 9% linters, with 4% lost during processing (Cherry, 1983). Food products are limited to highly processed products because of the presence of natural toxicants (gossypol) and anti-nutrients (cyclopropenoid fatty acids) in unprocessed cottonseed. Cottonseed oil has been used safely for human food for over a century. Meal and hulls are mainly used as livestock feed.

The fatty acid profile of cottonseed oil comprises mainly oleic and linoleic acids. The natural oil has a strong and unpleasant flavour and requires a deodorisation process to render it palatable. 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 cyclopropanoid 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.

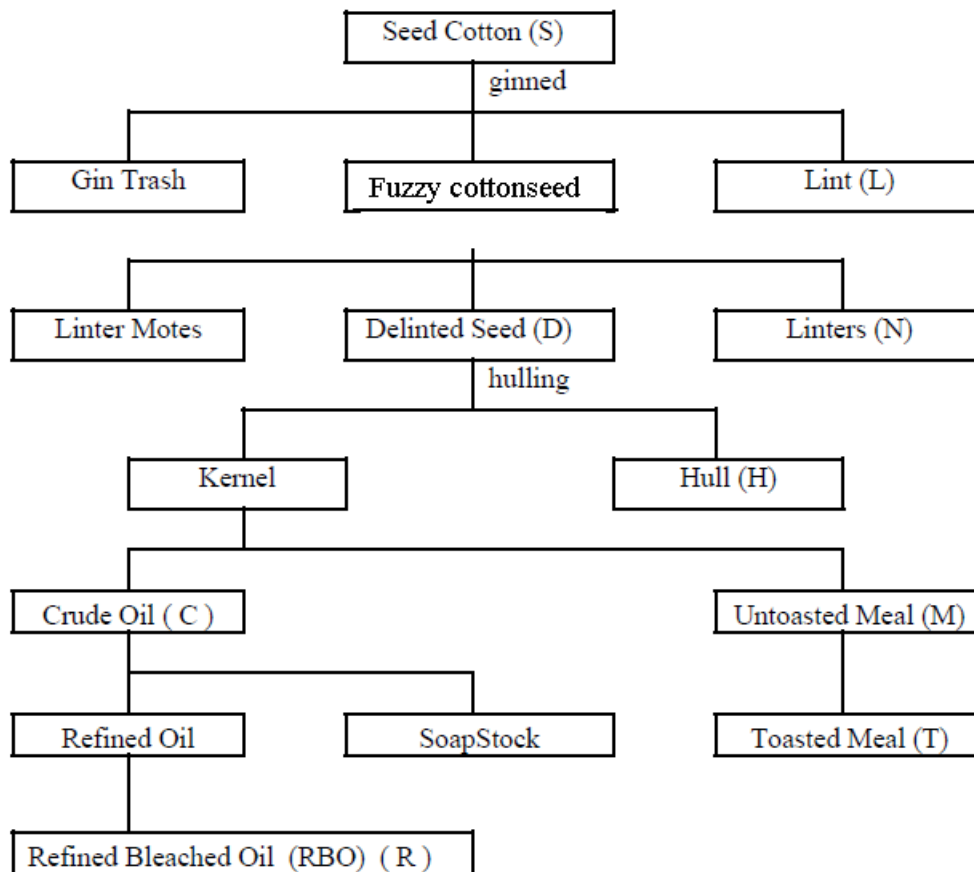


Figure 1: The major processed fractions obtained from cottonseed

The material left after extraction of crude cottonseed oil is cottonseed meal. This product is not used for human consumption in Australia or New Zealand. Cottonseed meal is permitted to be used for human food (after processing) in the U.S. and other countries, but is primarily sold for stock feed. The levels of gossypol in the meal after extraction are reduced by approximately half.

Cotton is not grown in New Zealand. Australia has significant plantings of the crop although the area varies annually due largely to prevailing environmental factors. In the 2006 – 2007 season, 92% of the commercial cotton planted in Australia was genetically modified (Molony and Hassall, 2008) and all traits were for protection against insect pests and/or tolerance to a herbicide (OGTR, 2008b). Although fibre is seen as the main product, cotton is also Australia's major oilseed crop.

Most cottonseed is exported as fuzzy seed<sup>1</sup> destined for animal feedlots but a proportion of the seed is retained to produce oil, mainly for domestic use. In 2009, some 60,286 tonnes of oil was produced in Australia (FAOSTAT – available at <http://faostat3.fao.org/home/index>).

The cotton cultivar ‘Coker 130’ was used as the parental variety for the genetic modification described in this Application, and thus is regarded as the near-isogenic line for the purposes of the comparative assessment. It is not grown commercially in Australia. ‘Coker’ cultivars are U.S. cultivars that are widely used to produce GM cotton lines because they can be readily cultured and regenerated in the laboratory (OGTR, 2008a). Traits introduced into ‘Coker’ cultivars are transferred to commercial cultivars by backcrossing.

## **2.2 Donor organisms**

### **2.2.1 *Stenotrophomonas maltophilia***

The *dmo* gene is derived from the bacterium *Stenotrophomonas maltophilia* strain DI-6, isolated from soil at a dicamba manufacturing plant (Krueger, 1989). This gene and its source were assessed in Application A1063, food derived from herbicide-tolerant soybean line MON87708. The organism was originally named *Pseudomonas maltophilia*, but was subsequently changed to *Xanthomonas maltophilia*, before it was given its own genus (Palleroni and Bradbury, 1993).

*S. maltophilia* is an aerobic, gram negative bacterium commonly present in aquatic environments and soil. It is also commonly associated with plants (Ryan *et al.*, 2009) and has been isolated from the rhizosphere of wheat, maize, grasses, beet, cucumber, potato, strawberry, sugarcane and rapeseed, and has also been isolated from cottonseed, bean pods and coffee. *S. maltophilia* is widespread in moist sites in domestic houses, particularly in bathrooms and kitchens, as well as a variety of foods such as fruits, vegetables, frozen fish, milk and poultry (Denton *et al.*, 1998).

Although not regarded as an inherently virulent pathogen, *S. maltophilia* has an ability to colonise cells in the respiratory tract and surfaces of medical devices and, when coupled with intrinsic or acquired resistance to antimicrobials, has led to infections in immuno-compromised hospital patients (Looney, 2009), particularly those with ventilator tubes or catheters inserted for prolonged periods of time. Nevertheless, *S. maltophilia* is ubiquitous in the environment and in healthy individuals without causing infection, and is incidentally present in foods without any adverse associations.

### **2.2.2 *Streptomyces hygroscopicus***

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

Although this organism is not used in the food industry directly, the *bar* gene from *S. hygroscopicus* has been used to confer glufosinate ammonium-tolerance in food producing crops over the past decade. The *pat* gene from the closely related species *S. viridochromogenes* produces a protein that is structurally and functionally equivalent to the protein encoded by the *bar* gene (Wehrmann *et al.*, 1996).

---

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



### 3. Molecular characterisation

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

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

#### Studies submitted:

Arackal S.M., Garnaat C.W., Lawry K.R., Ralston L.F., Tian Q. (2011) Molecular Characterisation of Dicamba Glufosinate Tolerant Cotton MON88701. Monsanto Company, MSL0023280 (unpublished).  
Arackal S.M., Deffenbaugh A., Garnaat C.W., Niemeyer K.E., Ralston L.F., Tian Q. (2011) Stability of the DNA Insert and Expression of MON88701 DMO and PAT (bar) Proteins in MON88701. Monsanto Company, MSL0023322 (unpublished).  
Arackal S.M. and Tian Q. (2011) Segregation Analysis of the Coding Sequences Present in Herbicide Tolerant Cotton MON88701 Across Multiple Generations. Monsanto Company, RPN-2011-0089.  
Tu H. and Silvanovich, A. (2011) Bioinformatics Evaluation of the Transfer DNA Insert in MON88701 Utilizing the AD\_2011, TOX\_2011 and PRT\_2011 Databases. Monsanto Company, MSL0023565 (unpublished).  
Kang H.T. and Silvanovich A. (2012) Amended Report for MSL0023585: Bioinformatics Evaluation of DNA Sequences Flanking the 5' and 3' Junctions of Inserted DNA in MON88701: Assessment of Putative Polypeptides. Monsanto Company, MSL0024371 (unpublished).

#### 3.1 Method used in the transformation

MON88701 cotton was developed using *Agrobacterium*-mediated transformation of germinated seeds of the cotton cultivar 'Coker 130' according to published methods (Duncan 2010; Duncan and Ye 2011). The genes of interest (*dmo* and *bar*) were carried on plasmid vector PV-GHHT6997 which is described in more detail below. After co-culturing with the *Agrobacterium* carrying the vector, hypocotyl segments were placed on a sequence of callus-inducing media containing carbenicillin and cefotaxime to inhibit the growth of excess *Agrobacterium*, and glufosinate to inhibit growth of untransformed cells.

Following shoot regeneration and transfer to soil, R<sub>0</sub> plants with normal phenotypic characteristics were selected for further assessment. Suitable R<sub>0</sub> plants were self-pollinated to produce R<sub>1</sub> seed. R<sub>0</sub> and R<sub>1</sub> plants were evaluated for tolerance to dicamba and glufosinate and screened for the presence of the T-DNA. Subsequently, the *dmo* and *bar* homozygous positive R<sub>1</sub> plant was self-pollinated to give rise to R<sub>2</sub> plants. Homozygous positive R<sub>2</sub> plants containing only a single T-DNA insertion were identified by a combination of analytical techniques including dicamba and glufosinate sprays, PCR, and Southern blot analysis. This selection process resulted in the identification of MON88701 as the event with the most desired phenotypic and molecular characteristics.

#### 3.2 Introduced gene construct

The plasmid used in the transformation of cotton to produce MON88701 was PV-GHHT6997. The T-DNA (denoted by the Left and Right Border sequences) contains the two genes of interest, *dmo* and *bar*, and necessary regulatory elements for expression in plant cells. Information on the genetic elements in plasmid PV-GHHT6997 (based on standard cloning vector pBR322) is summarised in Table 1.

**Table 1: Description of the genetic elements contained in PV-GHHT6997**

Genetic element	bp Location on Plasmid	Source	Description & Function	References
<b>T-DNA</b>				
<b>Right Border Region</b>	1-331	<i>Agrobacterium tumefaciens</i>	<ul style="list-style-type: none"> <li>Right border sequence</li> <li>Required for the transfer of the T-DNA into the plant genomic DNA</li> </ul>	Depicker <i>et al.</i> 1982; Zambryski <i>et al.</i> 1982
<i>Intervening sequence</i>	332 - 433		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>P-PC1SV</b>	434-866	Peanut chlorotic streak caulimovirus	<ul style="list-style-type: none"> <li>Promoter from the Full-Length transcript (FLt)</li> <li>Directs transcription of genes in plant cells</li> </ul>	Maiti and Shepherd, 1998
<i>Intervening sequence</i>	867 - 872		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>L-TEV</b>	873-1004	Tobacco etch virus (TEV)	<ul style="list-style-type: none"> <li>5' untranslated region (UTR) leader sequence from the viral RNA</li> <li>Regulates gene expression</li> </ul>	Niepel and Gallie, 1999
<i>Intervening sequence</i>	1005		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>TS-CTP2</b>	1006 - 1233	<i>Arabidopsis thaliana</i>	<ul style="list-style-type: none"> <li>Targeting sequence of the <i>ShkG</i> gene encoding the EPSPS transit peptide region</li> <li>Directs transport of the protein into the chloroplast</li> </ul>	Herrmann <i>et al.</i> 1995; Klee <i>et al.</i> 1987
<b>CS-dmo</b>	1234-2256	<i>Stenotrophomonas maltophilia</i>	<ul style="list-style-type: none"> <li>Codon optimised coding sequence for the dicamba mono-oxygenase (DMO) protein</li> <li>Confers tolerance to dicamba</li> </ul>	Herman <i>et al.</i> 2005; Wang <i>et al.</i> 1997
<i>Intervening sequence</i>	2257-2310		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>T-E6</b>	2311-2625	<i>Gossypium barbadense</i> (cotton)	<ul style="list-style-type: none"> <li>3' UTR sequence of the <i>E6</i> gene encoding a fibre protein involved in early fibre development</li> <li>Directs polyadenylation of the mRNA transcript</li> </ul>	John, 1996
<i>Intervening sequence</i>	2626-2637		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>P-e35S</b>	2638-3249	Cauliflower mosaic virus (CaMV)	<ul style="list-style-type: none"> <li>Promoter from the 35S RNA containing the duplicated enhancer region</li> <li>Directs transcription of genes in plant cells</li> </ul>	Odell <i>et al.</i> 1985; Kay <i>et al.</i> 1987
<i>Intervening sequence</i>	3250-3252		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>L-Hsp70</b>	3253-3348	<i>Petunia hybrida</i>	<ul style="list-style-type: none"> <li>5' UTR leader sequence of the <i>Hsp70</i> gene encoding a plant heat shock protein 70 (HSP70)</li> <li>Regulates gene expression</li> </ul>	Rensing and Maier, 1994; Winter <i>et al.</i> 1988
<i>Intervening sequence</i>	3349-3354		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>CS-bar</b>	3355 - 3906	<i>Streptomyces hygroscopicus</i>	<ul style="list-style-type: none"> <li>Coding sequence of the phosphinothricin N-acetyltransferase (PAT) protein</li> <li>Confers tolerance to glufosinate ammonium herbicide</li> </ul>	Thompson <i>et al.</i> (1987)
<i>Intervening sequence</i>	3907-3911		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	
<b>T-nos</b>	3912-4164	<i>Agrobacterium tumefaciens</i>	<ul style="list-style-type: none"> <li>3' UTR sequence of the nopaline synthase (<i>nos</i>) gene from pTI encoding NOS</li> <li>Directs polyadenylation of the mRNA transcript</li> </ul>	Bevan <i>et al.</i> 1983 Fraley <i>et al.</i> 1983
<i>Intervening sequence</i>	4165-4183		<ul style="list-style-type: none"> <li>Sequence used in DNA cloning</li> </ul>	

Genetic element	bp Location on Plasmid	Source	Description & Function	References
Left Border Region	4184-4625	<i>Agrobacterium tumefaciens</i>	<ul style="list-style-type: none"> <li>• DNA region containing the Left Border sequence</li> <li>• Required for transfer of T-DNA into the plant genome.</li> </ul>	Barker <i>et al.</i> 1983
<b>Plasmid Vector Backbone</b>				
Vector sequences outside of the T-DNA (not transferred to the plant)	4626 - 9379		<ul style="list-style-type: none"> <li>• <i>oriV</i> – origin of replication from the broad host range plasmid RK2 for maintenance of plasmid in <i>Agrobacterium</i></li> <li>• <i>CS-rop</i> – coding sequence for repressor of primer protein from ColE1 for maintenance of copy number in <i>E. coli</i></li> <li>• Ori- pBR322 – origin of replication from plasmid pBR322 for maintenance of plasmid</li> <li>• <i>aadA</i> – bacterial gene conferring resistance to antibiotics spectinomycin and streptomycin.</li> <li>• Intervening sequences</li> </ul>	Stalker <i>et al.</i> 1981 Giza and Huang, 1989 Fling <i>et al.</i> 1985

A schematic diagram of plasmid PV-GHHT6997 (approximately 9.4 kb), shown in Figure 2, shows the arrangement of genes and regulatory elements within the T-DNA region and other elements outside of this region essential for maintenance of the plasmid in bacteria. One of the elements on the plasmid backbone is the selectable marker gene *aadA*, encoding resistance to the antibiotics spectinomycin and streptomycin, and used for cloning purposes only in *E. coli* and *Agrobacterium*. The elements outside of the T-DNA region were not intended to be transferred into the cotton genome, and this was confirmed by Southern blot analysis (see below).

### 3.2.1 *dmo* gene cassette

Expression of the *dmo* gene is under the control of the peanut chlorotic streak caulimovirus (PC1SV) promoter, the tobacco etch virus (TEV) 5' leader sequence, and the 3' untranslated sequence of the *E6* gene (encoding a fibre protein) from cotton. The addition of a target sequence encoding a chloroplast transit peptide CTP2 from *Arabidopsis* in front of the bacterial *dmo* gene, produces a precursor DMO protein of 416 amino acids, which is targeted to chloroplasts. On uptake into the chloroplast, 67 amino acids are cleaved from the precursor protein to produce the mature DMO of 349 amino acids.

The coding sequence of the bacterial *dmo* gene was optimised for expression in plants and therefore does not correspond precisely with the DNA sequence from the source organism, *Stenotrophomonas maltophilia*.

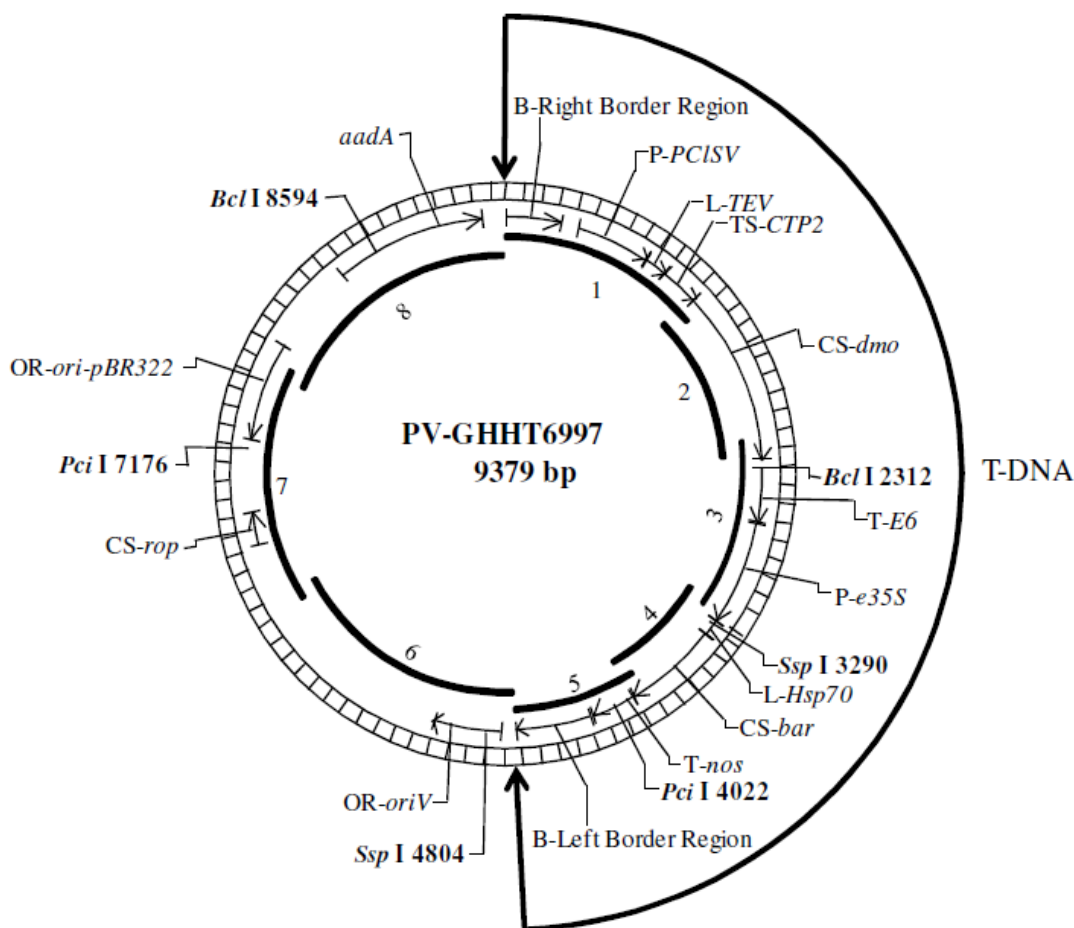


Figure 2: Circular map of transforming plasmid PV-GHHT6997 showing organisation and orientation of all genetic elements. The numbers inside the circle refer to the size and position of eight probes used to characterise the insert by Southern blot analysis.

### 3.2.2 *bar* gene cassette

The *bar* gene from *Streptomyces hygroscopicus* encodes the PAT protein and confers tolerance to herbicides containing glufosinate ammonium (phosphinothricin). In MON88701 cotton, the PAT protein is a single polypeptide of 183 amino acids with molecular mass of approximately 21 kDa.

Expression of the *bar* gene in MON88701 is under the control of the constitutive *e35S* promoter, the *Hsp70* leader, and the *nos* 3' termination element (see Table 1). The promoter is derived from the 35S RNA of the cauliflower mosaic virus (CaMV) and contains the duplicated enhancer region. All regulatory elements in this cassette have been widely used for genetic modification of crop plants and have been assessed previously in other GM crops.

### 3.3 Breeding process and analyses

As described in section 3.1,  $R_0$  (the original transformant) and  $R_1$  plants were evaluated for tolerance to dicamba and glufosinate, and screened for the presence of the T-DNA. Once the *dmo* and *bar* homozygous positive  $R_1$  plant was selected, it was self-pollinated to give rise to  $R_2$  plants. Homozygous positive  $R_2$  plants containing only a single T-DNA insertion were identified by a combination of analytical techniques, PCR, and Southern blot analysis.

After another round of self-pollination to yield R<sub>3</sub> homozygous plants, MON88701 was selected for commercial development. The R<sub>3</sub> generation was also used for the molecular characterisation. Plants from the R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>6</sub> generations (each produced by self-pollination) were used to analyse stability of the insert across multiple generations. The R<sub>5</sub> and R<sub>6</sub> generations were used to analyse the expression of the traits and for the compositional analysis (plant tissues other than seed were obtained from R<sub>5</sub>, while seed was taken from R<sub>6</sub>).

### **3.4 Characterisation of the genes in the plant**

Analyses were undertaken in order to fully characterise at the molecular level the genetic modification in cotton line MON88701. These analyses focussed on the exact nature of the inserted genetic elements, and whether any unintended genetic re-arrangements may have occurred as a consequence of the transformation procedure. A range of techniques including Southern blots, PCR and DNA sequence analysis were used for the characterisation.

#### **3.4.1 *Transgene copy number and insertion integrity***

Southern blots were used to determine the number of T-DNA insertions and the sequence integrity of the introduced DNA in MON88701, and test for the presence or absence of plasmid vector backbone sequences that could have been transferred to the plant. Genomic DNA for these analyses was extracted from leaf tissue from the R<sub>3</sub> generation of MON88701 cotton plants, and from Coker 130 plants as the conventional parental control.

Eight probes spanning the entire plasmid were prepared by PCR amplification using PV-GHHT6997 as the template, and radiolabelled. Probes 1-5 (as shown on Figure 2) were used to detect the T-DNA, while probes 6-8 were used to check for the presence of plasmid backbone. The same restriction enzymes were used to cleave MON88701 and parental control DNA, which allowed direct analysis of the distinctive banding pattern obtained with each sample. The restriction enzyme sets were chosen such that each set cleaves once within the inserted T-DNA and at least once within the known DNA flanking the 5' and 3' ends of the insert. In addition, each Southern blot included at least one negative control and one or more positive controls. Parental plant genomic DNA was spiked with digested plasmid DNA from PV-GHHT6997 to serve as a positive hybridisation control.

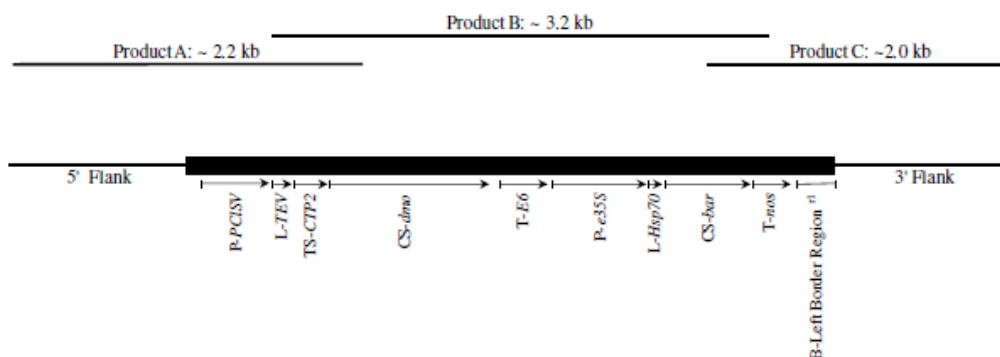
The strategy of using short and long run times on duplicate samples was adopted to maximise the possibility of detecting an insertion elsewhere in the genome, which could be overlooked if that band co-migrated on the gel with another (expected) band. The long run allowed for greater resolution of large molecular weight DNA, whereas the short run allowed for separation of small molecular weight DNA on the gel, and more accurate size estimates.

Analysis of multiple Southern blots indicated that one copy of the *dmo* expression cassette and one copy of the *bar* expression cassette were integrated as a single T-DNA at a single locus in MON88701. No additional DNA fragments and no sequences derived from the plasmid backbone were detected in MON88701 genomic DNA.

#### **3.4.2 *DNA sequence analysis***

Analysis of PCR products and DNA sequencing complement the results obtained with Southern blots, and allow the exact nature of the T-DNA insert in MON88701 to be described. This allows confirmation of the organisation and sequence of each genetic element transferred into the plant, relative to the sequences in the T-DNA region of the transforming plasmid.

Genomic DNA was extracted from leaf tissue of MON88701 cotton plants and the conventional control. PCR primer pairs were designed to amplify three overlapping DNA regions spanning the entire length of the insert and the adjacent plant genomic DNA flanking the 5' and 3' ends of the insert, as depicted in Figure 3. PCR was also performed on conventional control DNA using a primer pair that hybridised to genomic sequence in the 5' and 3' flanking regions, to obtain an amplicon corresponding to the site of insertion in the untransformed parental control.



*Figure 3: The upper illustration shows the approximate size and position of three PCR amplicons (Products A, B, C) used for DNA sequencing of the T-DNA and flanking regions in MON88701 cotton. The lower illustration shows the size and order of genes in the T-DNA insertion.*

The DNA sequence analysis determined that the insert in MON88701 is 4105 bp and is identical to the T-DNA sequence of the transforming plasmid PV-GHHT6997. Approximately 1100 bp of genomic sequence was obtained at both ends of the T-DNA insert. Alignments between the parental control sequence obtained from this analysis and the sequences immediately flanking the 5' and 3' ends of the MON88701 insert were separately performed to determine the integrity and genomic organisation of the insertion site in MON88701. The alignment showed sequence identity except for a small deletion (123 bp) in MON88701, which occurred at the insertion site. There are numerous examples of these types of deletions in the literature. Based on the normal phenotype of MON88701 cotton, the small deletion raises no safety concerns.

### **3.4.3 Open reading frame (ORF) analysis**

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

Putative peptides consisting of a minimum length of eight amino acids were compared to toxin, allergen and all protein databases using typical bioinformatics tools. The FASTA sequence alignment tool was used to assess structural relatedness between the query sequence from MON88701 and any protein sequences in the AD\_2011, TOX\_2011 and PRT\_2011 databases. The extent of structural relatedness was evaluated by detailed visual inspection of the alignment, the calculated percent identity and alignment length.

Alignments with an *E*-score less than  $1 \times 10^{-5}$  may reflect sequences with shared structure and function (Ladics *et al.* 2007). In addition to structural similarity, each putative polypeptide was screened for short peptide matches using a pair-wise comparison algorithm. In terms of potential allergenicity, eight contiguous and identical amino acids were defined as the minimum requirements for a possible immunologically relevant epitope (Silvanovich *et al.* 2006).

The results of comparative searches of the allergen and toxin databases revealed no relevant structural similarity to known allergens or toxins for any of the putative polypeptides in the T-DNA insert reading frames. No short peptide matches of eight amino acids were found between MON88701 sequences and proteins in the allergen database. The PRT\_2011 database was interrogated using the same process, and the DMO and PAT proteins were positively identified, as expected. Alignments showed 100% identity over 340 amino acids for the DMO protein, and over 183 amino acids for the PAT protein. Other alignments with significant *E*-scores were examined in more detail but none of these were considered to indicate a putative translation product with adverse biological activity.

Putative polypeptides encoded by DNA spanning the 5' and 3' genomic junctions of the MON88701 inserted DNA were investigated in the same manner using bioinformatics tools. No relevant structural similarity to known toxins or allergens, nor proteins with adverse biological activity were identified for any of the ORFs. No short matches of eight amino acids were observed between junction sequences in MON88701 and proteins in the allergen database.

Overall, the results from the bioinformatic analyses demonstrate that the insertion of the T-DNA in MON88701 is unlikely to give rise to novel polypeptides with significant similarity to known proteins showing toxic, allergenic or other biologically adverse properties.

### **3.5 Stability of the genetic changes**

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

#### **3.5.1 Patterns of inheritance**

During development of MON88701 cotton, phenotypic and genotypic segregation data were recorded to assess the inheritance and stability of the inserted T-DNA using Chi-square ( $\chi^2$ ) analysis over several generations. The  $\chi^2$  analysis is based on comparing the observed segregation ratio to the expected segregation ratio according to Mendelian principles.

Using traditional breeding techniques, homozygous positive R<sub>3</sub> plants (see Section 3.3) were crossed to a proprietary inbred cotton, which did not contain the *dmo* or *bar* coding sequence, to produce hemizygous F<sub>1</sub> seed. The F<sub>1</sub> plants were crossed again with a proprietary inbred cotton to produce BC1F<sub>1</sub> seed. Plants from the BC1F<sub>1</sub> generation were assessed using a glufosinate herbicide application to select for plants expressing the MON88701 T-DNA. The plants that survived the herbicide application were confirmed as hemizygous for the introduced traits. The hemizygous BC1F<sub>1</sub> plants were self-pollinated to produce BC1F<sub>2</sub> seed.

Plants from this generation were assessed again with application of glufosinate and the surviving plants were analysed for the MON88701 T-DNA insert using molecular techniques.

Inheritance of the MON88701 T-DNA was assessed in the R<sub>1</sub>, BC1F<sub>1</sub> and BC1F<sub>2</sub> generations, evaluating 173, 261 and 118 individual plants respectively. The results of the  $\chi^2$  analysis showed the observed segregation ratio was consistent with the expected segregation ratio for each generation of plants. These results support the conclusion that the T-DNA in MON88701 is at a single locus within the cotton genome and is inherited according to Mendelian principles of inheritance.

### **3.5.2 Generational stability**

To demonstrate stability of the genetic modification, Southern blot analysis was performed using genomic DNA extracted from leaf tissue from five breeding generations of MON88701. As described in section 3.3, plants from the R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>6</sub> generations (each produced by self-pollination) were analysed for stability of the T-DNA insert by using two probes simultaneously. Any instability or unexpected movement associated with the insert would be readily detected as extra bands within the MON88701 fingerprint on the Southern blot. The positive control in this system was conventional parental DNA spiked with PV-GHHT6997. The negative control was parental DNA, digested and probed in the same way as other samples. The Southern blot results showed exactly the same pattern of bands for all generations tested, and all consistent with those produced by plants of the R<sub>3</sub> generation. As noted elsewhere, the R<sub>3</sub> generation was selected for commercial development and was fully characterised in other studies.

### **3.5.3 Phenotypic analysis**

Leaf tissue from the same five generations (R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>6</sub>) and the conventional control was used in Western blots to assess the presence of the DMO and PAT proteins across multiple generations of MON88701. The DMO and PAT proteins were separately produced in *E. coli* to be used as reference protein standards for the identification of the DMO and PAT proteins produced in MON88701 cotton plants. The results of the Western blots showed that successive generations of MON88701 plants produced DMO and PAT proteins that were indistinguishable from the *E. coli* reference proteins in terms of their mobility through polyacrylamide gels.

### **3.6 Antibiotic resistance marker genes**

No antibiotic resistance marker genes are present in MON88701 cotton. The herbicide tolerance traits were used to select the appropriate lines of interest. In addition, the molecular analysis shows that no plasmid backbone was transferred to the cotton genome during transformation.

### **3.7 Conclusion**

Cotton line MON88701 contains two new genes derived from bacteria that were introduced on a single expression cassette via *Agrobacterium*-mediated transformation. The *dmo* gene from *Stenotrophomonas maltophilia* encodes the DMO protein, a mono-oxygenase to inactive dicamba herbicides, and the *bar* gene from *Streptomyces hygroscopicus* encodes phosphinothricin N-acetyltransferase (PAT), an enzyme conferring tolerance to herbicides containing glufosinate ammonium (phosphinothricin). Detailed molecular analyses of MON88701 cotton indicate that one complete copy of the two-gene expression cassette is present at a single insertion site in the cotton genome. Plasmid backbone analysis shows no extraneous sequences derived from the plasmid were incorporated into the cotton genome.



The introduced genetic elements and the expression of new proteins in MON88701 cotton were shown by phenotypic analysis and molecular techniques to be stably inherited from one generation to the next across multiple generations. The pattern of inheritance supports the conclusion that the herbicide-tolerance traits are a single locus within the cotton genome and are passed on in accordance with Mendelian principles of inheritance.

#### 4. Characterisation of novel proteins

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

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

The two proteins newly expressed in MON88701 cotton are DMO and PAT. Laboratory analyses and bioinformatics studies were done to determine the identity, physicochemical properties, *in planta* expression, bioactivity and potential toxicity and allergenicity of these two proteins. Because the expression of transgenic proteins *in planta* is usually too low to allow purification of sufficient quantities for these studies, a bacterial expression system was used to generate larger quantities of both proteins. The equivalence of the proteins produced in *E. coli* to the plant-produced proteins was determined as part of the protein characterisation.

##### 4.1 Function of the DMO protein in MON88701

###### Studies submitted:

D Chandu, JM Harrison, TC Lee, KS Crowley and E Bell (2012). Characterisation of Dicamba Mono-oxygenase (DMO) Protein Purified from the Cottonseed of MON88701 and Comparison of the Physicochemical and Functional Properties of the Plant-Produced and *Escherichia coli*-Produced DMO Proteins. Monsanto Company MSL0023517 (unpublished)

D Chandu, JM Harrison, TC Lee, KS Crowley and E Bell (2012). Characterisation of Phosphinothricin N-Acetyltransferase (*bar*) Protein Purified from the Cottonseed of MON88701 and Comparison of the Physicochemical and Functional Properties of the Plant-Produced and *Escherichia coli*-Produced PAT (*bar*) Proteins. Monsanto Company MSL0023428 (unpublished)

Tolerance to dicamba in cotton line MON88701 is conferred by the expression in the plant of DMO, a mono-oxygenase enzyme, that catalyses the demethylation of dicamba (3,6-dichloro-2-methoxy benzoic acid) to a non-herbicidal compound DCSA (3,6-dichlorosalicylic acid) and formaldehyde. The active form of the enzyme is a trimer of DMO monomers. This trimeric quaternary structure is the native form of the enzyme observed during crystallisation and is an absolute requirement for electron transfer (oxidation of NADH) and catalysis.

Wildtype DMO was initially purified from the *S. maltophilia* strain D1-6 that was isolated from soil at a dicamba manufacturing plant (Krueger *et al* 1989).

The newly expressed DMO protein is active in the chloroplast. To target the protein to this organelle, additional coding sequence (CTP2 derived from *Arabidopsis*, Table 1) was included in the genetic construct to enable translation of a precursor protein with a chloroplast transit peptide of 76 amino acids at the N-terminus of the DMO protein. Although leader (transit) peptides are typically precisely removed from the precursor protein as uptake occurs, in some cases there is alternate processing resulting in forms of the protein with parts of the transit peptide remaining at the N-terminus. Such alternate processing of DMO has been observed with uptake into chloroplasts in MON88701 cotton.

Analysis of cottonseed extracts from MON88701 determined that the chloroplast-processed DMO protein has retained nine amino acids (VMSSVSTAC) on the N-terminus originating from the chloroplast transit peptide. The resulting mature polypeptide corresponds to 349 amino acids and has an apparent molecular weight of 39.5 kDa. The sequence of the 416 amino acids comprising the full DMO precursor protein in MON88701 was deduced from the coding nucleotide sequence in PV-GHHT6997 and is shown in Figure 4 below, along with the position of the amino acids retained from the transit peptide.

Except for the nine amino acids derived from the C-terminus of the chloroplast transit peptide and an additional leucine at position two, the mature DMO protein in MON88701 has an identical sequence to the wildtype DMO protein from the DI-6 strain of *S. maltophilia* (Herman *et al*. 2005). The differences between the wildtype and MON88701 DMO proteins were not expected to result in changes in overall structure, activity or enzyme specificity because the N-terminus and position two are sterically distant from the catalytic site of the enzyme (D'Ordine *et al* 2009; Dumitru *et al* 2009). Several different approaches outlined below were used to demonstrate that the DMO protein expressed in MON88701 cotton has the expected functionality and exhibits equivalent physicochemical properties to a bacterially-produced DMO protein used as reference material in later studies.

```

1  MAQVSRICNG VQNPSLISNL SKSSQRKSPL SVSLKTQQHP RAYPISSSWG
51  LKKSGMTLIG SELRPLKVMS SVSTACMLTF VRNAWYVAAL PEELSEKPLG
101 RTILDTPALAL YRQPDGVVAA LLDICPHRFA PLSDGILVNG HLQCPYHGLE
151 FDGGGQCVHN PHGNGARPAS LNVRSFPVVE RDALIWIWPG DPALADPGAI
201 PDFGCRVDPA YRTVGGYGHV DCNYKLLVDN LMDLGHAQYV HRANAQTDAF
251 DRLEREVIVG DGEIQALMKI PGGTPSVLMA KFLRGANTPV DAWNDIRWNK
301 VSAMLNFIAV APEGTPKEQS IHSRGTHILT PETEASCHYF FGSSRNFGID
351 DPEMDGVLRs WQAQALVKED KVVVEAIERR RAYVEANGIR PAMLSCEAA
401 VRVSREIEKL EQLEAA

```

Figure 4: The full precursor DMO protein of 416 amino acids, as deduced from the nucleotide sequence of PV-GHHT6997. The first 76 amino acids corresponding to the chloroplast transit peptide (CTP2) are underlined. In MON88701, the mature DMO protein is comprised of 349 amino acids, starting at valine at position 68. The double underlined sequence corresponds to the nine amino acids from CTP2 that are retained at the N-terminus of the mature DMO protein upon uptake into the chloroplast.

#### 4.1.1 DMO specificity

The Applicant investigated whether the active site in the DMO enzyme in MON88701 would interact with substrates other than dicamba, particularly a number of chemically-similar endogenous compounds in plants. Dicamba interacts with amino acids in the catalytic site of DMO through the carboxylic acid moiety and its chlorine atoms, which are required for catalysis. The Applicant identified that chlorinated compounds with structures similar to dicamba are not common in plants and other eukaryotes.

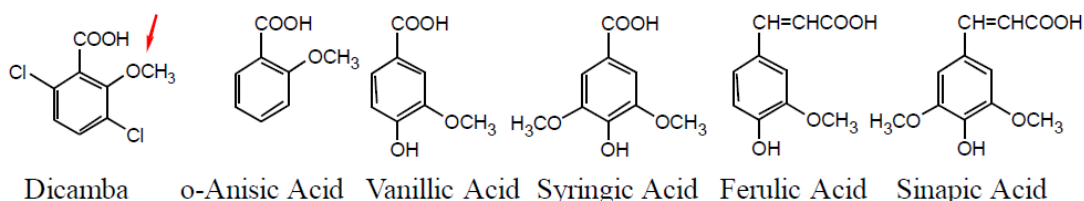


Figure 5: Dicamba and potential endogenous substrates tested in vitro for binding with DMO. The arrow indicates the methyl group removed by DMO.

A set of potential endogenous compounds found in cotton, corn and soybean was selected for evaluation based on structural similarity to dicamba. The compounds tested were o-anisic acid (2-methoxybenzoic acid), vanillic acid (4-hydroxy-3-methoxybenzoic acid), syringic acid (3,5-dimethoxy-4-hydroxybenzoic acid), ferulic acid [3-(4-hydroxy-3-methoxy-phenyl)prop-2-enoic acid] and sinapic acid [3-(4-hydroxy-3,5-dimethoxyphenyl) prop-2-enoic acid] (see Figure 4). The assay mixture included NADH, reductase, ferredoxin and DMO, identical to wildtype enzyme except for the N-terminal histidine tag used for purification purposes.

The disappearance of potential substrates and the formation of potential oxidation products were monitored using LC-UV and LC-MS. In these experiments, dicamba was the only compound in this group to be metabolised by the histidine-tagged DMO enzyme.

#### 4.1.2 Characterisation and equivalence of DMO produced in the laboratory

In order to accept the relevance of the safety data generated using *E. coli*-produced MON88701 DMO, it was necessary to demonstrate equivalence with the DMO protein produced in MON88701 cotton. Five separate analytical techniques were used to characterise the DMO proteins from both sources. Four of these analyses could be used to demonstrate equivalence of the plant- and *E. coli*-produced DMO proteins. A small quantity of the DMO protein was purified from MON88701 cottonseed to use as reference material. The total protein concentration of the MON88701 DMO was measured using standard techniques to be 0.15 mg/ml.

The techniques used were:

- MALDI-TOF Tryptic Mass Fingerprint Analysis (MALDI-TOF MS)
- Western blot analysis
- SDS-PAGE
- Glycosylation analysis
- Enzyme activity assay

##### 4.1.2.1 MALDI-TOF Tryptic Mass Fingerprint Analysis

The identity of the MON88701 DMO protein was confirmed using MALDI-TOF MS analysis of peptides derived from MON88701 DMO protein digested with trypsin.

This standard technique, also referred to as peptide mapping, matches a sufficient number of observed tryptic peptide fragment masses with those predicted from the amino acid sequence of the DMO protein. Protein identification is considered reliable with this method where  $\geq 40\%$  of the protein sequence is identified by matching the observed masses with the expected masses of the tryptic peptide fragments.

This method identified 19 unique peptides that corresponded to the theoretical masses expected to be produced by trypsin digestion of the DMO protein. This coverage corresponded to 66.5% of the MON88701 DMO protein sequence (232 out of 349 amino acids), and serves as confirmation of the identity of the MON88701 DMO protein.

#### 4.1.2.2 Western blot analysis

Western blot analysis was performed to confirm the identity of the MON88701 DMO protein and compare the immunoreactivity of DMO produced in MON88701 cotton and in *E. coli*. Duplicate samples of three different amounts (0.5, 2 and 6 ng) each of plant- and *E. coli*-produced DMO were loaded on the gel. The membrane was probed with goat anti-DMO polyclonal antibodies. The signal intensities of the immunoreactive bands on the blot were quantified.

The identity of the MON88701 DMO protein was confirmed using antibodies specific for the DMO protein. The Western blot analysis showed the immunoreactive properties of the plant- and *E. coli*-produced DMO protein were equivalent.

#### 4.1.2.3 SDS PAGE

SDS-PAGE was used to determine and compare the apparent molecular weight (mobility) and purity of the plant- and *E. coli*-produced DMO proteins. Based on the results, the MON88701 DMO protein had an apparent molecular weight of 39.5 kDa, whereas the *E. coli*-produced DMO corresponded to 38.7 kDa. The apparent molecular weight of the DMO protein produced in MON88701 cotton and in *E. coli* were considered equivalent within the acceptable limits of this method of analysis. The average purity determined by densitometric analysis of multiple loadings of the MON88701 DMO was 97%.

#### 4.1.2.4 Glycosylation analysis

Glycosylation analysis was used to determine whether the DMO protein produced in MON88701 cotton undergoes post-translational modification with covalently bound carbohydrate moieties. Transferrin, a known glycosylated protein, was used as a positive control in the assay, and the DMO protein produced in *E. coli* was also analysed (assumed as a negative control). No glycosylation signal was observed in the samples corresponding to the DMO protein produced in *E. coli* and that produced in MON88701 cotton. These results are consistent with the DMO protein produced in MON88701 as being non-glycosylated.

#### 4.1.2.5 Enzyme activity

The specific activity of MON88701 DMO and the *E. coli*-produced DMO was determined by quantifying the conversion of 3,6-dichloro-2-methoxybenzoic acid (dicamba) to 3,6-dichlorosalicylic acid (DCSA) via HPLC separation and fluorescence detection. The specific activity was calculated based on the amount of purity-corrected MON88701 DMO protein added to the reaction mixture and expressed as nmol of DCSA produced per minute per mg of MON88701 DMO ( $\text{nmol} \cdot \text{minute}^{-1} \cdot \text{mg}^{-1}$ ).

The specific activities of the MON88701 and *E. coli* – produced DMO proteins were determined to be 5.48 nmol DCSA/minute/mg and 7.23 nmol DCSA/minute/mg, respectively. These activity levels were within acceptable limits of variability, and therefore support the conclusion that the MON88701 and *E.coli* DMO protein preparations were equivalent in terms of functional activity.

## **4.2 Function of the PAT protein in MON88701**

*Streptomyces hygroscopicus* produces an antibiotic compound known as bialaphos. Now also used as a non-selective herbicide, bialaphos is a tripeptide composed of two L-alanine residues and an analogue of glutamate known as L-phosphinothricin (L-PPT) (see Thompson *et al.*, 1987), and more recently known also as glufosinate ammonium. Free L-PPT released from bialaphos by peptidases (or applied directly as a synthetic herbicide) inhibits the enzyme glutamine synthetase which in turn leads to rapid accumulation of ammonia and subsequent death of plant cells.

In order to avoid the inherent toxicity of its own product, *S. hygroscopicus* also produces the PAT protein, which chemically inactivates bialaphos. Expression of the *bar* gene from *S. hygroscopicus* in transgenic plants confers tolerance to glufosinate ammonium based herbicides (Murakami *et al.*, 1986).

Phosphinothricin acetyltransferase (PAT) is an acetyl transferase enzyme with specificity for both L-PPT and demethylphosphinothricin (DMPT). In the presence of acetyl-CoA, PAT catalyses the acetylation of the free amino group of L-PPT to N-acetyl-L-PPT, a herbicidally-inactive compound. Plants expressing PAT therefore tolerate applications of L-PPT (glufosinate ammonium based herbicide) without accumulating toxic levels of ammonia. The kinetics and substrate specificity of the PAT enzyme are well characterised; it has a high specificity for L-PPT and has been shown to have a very low affinity to related compounds and amino acids, and even excess glutamate is unable to block the L-PPT-acetyltransferase reaction (Thompson *et al.*, 1987).

### **4.2.1 Previous assessment of PAT**

The *bar* gene from *Streptomyces hygroscopicus* and the *pat* gene from *Streptomyces viridochromogenes* both encode the PAT protein and have been used in other GM crops to confer tolerance to glufosinate-ammonium herbicides. Consequently, the PAT protein has been assessed previously by FSANZ in sixteen GM food applications over an extended period since 2000. Despite this history of assessment, the Applicant was required to fully characterise the PAT protein newly expressed in MON88701 cotton and determine its equivalence to the PAT protein produced in *E. coli*, as the latter was used in a range of studies supporting the assessment of MON88701 cotton.

### **4.2.2 Characterisation and equivalence of PAT produced in the laboratory**

As described above for the DMO protein, separate analytical techniques were used to characterise the PAT protein produced in MON88701 cotton and compare it directly with that produced in *E. coli* to demonstrate that the proteins from both sources are chemically and functionally equivalent. A small quantity of the PAT protein was purified from approximately 1 kg of MON88701 cottonseed for these studies.

For the PAT protein, the six techniques used were:

- N-terminal sequencing
- MALDI-TOF Tryptic Mass Fingerprint Analysis (MALDI-TOF MS)
- Western blot analysis

- SDS-PAGE
- Glycosylation analysis
- Enzyme activity assay

#### 4.2.2.1 N-terminal sequence analysis

N-terminal sequencing of the first 15 amino acids (carried out by automated Edman degradation chemistry) was used to confirm the identity of the MON88701 PAT protein. The expected amino acid sequence was deduced from the nucleotide sequence of the *bar* gene. The results of the N-terminal sequencing showed PAT from MON88701 corresponded to the expected amino acid sequence, minus the N-terminal methionine residue (residues 2 to 16 were an exact match). Since protein N-terminal methionine excision is a common cotranslational process that occurs in most newly synthesized proteins, the results provide confirmation of the identity of the PAT protein expressed in MON88701 cotton.

#### 4.2.2.2 MALDI-TOF Tryptic Mass Fingerprint Analysis

Ten unique peptides were identified that corresponded to the theoretical masses expected to be produced by trypsin digestion of the PAT protein. In total, this experimentally determined tryptic mass coverage corresponded to 84.7% of the PAT protein sequence (155 out of 183 amino acids), and serves as additional confirmation of the identity of the PAT protein in MON88701.

#### 4.2.2.3 Western blot analysis

Western blot analysis used anti-PAT (*bar*) polyclonal antibodies raised in goat. Duplicate samples of three different amounts (2, 4 and 6 ng) each of plant- and *E. coli*- produced PAT proteins were loaded on the gel. The signal intensities of the immunoreactive bands on the blot were quantified. This method of analysis determined that the immunoreactive properties of both the plant- and *E. coli*-produced PAT proteins were equivalent, and confirmed the identity of the PAT protein expressed in MON88701 cotton.

#### 4.2.2.4 SDS PAGE

The plant- and *E. coli*-produced PAT proteins were analysed by SDS-PAGE to compare their apparent molecular weights (MW) and purity. After staining, the MON88701-produced PAT protein was shown to have an apparent MW of 24 kDa, which compared closely with the MW of the *E. coli*-produced PAT (25.0 kDa). The average purity, rounded to the nearest whole number, was determined by densitometric analysis of multiple loadings of the MON88701 PAT to be around 100%.

#### 4.2.2.5 Glycosylation analysis

The possibility of post-translational modification of the PAT protein in MON88701 was tested using an ECL™ Glycoprotein Detection Module (GE Healthcare). As for the DMO protein, a known glycosylated protein, transferrin, was used as a positive control and the PAT protein produced in *E. coli* was included in the assay. A clear positive glycosylation signal was observed at the expected MW (approx. 80 kDa) for the transferrin samples. No glycosylation signal was observed in the samples corresponding to the PAT protein produced in *E. coli* and that produced in MON88701 cotton. These results are consistent with the PAT protein produced in MON88701 cotton as being non-glycosylated.

#### 4.2.2.6 Enzyme activity

The specific activities of the PAT enzyme produced in MON88701 cotton and *E.coli* were determined spectrophotometrically by measurement of the release of coenzyme A (CoA) from acetyl-CoA upon transfer of an acetyl- group to phosphinothricin.

Based on this assay system, the specific activities of the MON88701- and *E. coli* – produced PAT proteins were determined to be 36.4  $\mu\text{mol/minute/mg}$  and 46.2  $\mu\text{mol/minute/mg}$  respectively. These activity levels were within acceptable limits of variability, and therefore support the conclusion that the MON88701 and *E.coli* PAT proteins are equivalent in terms of functional enzyme activity.

### 4.3 Novel protein expression in plant tissues

#### Study submitted:

A.E. Deffenbaugh and K.E. Niemeyer (2012). Amended Report for MSL0024006: Assessment of MON88701 DMO and PAT (*bar*) Protein Levels in Tissues from Dicamba Glufosinate Tolerant Cotton (MON88701) Produced in U.S. Field Trials during 2010. MSL 0024523 (unpublished).

Expression of the DMO and PAT proteins is expected in all plant tissues since the *dmo* and *bar* genes are coupled with promoters from common plant viruses that give rise to constitutive expression (refer to Table 1).

Tissues were harvested from MON88701 cotton plants grown in four replicate plots planted in a randomised complete block field design during the 2010 growing season from the following eight field sites in the U.S.: Arkansas, Georgia, Kansas, Louisiana, North Carolina, New Mexico, South Carolina and Texas. The plots were treated at the 3-5 leaf stage with glufosinate ammonium herbicide at the label rate (0.5 lbs active ingredient/acre), and at the 6-10 leaf stage with dicamba herbicide at the proposed label rate (0.5 lbs acid equivalent/acre). Tissue samples collected included seed, pollen, root and over-season leaf (OSL-1 to OSL-4). All protein levels for all types of tissues were calculated on a fresh weight basis. Moisture content was then measured for all tissue types and protein levels converted and reported on a dry weight basis.

#### 4.3.1 Expression of DMO in MON88701 cotton

The levels of DMO in collected plant tissues from MON88701 were determined by a validated enzyme-linked immunosorbent assay (ELISA). The results are presented in Table 2. Moisture content was not measured for pollen due to a limited amount of material collected.

As shown in Table 2, the highest levels of DMO protein were found in leaf, ranging from OSL-2 to OSL-4 stages; the lowest levels of DMO protein were found in the seed (on a dry weight basis). On a fresh weight basis, the lowest levels were measured in pollen and root, however the levels in seed were also lower than in leaf tissue at all growth stages.

**Table 2: Summary of DMO Protein Levels in Tissues from MON88701 cotton - U.S. Field Trials conducted in 2010**

Tissue <sup>1</sup>	Development Stage <sup>2</sup>	Days After Planting (DAP)	MON 88701 DMO Mean (SD) Range (µg/g fw) <sup>3</sup>	MON 88701 DMO Mean (SD) Range (µg/g dw) <sup>4</sup>	LOD/LOQ <sup>5</sup> (µg/g fw)
OSL-1	2-4 leaf	14-25	27 (7.6) 13 – 42	180 (52) 110 – 280	0.168/0.313
OSL-2	4-7 leaf	25-37	41 (12) 19 – 65	240 (69) 110 – 380	0.168/0.313
OSL-3	9 leaf - Full flower	35-99	52 (17) 24 – 97	240 (75) 91 – 410	0.168/0.313
OSL-4	Full flower – Cutout	70-121	57 (18) 0.70 – 91	230 (59) 2.8 – 310	0.168/0.313
Root	50% open flower – Full flower	62-99	14 (3.7) 8.2 – 21	43 (12) 26 – 72	0.136/0.313
Pollen	50% open flower – Full Flower	68-99	14 (28) 0.31 – 110	NA (NA) NA	0.043/0.125
Seed	Maturity	148-183	20 (4.6) 8.2 – 29	21 (5.0) 8.9 – 33	0.059/0.313

<sup>1</sup>OSL=over-season leaf. Seed=black seed (ginned and delinted)

<sup>2</sup>Stage of development of crop when tissue collected.

<sup>3</sup>Protein levels are expressed in microgram of protein per gram of tissue on a fresh weight (fw) basis. For each tissue across all sites n=32, except OSL-3 n=31, OSL-4 n=28, pollen n=29.

<sup>4</sup>Protein levels expressed on a dry weight (dw) basis.

<sup>5</sup>LOQ=limit of quantitation; LOD=limit of detection

#### 4.3.1.1 Dietary exposure to DMO

Cottonseed is the source of the only human food currently produced from MON88701 cotton. This is used to produce refined, bleached and deodorised (RBD) oil, and to a smaller extent, linters which are highly processed and consist of nearly pure cellulose (>99%). The mean level of DMO in MON88701 cottonseed was 21 µg/g dw (Table 2). As the mean percent dry weight of total protein in MON88701 cottonseed is 28%, the amount of DMO protein was calculated to be 0.008% of total protein (80 ppm) in MON88701 cottonseed. In view of the low levels of DMO in cottonseed, its presence in refined, bleached and deodorised (RBD) oil will be difficult to detect. This suggests that dietary exposure to DMO is likely to be negligible.

#### 4.3.2 Expression of PAT in MON88701 cotton

The levels of PAT protein in collected plant tissues from MON88701 were determined using a validated ELISA method of analysis. The results are presented in Table 3. As shown in the Table, the mean PAT protein levels were highest in the seed (dry-weight and fresh-weight basis) and lowest in the pollen and root (fresh-weight only).



**Table 3: Summary of PAT Protein Levels in Tissues from MON88701 cotton - U.S. Field Trials conducted in 2010**

Tissue <sup>1</sup>	Development Stage <sup>2</sup>	Days After Planting (DAP)	PAT ( <i>bar</i> ) Mean (SD) Range (µg/g fw) <sup>3</sup>	PAT ( <i>bar</i> ) Mean (SD) Range (µg/g dw) <sup>4</sup>	LOD/LOQ <sup>5</sup> (µg/g fw)
OSL-1	2-4 leaf	14-25	0.84 (0.21) 0.46 – 1.4	5.5 (1.5) 3.7 – 9.1	0.162/0.188
OSL-2	4-7 leaf	25-37	1.1 (0.26) 0.68 – 1.6	6.4 (1.4) 3.8 – 9.4	0.162/0.188
OSL-3	9 leaf – Full flower	35-99	1.0 (0.34) 0.34 – 1.7	4.8 (2.0) 1.3 – 10	0.162/0.188
OSL-4	Full flower – Cutout	70-121	0.78 (0.29) 0.42 – 1.7	3.2 (1.2) 2.0 – 6.7	0.162/0.188
Root	50% open flower – Full flower	62-99	0.56 (0.18) 0.27 – 0.89	1.8 (0.75) 0.93 – 3.3	0.096/0.188
Pollen	50% open flower – Full flower	68-99	0.56 (0.24) 0.27 – 0.90	NA (NA) NA	0.021/0.188
Seed	Maturity	148-183	6.1 (0.95) 4.8 – 8.8	6.6 (1.1) 5.2 – 9.6	0.032/0.188

<sup>1</sup>OSL=over-season leaf. Seed=black seed (ginned and delinted)

<sup>2</sup>Stage of development of crop when tissue collected.

<sup>3</sup>Protein levels are expressed in microgram of protein per gram of tissue on a fresh weight (fw) basis. For each tissue across all sites n=32, except OSL-1 n=28, OSL-3 n=31, OSL-4 n=27. For pollen samples n=6 because 26 samples expressed below LOQ.

<sup>4</sup>Protein levels expressed on a dry weight (dw) basis.

<sup>5</sup>LOQ=limit of quantitation; LOD=limit of detection

#### 4.3.2.1 Comparison of PAT expression in other GM cotton lines

A comparison of the levels of PAT expression in two separate and independently developed GM cotton lines indicates that the levels of PAT protein are similar in the seeds and pollen. The PAT protein is expressed from the *pat* gene (*S. viridochromogenes*) in herbicide-tolerant and insect-protected cotton line GHB119 (FSANZ Application A1073). While detected in all GHB119 plant parts, PAT protein measurements were lowest in seed (2.4 µg/g fw) and pollen (0.5 µg/g fw).

#### 4.3.2.2 Dietary exposure to PAT

As seen in Table 3, the mean level of PAT in MON88701 cottonseed was 6.6 µg/g dw. Based on the percentage of total protein in MON88701 cottonseed being 28%, this level represents 0.002% of total protein (20 ppm) in MON88701 cottonseed. Dietary exposure to the PAT protein is therefore likely to be negligible from consumption of RBD oil or products containing cellulose derived from MON88701 cotton.

### **4.3.3 Conclusion on identity and function of newly expressed proteins**

The studies described above confirmed the identity and expected functional properties of the newly expressed DMO and PAT proteins in MON88701 cotton, and determined the equivalence of the proteins produced in MON88701 and in a bacterial expression system using *E. coli*. Neither PAT nor DMO are post-translationally modified in MON88701. Both proteins are expressed in various plant tissues. Compared with leaves at various stages of development, roots and pollen, the seeds of MON88701 contain the lowest levels of the DMO protein and the highest levels of the PAT protein.

## **4.4 Potential toxicity of the newly expressed proteins**

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

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

### **4.4.1 History of human consumption**

#### **4.4.1.1 DMO**

The organism from which the *dmo* gene was isolated, *S. maltophilia*, occurs ubiquitously in the environment and is found on plant species many of which are commonly eaten (Ryan *et al.*, 2009). Isolates have been identified in sources such as ready-to-eat salads (Qureshi *et al.*, 2005), fish such as yellowtail and tuna, as well as drinking water dispensers (Sacchetti *et al.*, 2009) and milk (Cairo *et al.*, 2008).

The DMO protein is classified as an oxygenase; enzymes in this class are widely distributed in many universal metabolic pathways. Non-heme iron oxygenases, where iron is involved in the catalytic site, are an important class, and within this class are Rieske non-heme iron oxygenases, which contain a Rieske iron-sulphur [2Fe-2S] cluster. The DMO protein expressed in MON88701 is an example of Rieske type non-heme iron oxygenases, which are found in diverse organisms ranging from bacteria to plants consumed by humans and other animals. Humans therefore have extensive prior dietary exposure to oxygenases of this class, and likely exposure to the DMO protein expressed in MON88701 through indirect ingestion of the bacterial source without adverse effects.

#### **4.4.1.2 PAT**

*Streptomyces hygroscopicus* (*bar*) and *S. viridochromogenes* (*pat*) are common soil bacteria, therefore humans have a long history of exposure to the PAT protein through the consumption of plant roots and vegetables. In addition, since 1995, humans have also been directly exposed to the PAT protein through the consumption of a variety of foods derived

from previously approved GM soybean, cotton, corn and canola crops tolerant to glufosinate ammonium. There is no evidence of toxicity associated with the PAT protein as a result of dietary exposure from any of these sources.

#### **4.4.2 Similarities with known protein toxins**

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

- H. Tu & A. Silvanovich (2012). Bioinformatics Evaluation of the DMO protein in MON88701 Utilizing the AD\_2011, TOX\_2011 and PRT\_2011 Databases. Monsanto Study MSL0023516 (unpublished).
- H. Tu & A. Silvanovich (2011). Bioinformatics Evaluation of the PAT (*bar*) Protein in MON88701 Utilizing the AD\_2011, TOX\_2011 and PRT\_2011 Databases. Monsanto Study MSL0023528 (unpublished).
- C. Herouet-Guicheney (2006). Phosphinothricin Acetyltransferase (PAT) Protein *bar* Gene Product: Overall Amino Acid Sequence Homology Search with Known Toxins and Allergens. Study Report SA 06001, Bayer CropScience (unpublished)<sup>2</sup>

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

##### **4.4.2.1 DMO**

The DMO protein sequence expressed in MON88701 (that is, with addition of nine amino acids from *CTP2*), was compared to allergen (AD\_2011), toxin (TOX\_2011) and public domain (PRT\_2011) sequence databases using the FASTA sequence comparison algorithm. The FASTA program directly compares amino acid sequences (primary protein structure) and the alignment data may be used to infer shared higher order structural similarities between two protein sequences (secondary or tertiary protein structure). Proteins that share a high degree of similarity throughout the entire sequence are often homologous, and often have common secondary structures, common three-dimensional configuration and consequently, may share similar functions.

The TOX\_2011 database contained 10,570 protein sequences and was a selected sub-set of sequences derived from the PRT\_2011 database. Examination of the bioinformatics results showed no biologically relevant sequence similarities between the MON88701 DMO and known toxins or proteins with anti-nutritional properties.

##### **4.4.2.2 PAT**

The Applicant compared the sequence of the PAT protein expressed from the *bar* gene in MON88701 to allergen (AD\_2011), toxin (TOX\_2011) and public domain (PRT\_2011) sequence databases using the same FASTA sequence comparison algorithm. The results indicated no biologically relevant similarities to known toxins or other proteins of concern.

In previous assessments of glufosinate ammonium tolerant crops, the PAT (*bar*) protein sequence was compared with protein sequences present in a number of large public reference databases: eg Uniprot\_Swissprot, Uniprot\_TrEMBL, PDB (Protein Data Bank), DAD (DNA Data Bank of Japan Aminoacid Database) and GenPept. The results of the overall homology search with the PAT protein showed no similarity with known toxins; similarity only with other acetyltransferase proteins was revealed however none of these proteins are associated with any toxicity.

---

<sup>2</sup> This study was submitted in Application A1028 and considered again in support of Application A1040.

#### 4.4.3 Stability to thermal treatment

Studies on the thermolability of a protein provide an indication of the stability of the protein under cooking/processing conditions. The Applicant conducted separate studies on DMO and PAT to investigate the effect of heat on the structure and function of both proteins.

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

R. Hernan, R. Frazier & E. Bell (2011). Effect of Heat Treatment on the Functional Activity of *E. coli*-Produced MON88701 DMO Protein. Monsanto Company MSL0023606 (unpublished)

K.J. Mathis and E. Bell (2011). The Effect of Heat Treatment on the Functional Activity of *E. coli*-Produced Phosphinothricin N-acetyltransferase (*bar*) Protein. Monsanto Company MSL0023584 (unpublished)

##### 4.4.3.1 DMO

Aliquots of the DMO protein produced in *E. coli* were heated to 25, 37, 55, 75 and 95 degrees (°C) for either 15 or 30 minutes. The integrity of the protein was analysed by gel electrophoresis (SDS-PAGE) and by a dicamba mono-oxygenase activity assay.

The activity of the reference sample was measured as 8.23 nmol DCSA/minute/mg DMO protein. The control sample stored on ice had activity of 6.74 nmol DCSA/minute/mg DMO protein. There was no change in the activity of the bacterially-produced DMO protein when heated at 25°C and 37°C for incubation times of 15 and 30 minutes. The enzyme activity levels following incubation at 55°C, 75°C and 95°C were below the limit of quantification.

No significant changes to band intensity were observed on SDS-PAGE following incubation at all temperatures for 15 minutes or 30 minutes. These data indicate that the functional activity of DMO is diminished with denaturation of the protein at temperatures higher than 37°C.

##### 4.4.3.2 PAT

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

As for the experiments with the DMO protein, aliquots of the PAT protein produced in *E. coli* were incubated at 25, 37, 55, 75 or 95 °C for either 15 or 30 minutes. The integrity of the protein was analysed qualitatively by SDS-PAGE and quantitatively by a PAT enzyme activity assay. The specific activities of the test and reference substances were calculated using the molar absorptivity of product released during the assay and expressed as µmol/minute/mg of PAT (*bar*).

The activity of the reference sample was measured as 28.1 µmol/minute/mg of PAT (*bar*). The control sample stored on ice had activity of 27.2 µmol/minute/mg of PAT (*bar*). There was a modest reduction in PAT activity following incubation at 25°C and 37°C for 15 or 30 minutes, with ≥79% activity remaining relative to the control sample. At 55°C, a significant reduction in PAT enzyme activity was observed, with 40% remaining after 15 minutes incubation and 24% remaining after 30 minutes incubation. At the higher temperatures of 75°C and 95°C, >90% of PAT enzyme activity was lost after 15 and 30 minutes incubation.

The SDS-PAGE indicated no significant decrease in band intensity of the *E. coli*-produced PAT protein when heated at all temperatures for either 15 or 30 minutes. The results of the heat stability experiments indicate that the functional activity of the PAT protein is gradually diminished with increasing temperature, with approximately three-quarters of the enzyme activity lost after heating for 30 minutes at temperatures higher than 37°C. These data are consistent with similar experiments with the PAT (*bar*) protein considered in support of Application A1028<sup>3</sup>.

#### 4.4.4 Acute oral toxicity studies

**Studies submitted:**

J.W. Smedley (2012). An Acute Toxicity Study of *E. coli*-produced MON88701 DMO Administered by the Oral Gavage Route to Mice. CRO-2011-035. (Monsanto Company, unpublished)

J.W. Smedley (2012). An Acute Toxicity Study of *E. coli*-produced Phosphinothricin N-acetyltransferase (PAT[*bar*]) Protein Administered by the Oral Gavage Route to Mice. CRO-2011-007. (Monsanto Company, unpublished)

Separate acute oral toxicity studies in mice using respective *E. coli*-produced test substances DMO and PAT were submitted by the Applicant but are not included in this safety assessment. Such studies are only deemed necessary if the results of the biochemical, bioinformatic, digestibility or stability studies indicate further investigation of potential toxicity is warranted. In this case, on the basis of the data assembled from the other studies, and on previous conclusions from several safety assessments of the PAT (*bar*) and DMO proteins, no safety concerns were identified.

#### 4.5 Potential allergenicity of the newly expressed proteins

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

The allergenic potential of DMO and PAT proteins was assessed by consideration of:

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

<sup>3</sup> D.J. Esdaile (2002). Phosphinothricin Acetyltransferase (PAT) Bar Gene Product Heat Stability Study. Study SA 02175, Bayer CropScience (unpublished).

#### 4.6.1 Source of each protein

As described in Section 4.2, neither the DMO nor PAT proteins are derived from bacterial sources associated with allergenicity in humans.

#### 4.6.2 Similarity to known allergens

**Studies submitted:**

H. Tu & A. Silvanovich (2012). Bioinformatics Evaluation of the DMO Protein in MON88701 Utilizing the AD\_2011, TOX\_2011 and PRT\_2011 Databases. Monsanto Company MSL0023516 (unpublished).

H. Tu & A. Silvanovich (2011). Bioinformatics Evaluation of the PAT (*bar*) Protein in MON88701 Utilizing the AD\_2011, TOX\_2011 and PRT\_2011 Databases. Monsanto Study MSL0023528 (unpublished).

Bioinformatics contributes to the weight of evidence approach for assessing potential allergenicity of novel proteins introduced to GM plants (Goodman, 2006; Thomas *et al.*, 2005). As with the bioinformatic analysis that looked at similarities with known protein toxins (refer to Section 4.5.2), this analysis compares the amino acid sequence of the newly expressed proteins with sequences of known allergens in order to identify primary structural similarities that may indicate a potential for cross-reactivity with allergenic proteins. The Codex guideline for the evaluation of potential allergenicity of newly expressed proteins indicates the possibility for cross-reactivity if the newly expressed protein is found to have at least 35% amino acid identity with an allergen over any segment of at least 80 amino acids (Codex, 2009). A sliding window search of eight linear contiguous amino acids is also used to identify short peptides of possible immunological relevance in otherwise unrelated proteins. The FASTA sequence alignment tool was used for this purpose.

The sequence of the 'mature' DMO protein (including nine amino acids remaining from the cleavage of the CTP and the additional leucine at position two) was analysed against the allergen, gliadin and glutenin sequence database (AD\_2011) obtained from the Food Allergy Research and Resource Program Database (FARRP, 2011). The AD\_2011 database contains 1491 sequences. No alignment of 35% identity over 80 amino acids was present, indicating that the DMO protein in MON88701 does not share meaningful similarity with sequences in the allergen database. In addition, there were no matches of eight contiguous amino acids detected in the comparison of DMO protein sequence to proteins in the allergen database. Taken together, these data show that the DMO protein lacks both structurally and immunologically relevant similarities to known allergens, gliadins and glutenins.

Similarly, for the PAT protein, no alignments with known allergens were identified and no short (eight amino acids) peptide matches were shared between the PAT (*bar*) protein sequence and proteins in the allergen database. These data are consistent with previous assessments of the PAT protein indicating a lack of potential allergenicity in humans.

#### 4.6.3 In vitro digestibility

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

*In vitro* digestibility of the DMO and PAT proteins was assessed in assays using simulated gastric fluid (SGF) containing pepsin, and simulated intestinal fluid (SIF) containing pancreatin (an enzyme mixture). The SGF assay protocol has been standardised based on results obtained from an international, multi-laboratory ring study (Thomas *et al.* 2004). This study showed that the results of *in vitro* pepsin digestion assays were reproducible when a standard protocol was followed utilising a physiologically relevant acidic pH to simulate conditions in the stomach. The SIF study is considered less relevant because an ingested protein would first be exposed to pepsin-mediated hydrolysis in the acidic environment of the stomach before being exposed to further digestion in the small intestine.

**Studies submitted:**

B. Chen, K.S Crowley & E. Bell (2011). Assessment of the *in vitro* Digestibility of *Escherichia coli* (*E.coli*)-produced MON88701 Decamba Mono-oxygenase (DMO) Protein in Simulated Gastric and Simulated Intestinal Fluids. Monsanto Company MSL0023579 (unpublished)

R. Wang, K.S Crowley, R. Hernan & E. Bell (2011). Assessment of the *in vitro* Digestibility of Phosphinothricin N-Acetyltransferase (*bar*) Protein in Simulated Gastric and Simulated Intestinal Fluids. Monsanto Company MSL0023567 (unpublished)

#### 4.6.3.1 DMO digestibility studies

The digestibility of the *E. coli*-produced DMO protein was determined by analysing digestion mixtures incubated with SGF at 37°C for specific time intervals ranging from 30 seconds to 60 minutes. The products were analysed visually by SDS-PAGE and by Western blot probed with anti-DMO polyclonal antibody.

Visual examination of the gel showed that the intact DMO protein was completely digested within 30 seconds of incubation in SGF. The presence of the intact DMO protein band in the sample without pepsin indicates the protein was stable to the acid pH of the assay conditions and that the observed degradation over the time course was due to the activity of pepsin.

The results of the Western blot confirmed the observations from the protein gel. The DMO protein was digested to below the LOD (estimated at 0.2 ng) within 30 seconds of incubation in SGF. No immunoreactive bands were visible at other timepoints, and no other fragments were observed. Based on calculations for the LOD and the amount of DMO protein loaded onto the gel, it was concluded that more than 99% of the DMO protein was digested within 30 seconds of exposure to SGF.

Digestibility of the DMO protein in SIF over various timepoints between 5 minutes and 24 hours was analysed by Western blot. The DMO protein was digested to a level below the LOD within five minutes of incubation in SIF, the first timepoint measured. Taken together, these studies demonstrate that the DMO protein is readily susceptible to complete digestion in a simulated gastrointestinal system. These data also support the conclusion that if ingested as a component of the diet, it is unlikely that intact DMO protein would be in contact with the intestinal mucosa.

#### 4.6.3.2 PAT digestibility studies

The *in vitro* digestibility of the *E. coli*-produced PAT protein was determined by analysing digestion mixtures incubated with SGF at 37°C for specified time intervals: 30 seconds, 2, 5, 10, 20, 30 and 60 minutes. The products were analysed visually by SDS-PAGE and by Western blot probed with anti-PAT polyclonal antibody. The LOD of intact PAT protein on the SDS gels was determined at a loading of 13 ng protein. On Western blot, the LOD of the intact PAT protein was observed at a loading of 0.16 ng protein.

Both the SDS gels and Western blotting indicated that the PAT protein was rapidly hydrolysed in SGF, with more than 98% digestion within 30 seconds of incubation in SGF. In addition, no peptide fragments were observed at any time point, indicating complete digestion of the protein at this first measurement point.

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

Digestibility of the PAT protein incubated at 37°C in SIF over a series of time points (5, 15, 30, 60 minutes and 2, 4, 8 and 24 hours) was analysed by Western blot. The PAT protein was hydrolysed to a level below the LOD within five minutes of incubation in SIF. Based on the calculated LOD, more than 98% of the intact PAT protein was digested after five minutes exposure to digestive enzymes. A faint immunoreactive band corresponding to a fragment less than 10kDa was observed at the five and 15 minute time points in SIF, however this was no longer visible by the 30 minute time point. These results are confirmation of previous *in vitro* digestibility experiments from which it was concluded that PAT would be susceptible to rapid digestion in the human gastrointestinal tract.

#### 4.6 Herbicide metabolites

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

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

Dicamba mono-oxygenase (DMO) is an enzyme that catalyzes the demethylation of dicamba to the non-herbicidal compound 3,6-dichlorosalicylic acid (DCSA) and formaldehyde, allowing a GM plant to function in the presence of the herbicide. In herbicide-tolerant crops such as MON88701, residue data are needed to confirm the concentration of GM trait-specific metabolites, relative to the parent herbicide and its major metabolites on conventional crops (see Section 4.6.1).

Dicamba is the ISO common name for 3,6-dichloro-2-methoxybenzoic acid (IUPAC). In Australia, dicamba has been registered for use on cereal crops such as wheat, and sugarcane by the Australian Pesticides and Veterinary Medicines Authority (APVMA)<sup>4</sup>. The MRL for use of dicamba on cereal grains is 0.05 ppm and for sugarcane is 0.1 ppm, however dicamba is not currently registered for use in or on cotton, and consequently there is no corresponding MRL for cotton commodities. For imported commodities only, Australia and New Zealand can recognise residue limits set internationally by Codex or in certain cases adopt the limits set by the country in which the commodity is grown, subject to dietary exposure assessments for Australia and New Zealand.

---

<sup>4</sup> Australia New Zealand Food Standards Code, Standard 1.4.2 Maximum Residue Limits, Schedule 1



#### 4.6.1 Residues of dicamba and metabolites in MON88701 cottonseed

##### Study submitted:

D. L. Maher (2012). Summary of the Magnitude of Residues of Dicamba in Undelinted Cotton Seed after Application to MON88701. Monsanto Company MSL0024066 (unpublished).

A field study was conducted at 13 sites across the United States to determine the residue levels of dicamba and its metabolites in MON88701 raw agricultural commodities (undelinted cottonseed) and processed commodities, following application of dicamba formulations. The sites were typical of the major cotton producing regions of that country. The formulation used in the treatments was MON 54140, which is a formulation containing the diglycolamine salt of dicamba. The maximum labelled rate and timing for applications of dicamba to MON88701 cotton were varied for different treatments, however the data presented were from the treatment options expected to provide the maximum residues under the proposed label rates. As indicated in Tables 4 and 5, these were treatments 3 and 4 (data from treatment 2 are not relevant for this assessment and are not presented).

Residue analysis quantified DCSA (3,6-dichloro-2-hydroxybenzoic acid) and DCGA (2,5-dichloro-3,6-dihydroxybenzoic acid), along with dicamba and 5-hydroxydicamba using liquid chromatography-tandem mass spectrometry (LC-MS/MS). The analytical method was validated down to a lower limit of 0.02 ppm for dicamba, 5-hydroxydicamba, DCSA and DCGA in cotton undelinted seed for method validation.

**Table 4: Applications of Dicamba to MON88701 cotton in United States field trials (2010)**

Treatment	Application Rate lb/a (kg/ha)				
	Pre-Emergence	6 Leaf	1st White Flower + 15 Days	First Open Boll	7 Days Preharvest
1	—	—	—	—	—
3	1.0 lb/a (1.12 kg/ha)	—	—	0.50 lb/a (0.56 kg/ha)	0.50 lb/a (0.56 kg/ha)
4	—	0.50 lb/a (0.56kg/ha)	0.50 lb/a (0.56 kg/ha)	0.50 lb/a (0.56 kg/ha)	0.50 lb/a (0.56 kg/ha)

The results for the median and range of residues in seed for each analyte and the total residue for cottonseed for dicamba and metabolites in MON88701, are presented in Table 5. Total residues are expressed as dicamba acid equivalents, and include the analytes in the proposed residue definition: dicamba, 5-hydroxydicamba and DCSA<sup>5</sup>.

<sup>5</sup> The two analytes that constitute the existing U.S. EPA Definition of the Residue for cotton seed are dicamba and 5-hydroxycamba, as per 40 CFR §180.227.

**Table 5: Summary of Dicamba Residues in MON88701 cotton seed**

Analyte	Treatment 3, ppm			Treatment 4, ppm		
	Mean	Median <sup>a</sup>	Range <sup>b</sup>	Mean	Median <sup>a</sup>	Range <sup>b</sup>
DCGA	0.03	0.02	<0.02-0.09	0.05	0.03	0.02-0.14
DCSA	0.05	0.03	<0.02-0.16	0.08	0.06	0.02-0.27
Dicamba	0.64	0.65	0.06-1.38	0.61	0.47	0.12-1.42
5-Hydroxy-dicamba	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02
Total <sup>c,d</sup>	0.72	0.71	0.10-1.57	0.71	0.54	0.17-1.72

<sup>a</sup> Median of residues across all sites in this treatment.

<sup>b</sup> Range of residues (minimum/maximum) across all sites.

<sup>c</sup> Individual analyte residues are expressed as each analyte *per se*.

<sup>d</sup> Total residues are expressed as dicamba acid equivalents. Total residue (ppm) = [Dicamba] + ([5-hydroxydicamba] x 0.933) + ([DCSA] x 1.068)

Cottonseed from treated and control (untreated) MON88701 plants was obtained from Treatment 4 plots and processed using small-scale commercial-type equipment. The residues in the seed meal were compared to the residues in the undelinted cottonseed prior to processing to determine a concentration factor. The results are presented in Table 6. In terms of use of cottonseed meal in animal feed, the conclusion from these analyses was that dicamba residues in dicamba-tolerant cottonseed and meal should not result in any impact on the current livestock dietary burden for dicamba in the United States.

**Table 6: Summary of Cotton Undelinted Seed Concentration Factors in Treatment 4 Processed Fractions from MON88701**

Process Fraction	Analyte/Residues (ppm) <sup>1</sup>				Total <sup>2,3</sup>	Conc. Factor <sup>4</sup>
	DCGA	DCSA	Dicamba	5-Hydroxy-dicamba		
Undelinted Seed	0.04	0.04	0.65	<0.02	0.71	-
Meal	<0.02	<0.02	0.11	<0.02	0.15	0.22

<sup>1</sup> Average of individual residues in undelinted cottonseed processed fraction sample analyses across two sites.

<sup>2</sup> Total Residue (ppm) = [Dicamba] + ([5-hydroxydicamba] x 0.933) + ([DCSA] x 1.068) as per proposed residue definition.

<sup>3</sup> Total residues are expressed as dicamba equivalents. Individual analytes are expressed as each analyte *per se*. Note: The lower limit of method validation of dicamba, 5-hydroxydicamba, DCSA and DCGA in cotton processed fractions is 0.02 ppm. Values below 0.02 are shown as <0.02 and included in the totals as 0.02.

<sup>4</sup> Concentration factor based on total residues calculated as proposed definition of residue.

#### 4.6.2 Safety of DCSA

The metabolism of dicamba in plants has been investigated in cereals (wheat, sugarcane) and in pulse/oilseed plants such as soybean and cotton. In plants, dicamba is first metabolised to 5-hydroxydicamba (hydroxylation step), or to DCSA via a demethylation step. Both of these metabolites are further degraded to DCGA. In the context of its use on conventional (non-GM) crops, dicamba and the major metabolite 5-hydroxydicamba and DCSA have been formally evaluated by agencies such as the US Environmental Protection Agency (EPA, 2009) and by the European Food Safety Authority (EFSA, 2011). DCSA is a known metabolite on soybean<sup>6</sup> and in livestock. DCSA is also the primary degradate in soil from aerobic metabolism of dicamba and is therefore not new to the environment; it is not persistent in the environment and has a low potential for leaching to ground water (EFSA, 2011). Based on a number of toxicity studies using Rainbow trout (*Oncorhynchus mykiss*), *Daphnia*, green algae (*Selenastrum caprinornutum*), *Lemna* and earthworm (*Eisenia fetida*), it was concluded by EFSA that DCSA “was not found to give rise to unacceptable risks”.

#### 4.6.3 Minor substances

Formaldehyde, as the other reaction product, is ubiquitous in the environment and atmosphere from biogenic sources such as plants and animals, and from industrial sources. It is rapidly degraded and any environmental effects of formaldehyde resulting from dicamba treatment of crop plants are considered to be negligible.

#### 4.6.4 Glufosinate residues

The enzyme activity of phosphinothricin N-acetyltransferase (PAT) results in the acetylation of the free amino group of glufosinate to produce the non-herbicidal N-acetyl glufosinate. This is a well known metabolite in glufosinate ammonium tolerant plants and has been previously considered in other glufosinate-tolerant crops<sup>7</sup>. In Australia, the import MRL (FSANZ) and the MRL for glufosinate in cottonseed (APVMA) is 3 ppm. The use pattern and rate of glufosinate application on MON88701 will follow the existing glufosinate-tolerant cotton uses outlined on the label, and residues are below the established pesticide residue tolerances for cottonseed. Additional data on the identity and levels of herbicide and any metabolites are therefore not given in this assessment.

### 4.7 Conclusion

The two newly expressed proteins in MON88701 cotton are DMO (conferring tolerance to dicamba) and PAT (conferring tolerance to glufosinate ammonium herbicides). Both proteins are expressed in mature cottonseeds at low levels, and trace amounts are detected in pollen.

The identity and physicochemical and functional properties of the DMO and PAT proteins were examined in detail to confirm their expression as expected in the plants. Separate investigations were conducted to determine their potential toxicity and allergenicity. The characterisation studies confirmed that the DMO and PAT proteins produced in MON88701 cotton conform in size and amino acid sequence to that expected, with the exception of an additional nine amino acids at the N-terminus of DMO, do not exhibit any post-translational modification including glycosylation and exhibit the expected enzyme activity.

---

<sup>6</sup> The metabolism of dicamba in herbicide-tolerant soybean MON87708 was discussed in the safety assessment of Application A1063

<http://www.foodstandards.gov.au/code/applications/Pages/applicationa1063food5198.aspx>

<sup>7</sup> See Application A533: Food derived from Glufosinate ammonium-tolerant Cotton Line LL25.

The DMO and PAT proteins have been investigated previously for potential toxicity and allergenicity and both proteins were found to be innocuous. Based on testing for potential toxicity and allergenicity repeated for this application, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens and *in vitro* digestibility studies demonstrated that both DMO and PAT proteins were rapidly degraded in simulated gastric and intestinal fluids. Both enzymes were rapidly inactivated by heating. Taken together, the evidence clearly supports the conclusion that DMO and PAT are not toxic nor likely to be allergenic in humans, and would be digested like other dietary proteins.

## **5. Compositional analysis**

The purpose of compositional analysis is to determine if any unexpected changes in composition have occurred to the food and to establish its nutritional adequacy. Compositional analysis can also be important for evaluating the intended effect where the genetic modification has resulted in a deliberate change to one or more nutrients in the food. In this case, MON88701 is herbicide tolerant and there was no intention to alter the nutrient composition of food derived from this plant line.

The focus of the compositional analysis is on those constituents most relevant to the safety of the food, or that may have an impact on the whole diet. Important analytes therefore include the key nutrients, toxicants and anti-nutrients relevant for the food in question. The key nutrients and anti-nutrients are those components in a particular food that may have a substantial impact in the overall diet. They may be major constituents (fats, proteins, carbohydrates or enzyme inhibitors/anti-nutrients) or quantitatively more minor constituents (minerals, vitamins). Key toxicants are those that have a level of toxicity and occur in amounts that may be significant to health (eg solanine in potatoes).

### **5.1 Key components**

Cottonseed oil is the primary cotton product used for human consumption. The key components to be analysed for the comparison of transgenic and conventional cotton are outlined in the OECD Consensus Document on Compositional Considerations for New Varieties of Cotton, and include proximates (cottonseed only), fatty acids, tocopherol, gossypol and the cyclopropenoid fatty acids - malvalic, sterculic and dihydrosterculic acids (OECD, 2004).

### **5.2 Study design and conduct for key components**

#### **Studies submitted:**

D. Howard, K.D. Miller & R. Sorbet (2012). Amended Report for MSL0024393: Compositional Analyses of Cottonseed Collected from MON88701 treated with Dicamba and Glufosinate Grown in the United States during 2010. Monsanto Company MSL0024606 (unpublished). Performing laboratories included Covance Laboratories Inc., Wisconsin, and Certus International Inc., Missouri.

The near isogenic conventional (non-GM) line Coker 130 was used for the direct comparison with MON88701 cotton. In addition, the Applicant included nine other unique conventional cotton varieties in this study, to provide data on natural variability for each compositional component analysed.

The field production was conducted at eight sites across agricultural areas in the United States: Arkansas, Georgia, Kansas, Louisiana, North Carolina, New Mexico, South Carolina and Texas. In addition to the transgenic and parental lines, four of the nine conventional reference varieties were grown at the sites.

All cotton plants including MON88701, conventional control and reference varieties were grown under normal agronomic field conditions, including maintenance pesticides as needed for their respective geographic region. The MON88701 plots were treated at the 3-5 leaf stage with glufosinate herbicide at the label rate (0.5 pound a.i. per acre) and at the 6-10 leaf stage with dicamba herbicide at the proposed label rate (0.5 pound a.i. per acre).

Cottonseed samples were harvested and ginned from all plots, and despatched to analytical laboratories under full identity labelling. Compositional analyses were carried out on acid-delinted cottonseed. Nutrients analysed were proximates (ash, carbohydrates by calculation, moisture, protein and total fat), acid detergent fibre (ADF), neutral detergent fibre (NDF), crude fibre (CF), total dietary fibre (TDF), amino acids, fatty acids (C8-C22), minerals (calcium, copper, iron, magnesium, manganese, phosphorus, potassium, sodium and zinc) and vitamin E (tocopherol). The antinutrients included in this analysis were gossypol and cyclopropenoid fatty acids (dihydrosterculic, malvalic and sterculic acids). Methods of compositional analysis were based on internationally recognised procedures (e.g. those of the Association of Official Analytical Chemists).

A total of 65 different analytical components were measured. All components were statistically analysed using a mixed-model analysis of variance with the Statistical Analysis Software<sup>8</sup> (SAS) MIXED procedure. Data from the eight replicated sites were analysed both separately and combined. In order to complete the statistical analysis for any component in this study, it was deemed that at least 50% of the values must be greater than the assay LOQ. If analytes had more than 50% of observations below the LOQ for that assay, they were excluded from the overall summary analysis. Using this criterion, the following 13 fatty acid analytes were excluded from the statistical summary: 8:0 caprylic acid, 10:0 capric acid, 12:0 lauric acid, 14:1 myristoleic acid, 15:0 pentadecanoic acid, 15:1 pentadecenoic acid, 17:0 heptadecanoic acid, 17:1 heptadecenoic acid, 18:3 gamma-linolenic acid, 20:1 eicosenoic acid, 20:2 eicosadienoic acid, 20:3 eicosatrienoic acid and 20:4 arachidonic acid. Values for all components were expressed on a dry weight basis with the exception of moisture, expressed as percent fresh weight, and fatty acids, expressed as percent of total fatty acids.

Nine sets of statistical comparisons of MON88701 cotton to the conventional control Coker 130 were conducted. One comparison was based on compositional data combined across all eight field sites (combined-site analysis) and eight separate comparisons were conducted on data from each individual field site. Compositional data from non-GM commercial varieties grown concurrently in the same trial as MON88701 and the control, were combined across all sites and used to calculate a 99% tolerance interval for each component to define the natural variability in cottonseed varieties that have a history of safe consumption. Any statistically significant differences between MON88701 and the control Coker 130 were also compared to this tolerance range, to assess whether the differences were likely to be biologically meaningful.

### **5.3 Analyses of key components**

The results and discussion of the compositional analyses are presented below for key constituents: proximates, fibre, amino acids, fatty acids, tocopherols and cyclopropenoid fatty acids.

#### **5.3.1 Proximates**

Results for the proximates analyses are shown in Table 7.

---

<sup>8</sup> SAS website - <http://www.sas.com/technologies/analytics/statistics/stat/index.html>

Statistically significant differences between MON88701 and Coker 130 were noted for all analytes except protein, but in all cases the magnitude of the differences was small (0.6% to 5.0%), and the mean levels in MON88701 were within the range for commercial non-GM cotton varieties.

**Table 7: Statistical Summary of Proximates in MON88701 cottonseed, Coker 130 and commercial cultivars – Combined-site Analysis**

Analytical Component (Units) <sup>1</sup>	MON 88701 <sup>2</sup> Mean (S.E.) <sup>3</sup> (Range)	Control <sup>4</sup> Mean (S.E.) (Range)	Difference (MON 88701 minus Control)			Commercial Tolerance Interval <sup>5</sup> (Range)
			Mean (S.E.) (Range)	95% Confidence Interval	Significance (p-Value)	
Proximate (% dw)						
Ash	4.31 (0.11) (3.77 - 4.74)	4.11 (0.11) (3.34 - 5.00)	0.21 (0.052) (-0.49 - 0.61)	0.094, 0.32	0.001	3.42, 4.65 (3.18 - 4.68)
Calories (Kcal/ 100g)	498.50 (1.65) (482.46 - 517.46)	495.24 (1.71) (487.70 - 512.65)	3.26 (1.29) (-14.30 - 18.37)	0.70, 5.82	0.013	457.61, 527.56 (466.09 - 509.91)
Carbohydrates	44.64 (0.56) (41.40 - 48.89)	45.83 (0.57) (42.14 - 50.30)	-1.19 (0.32) (-5.19 - 2.45)	-1.82, -0.56	<0.001	40.26, 56.45 (43.28 - 54.90)
Moisture (% fw)	7.15 (0.26) (5.93 - 9.67)	7.48 (0.27) (6.15 - 9.19)	-0.34 (0.11) (-1.82 - 0.79)	-0.56, -0.11	0.005	4.79, 9.92 (6.05 - 10.50)
Protein	27.91 (0.77) (22.71 - 31.47)	27.79 (0.77) (23.53 - 31.27)	0.13 (0.31) (-1.99 - 3.73)	-0.53, 0.78	0.685	22.30, 29.41 (20.58 - 29.28)
Total Fat	23.14 (0.31) (19.79 - 26.78)	22.31 (0.33) (20.71 - 25.20)	0.83 (0.26) (-2.89 - 3.86)	0.32, 1.34	0.001	15.01, 28.51 (16.58 - 25.25)

<sup>1</sup>dw = dry weight; fw = fresh weight; FA = fatty acid.

<sup>2</sup>MON 88701 was sprayed with dicamba and glufosinate.

<sup>3</sup>Mean (S.E.) = least-square mean (standard error).

<sup>4</sup>Control refers to the non-biotechnology derived, conventional control (Coker 130).

<sup>5</sup>With 95% confidence, interval contains 99% of the values expressed in the population of commercial substances. Negative limits set to zero.

### 5.3.2 Fibre content

Measurement of fibre components is of greater importance for animal feed. As seen from the results in Table 8, several fibre components were statistically different between MON88701 cotton and the control Coker 130, however the values for MON88701 were well within the reference range for commercial non-GM cotton.

**Table 8: Statistical Summary of Fibre and Ash in MON88701 cottonseed, Coker 130 and commercial cultivars – Combined-site Analysis**

Analytical Component (Units) <sup>1</sup>	MON 88701 <sup>2</sup> Mean (S.E.) <sup>3</sup> (Range)	Control <sup>4</sup> Mean (S.E.) (Range)	Difference (MON 88701 minus Control)			Commercial Tolerance Interval <sup>5</sup> (Range)
			Mean (S.E.) (Range)	95% Confidence Interval	Significance (p-Value)	
Fibre (% dw)						
Acid Detergent Fibre	25.27 (0.34) (23.26 - 27.74)	26.58 (0.35) (22.08 - 29.58)	-1.31 (0.35) (-5.42 - 1.77)	-2.06, -0.57	0.002	22.24, 31.96 (23.42 - 31.62)
Crude Fibre	18.17 (0.37) (15.97 - 21.66)	18.54 (0.38) (16.06 - 21.70)	-0.38 (0.32) (-3.36 - 4.75)	-1.02, 0.27	0.246	16.93, 22.68 (16.92 - 23.32)
Neutral Detergent Fibre	30.73 (0.51) (25.13 - 34.42)	32.59 (0.53) (28.87 - 35.89)	-1.86 (0.41) (-6.95 - 1.16)	-2.68, -1.05	<0.001	27.03, 42.49 (29.27 - 40.63)
Total Dietary Fibre	39.44 (0.39) (36.91 - 42.13)	41.12 (0.41) (39.05 - 44.37)	-1.68 (0.36) (-5.34 - 1.09)	-2.45, -0.91	<0.001	34.52, 52.58 (37.29 - 48.60)

<sup>1</sup>dw = dry weight; fw = fresh weight; FA = fatty acid.

<sup>2</sup>MON 88701 was sprayed with dicamba and glufosinate.

<sup>3</sup>Mean (S.E.) = least-square mean (standard error).

<sup>4</sup>Control refers to the non-biotechnology derived, conventional control (Coker 130).

<sup>5</sup>With 95% confidence, interval contains 99% of the values expressed in the population of commercial substances. Negative limits set to zero.



### 5.3.3 Fatty Acids

The major food from cotton plants is cottonseed oil, therefore fatty acid constituents of cottonseed are highly relevant for safety assessment. The levels of 22 fatty acids (C8 – C22) were measured in MON88701 cottonseed and in control Coker 130 seed, however only nine fatty acids were reliably measurable in cottonseed. As previously outlined, if levels were below the LOQ for the assay, the analyte was excluded from further statistical analysis.

As presented in Table 9, the levels of two fatty acids, 14:0 myristic acid and 18:2 linoleic acid, showed statistically significant differences between MON88701 and the control in the combined site analysis. The difference was small in each case: 2.7% decrease in 14:0 myristic acid, and 0.7% decrease for 18:2 linoleic acid. Considering the individual site data, significant differences for fatty acids between MON88701 and the control were not consistently observed at all trial sites, indicating normal biological variation rather than a pattern of change attributable to the genetic background. The mean values for these analytes were within the 99% tolerance interval established by non-GM commercial reference varieties grown concurrently in the same trial. In addition, the mean values for fatty acid components in MON88701 cottonseed were within the context of natural variability of commercial cotton composition as published in the scientific literature or available in the ILSI Crop Composition Database (ILSI, 2011).

Overall, the observed differences do not result in any impact on the nutritional quality of the oil from MON88701, as the values were well within the range for commercial non-GM cotton varieties.

**Table 9: Statistical Summary of major Fatty Acids in MON88701 cottonseed, Coker 130 and commercial cultivars – Combined-site Analysis**

Analytical Component (Units) <sup>2</sup>	MON 88701 <sup>2</sup> Mean (S.E.) <sup>3</sup> (Range)	Control <sup>4</sup> Mean (S.E.) (Range)	Difference (MON 88701 minus Control)			Commercial Tolerance Interval <sup>5</sup> (Range)
			Mean (S.E.) (Range)	95% Confidence Interval	Significance (p-Value)	
<b>Fatty Acid (% Total FA)</b>						
16:1 Palmitoleic	0.50 (0.0094) (0.44 - 0.54)	0.50 (0.0094) (0.45 - 0.54)	0.0022 (0.0038) (-0.025 - 0.039)	-0.0060, 0.010	0.572	0.39, 0.70 (0.44 - 0.67)
18:0 Stearic	2.54 (0.058) (2.29 - 2.85)	2.47 (0.058) (2.15 - 2.76)	0.068 (0.036) (-0.16 - 0.24)	-0.0091, 0.14	0.079	1.98, 2.95 (1.98 - 2.97)
18:1 Oleic	15.10 (0.26) (14.15 - 16.45)	14.96 (0.26) (14.06 - 16.44)	0.14 (0.070) (-0.48 - 0.75)	-0.0049, 0.29	0.057	11.38, 20.64 (13.71 - 18.39)
18:2 Linoleic	55.77 (0.39) (54.24 - 58.22)	56.15 (0.40) (54.04 - 57.93)	-0.39 (0.16) (-1.42 - 0.80)	-0.72, -0.053	0.026	47.49, 63.18 (49.78 - 59.61)
18:3 Linolenic	0.18 (0.022) (0.14 - 0.34)	0.17 (0.022) (0.12 - 0.30)	0.011 (0.0068) (-0.0073 - 0.052)	-0.0038, 0.025	0.136	0.060, 0.24 (0.10 - 0.29)
20:0 Arachidic	0.29 (0.0086) (0.23 - 0.32)	0.28 (0.0087) (0.23 - 0.32)	0.0044 (0.0047) (-0.027 - 0.046)	-0.0057, 0.015	0.364	0.17, 0.38 (0.20 - 0.36)
14:0 Myristic	0.77 (0.030) (0.66 - 0.95)	0.79 (0.031) (0.71 - 0.98)	-0.021 (0.0071) (-0.077 - 0.047)	-0.036, -0.0060	0.009	0.16, 1.37 (0.45 - 1.04)
16:0 Palmitic	23.95 (0.30) (22.34 - 25.28)	23.80 (0.30) (22.69 - 25.05)	0.15 (0.076) (-0.68 - 0.76)	-0.016, 0.31	0.073	16.54, 30.55 (19.11 - 26.73)
22:0 Behenic	0.15 (0.0051) (0.12 - 0.19)	0.15 (0.0051) (0.13 - 0.21)	-0.0035 (0.0029) (-0.049 - 0.032)	-0.0098, 0.0029	0.260	0.070, 0.21 (0.051 - 0.19)

<sup>1</sup>dw = dry weight; fw = fresh weight; FA = fatty acid.

<sup>2</sup>MON 88701 was sprayed with dicamba and glufosinate.

<sup>3</sup>Mean (S.E.) = least-square mean (standard error).

<sup>4</sup>Control refers to the non-biotechnology derived, conventional control (Coker 130).

<sup>5</sup>With 95% confidence, interval contains 99% of the values expressed in the population of commercial substances. Negative limits set to zero.

#### **5.3.4 Amino Acids**

Levels of 18 amino acids in acid-delinted cottonseed from MON88701 and the control Coker 130 were measured for comparison (see Table 10). Of these, arginine, methionine and proline showed statistically significant differences between the GM and non-GM lines in the combined-site analysis. For both arginine and proline, the magnitude of the differences between the mean values for MON88701 and the control were small decreases, 3.8% and 2.6% respectively. The mean values for MON88701 were within the 99% tolerance interval established by reference to conventional commercial cotton varieties grown concurrently in the same trial.

The mean level of methionine was increased 4.8% in MON88701 compared with the control. The observed difference in methionine in the combined-site analysis is not a safety concern because the mean levels were within the context of natural variation in methionine levels found in commercial cotton varieties as published in the scientific literature or as recorded in the ILSI Crop Composition Database (ILSI, 2011). As shown in the proximate analysis (Table 7), the levels of protein in MON88701 and the control Coker 130 are essentially the same. In addition, as indicated by the individual site data, small fluctuations (slight increases and decreases) in the levels of several amino acids commonly occur and are attributable to natural biological variation.



**Table 10: Statistical Summary of Amino Acids in MON88701 cottonseed, Coker 130 and commercial cultivars – Combined-site Analysis**

Analytical Component (Units) <sup>1</sup> Amino Acid (% dw)	MON 88701 <sup>2</sup> Mean (S.E.) <sup>3</sup> (Range)	Control <sup>4</sup> Mean (S.E.) (Range)	Difference (MON 88701 minus Control)			Commercial Tolerance Interval <sup>5</sup> (Range)
			Mean (S.E.) (Range)	95% Confidence Interval	Significance (p-Value)	
Alanine	1.06 (0.020) (0.91 - 1.14)	1.05 (0.020) (0.88 - 1.17)	0.0026 (0.0091) (-0.13 - 0.12)	-0.017, 0.022	0.775	0.86, 1.11 (0.83 - 1.22)
Arginine	3.03 (0.10) (2.33 - 3.60)	3.15 (0.10) (2.41 - 3.77)	-0.12 (0.033) (-0.47 - 0.39)	-0.19, -0.049	0.002	2.38, 3.47 (2.30 - 3.55)
Aspartic Acid	2.39 (0.062) (1.94 - 2.64)	2.41 (0.062) (1.92 - 2.74)	-0.015 (0.027) (-0.29 - 0.29)	-0.072, 0.042	0.575	1.94, 2.57 (1.79 - 2.72)
Cystine	0.41 (0.0091) (0.32 - 0.47)	0.40 (0.0094) (0.31 - 0.46)	0.0096 (0.0070) (-0.063 - 0.082)	-0.0043, 0.023	0.174	0.31, 0.45 (0.29 - 0.47)
Glutamic Acid	4.76 (0.13) (3.80 - 5.38)	4.84 (0.14) (3.66 - 5.70)	-0.079 (0.072) (-0.78 - 0.79)	-0.23, 0.077	0.295	3.74, 5.28 (3.39 - 5.45)
Glycine	1.10 (0.020) (0.93 - 1.19)	1.09 (0.020) (0.91 - 1.20)	0.0014 (0.011) (-0.13 - 0.14)	-0.021, 0.024	0.896	0.90, 1.14 (0.85 - 1.23)
Histidine	0.74 (0.019) (0.58 - 0.85)	0.75 (0.019) (0.61 - 0.84)	-0.0014 (0.0073) (-0.062 - 0.091)	-0.017, 0.014	0.854	0.59, 0.81 (0.57 - 0.84)
Isoleucine	0.91 (0.018) (0.75 - 1.01)	0.92 (0.018) (0.77 - 1.03)	-0.0066 (0.0079) (-0.077 - 0.096)	-0.023, 0.010	0.421	0.75, 0.96 (0.72 - 1.03)
Leucine	1.53 (0.032) (1.29 - 1.70)	1.54 (0.032) (1.28 - 1.69)	-0.0018 (0.013) (-0.14 - 0.16)	-0.029, 0.026	0.892	1.25, 1.62 (1.20 - 1.72)
Lysine	1.24 (0.025) (1.05 - 1.38)	1.23 (0.025) (1.06 - 1.39)	0.0069 (0.015) (-0.11 - 0.15)	-0.026, 0.039	0.658	1.01, 1.30 (0.99 - 1.44)
Methionine	0.40 (0.0079) (0.35 - 0.46)	0.38 (0.0084) (0.32 - 0.46)	0.018 (0.0081) (-0.066 - 0.12)	0.0023, 0.035	0.026	0.32, 0.38 (0.29 - 0.49)
Phenylalanine	1.43 (0.039) (1.14 - 1.66)	1.46 (0.039) (1.15 - 1.66)	-0.022 (0.014) (-0.18 - 0.19)	-0.052, 0.0084	0.144	1.12, 1.58 (1.10 - 1.63)
Proline	1.00 (0.029) (0.82 - 1.21)	1.03 (0.029) (0.81 - 1.25)	-0.027 (0.012) (-0.12 - 0.10)	-0.052, -0.0018	0.037	0.83, 1.08 (0.79 - 1.17)
Serine	1.08 (0.025) (0.90 - 1.23)	1.09 (0.026) (0.86 - 1.24)	-0.0036 (0.015) (-0.18 - 0.16)	-0.035, 0.028	0.807	0.83, 1.21 (0.81 - 1.24)
Threonine	0.87 (0.016) (0.74 - 0.94)	0.86 (0.016) (0.73 - 0.95)	0.0057 (0.0083) (-0.10 - 0.10)	-0.012, 0.023	0.504	0.72, 0.89 (0.67 - 0.96)
Tryptophan	0.41 (0.0092) (0.33 - 0.52)	0.42 (0.0095) (0.37 - 0.52)	-0.0061 (0.0066) (-0.081 - 0.078)	-0.019, 0.0071	0.361	0.34, 0.42 (0.31 - 0.46)
Tyrosine	0.81 (0.017) (0.67 - 0.92)	0.81 (0.018) (0.67 - 0.91)	-0.0011 (0.0083) (-0.074 - 0.12)	-0.019, 0.017	0.898	0.67, 0.84 (0.63 - 0.91)
Valine	1.21 (0.027) (1.00 - 1.40)	1.23 (0.027) (1.00 - 1.40)	-0.012 (0.011) (-0.090 - 0.12)	-0.036, 0.012	0.296	1.00, 1.28 (0.97 - 1.36)

<sup>1</sup>dw = dry weight; fw = fresh weight; FA = fatty acid.

<sup>2</sup>MON 88701 was sprayed with dicamba and glufosinate.

<sup>3</sup>Mean (S.E.) = least-square mean (standard error).

<sup>4</sup>Control refers to the non-biotechnology derived, conventional control (Coker 130).

<sup>5</sup>With 95% confidence, interval contains 99% of the values expressed in the population of commercial substances. Negative limits set to zero.

### 5.3.5 Minerals and vitamins

Levels of nine key minerals and vitamin E (tocopherol) were measured in seeds from MON88701 cotton and Coker 130, as presented in Table 11. In the combined-site analysis, the mean levels of phosphorus, calcium, magnesium, manganese and potassium were

higher (statistically significant) in MON88701 than in the control line, although the magnitude of the differences between the GM and non-GM control was small. The mean levels of zinc were 6.4% lower in MON88701 compared with the control. All of the nutrient mean values for MON88701 observed in the combined-site analysis were within the 99% tolerance interval established by conventional commercial reference varieties grown concurrently in the same trial, and there was no consistent pattern observed in the analysis of individual site data.

The combined-site data showed a small increase of 6.7% in the mean levels of vitamin E in MON88701 cottonseed compared with the non-GM control. Significant differences in mean values for vitamin E between MON88701 and the control were not consistently observed at individual sites, and the mean combined-site level was within the 99% tolerance interval for conventional reference varieties grown at the same time.

**Table 11: Statistical Summary of key Vitamins and Minerals in MON88701 cottonseed, Coker 130 and commercial cultivars – Combined-site Analysis**

Analytical Component (Units) <sup>1</sup>	MON 88701 <sup>2</sup> Mean (S.E.) <sup>3</sup> (Range)	Control <sup>4</sup> Mean (S.E.) (Range)	Difference (MON 88701 minus Control)			Commercial Tolerance Interval <sup>5</sup> (Range)
			Mean (S.E.) (Range)	95% Confidence Interval	Significance (p-Value)	
Mineral						
Calcium (% dw)	0.15 (0.0093) (0.10 - 0.22)	0.13 (0.0093) (0.081 - 0.19)	0.018 (0.0022) (-0.012 - 0.038)	0.013, 0.023	<0.001	0.058, 0.21 (0.081 - 0.18)
Copper (mg/kg dw)	8.90 (0.70) (5.22 - 11.91)	8.93 (0.70) (5.40 - 11.92)	-0.025 (0.16) (-2.59 - 1.29)	-0.34, 0.29	0.875	2.97, 12.86 (4.46 - 11.62)
Iron (mg/kg dw)	67.21 (4.40) (41.96 - 83.17)	71.33 (4.48) (45.03 - 95.10)	-4.12 (2.74) (-38.15 - 12.79)	-9.96, 1.71	0.153	47.30, 97.12 (39.49 - 114.34)
Magnesium (% dw)	0.40 (0.0083) (0.35 - 0.44)	0.38 (0.0084) (0.33 - 0.44)	0.021 (0.0032) (-0.036 - 0.054)	0.015, 0.028	<0.001	0.28, 0.47 (0.31 - 0.46)
Manganese (mg/kg dw)	12.81 (0.47) (10.18 - 14.81)	11.73 (0.48) (8.61 - 14.11)	1.08 (0.28) (-1.95 - 2.54)	0.48, 1.68	0.001	9.07, 17.33 (9.07 - 17.14)
Phosphorus (% dw)	0.72 (0.031) (0.56 - 0.84)	0.72 (0.031) (0.54 - 0.87)	0.0081 (0.0067) (-0.087 - 0.11)	-0.0053, 0.021	0.230	0.49, 0.87 (0.48 - 0.87)
Potassium (% dw)	1.12 (0.028) (0.98 - 1.24)	1.07 (0.028) (0.79 - 1.27)	0.053 (0.020) (-0.12 - 0.27)	0.0089, 0.097	0.021	0.92, 1.21 (0.90 - 1.26)
Sodium (% dw)	0.034 (0.0095) (0.018 - 0.12)	0.029 (0.0096) (0.0053 - 0.10)	0.0045 (0.0046) (-0.065 - 0.030)	-0.0053, 0.014	0.346	0, 0.066 (0.0054 - 0.077)
Zinc (mg/kg dw)	37.58 (2.01) (27.31 - 46.74)	40.14 (2.02) (28.22 - 52.95)	-2.57 (0.77) (-11.57 - 3.27)	-4.22, -0.91	0.005	27.27, 44.95 (25.07 - 48.49)
Vitamin (mg/kg dw)						
Vitamin E	140.14 (9.87) (86.23 - 179.34)	131.33 (9.88) (91.78 - 162.98)	8.80 (2.07) (-6.54 - 26.36)	4.39, 13.22	<0.001	41.91, 205.89 (84.07 - 162.76)

<sup>1</sup>dw = dry weight; fw = fresh weight; FA = fatty acid.

<sup>2</sup>MON 88701 was sprayed with dicamba and glufosinate.

<sup>3</sup>Mean (S.E.) = least-square mean (standard error).

<sup>4</sup>Control refers to the non-biotechnology derived, conventional control (Coker 130).

<sup>5</sup>With 95% confidence, interval contains 99% of the values expressed in the population of commercial substances. Negative limits set to zero.

### 5.3.6 Anti-nutrients

Cottonseed samples from MON88701 and the conventional control Coker 130 were analysed for five anti-nutrient compounds characteristic of cotton. As presented in Table 12, the components that showed statistically significant increases in mean values between MON88701 and the control in the combined-site analysis were the cyclopropenoid fatty acid dihydrosterculic acid, free gossypol and total gossypol, with the relative magnitude of the differences being 9.6%, 6.2% and 6.8% respectively.

**Table 12: Statistical Summary of Anti-Nutrients in MON88701 cottonseed, Coker 130 and commercial cultivars – Combined-site Analysis**

Analytical Component (Units) <sup>1</sup>	MON 88701 <sup>2</sup> Mean (S.E.) <sup>3</sup> (Range)	Control <sup>4</sup> Mean (S.E.) (Range)	Difference (MON 88701 minus Control)			Commercial Tolerance Interval <sup>5</sup> (Range)
			Mean (S.E.) (Range)	95% Confidence Interval	Significance (p-Value)	
Cyclopropenoid Fatty Acid (% Total FA)						
Dihydrosterculic Acid	0.15 (0.0034) (0.11 - 0.19)	0.14 (0.0037) (0.11 - 0.17)	0.013 (0.0044) (-0.026 - 0.068)	0.0044, 0.022	0.003	0.078, 0.25 (0.038 - 0.23)
Malvalic Acid	0.39 (0.015) (0.20 - 0.55)	0.37 (0.016) (0.26 - 0.49)	0.013 (0.015) (-0.16 - 0.16)	-0.016, 0.043	0.371	0.23, 0.54 (0.11 - 0.59)
Sterculic Acid	0.22 (0.0067) (0.13 - 0.29)	0.21 (0.0072) (0.17 - 0.27)	0.0067 (0.0081) (-0.085 - 0.078)	-0.0096, 0.023	0.412	0.17, 0.27 (0.061 - 0.34)
Gossypol (% dw)						
Free Gossypol	0.94 (0.037) (0.80 - 1.18)	0.89 (0.037) (0.68 - 1.20)	0.055 (0.020) (-0.086 - 0.20)	0.012, 0.099	0.016	0.099, 1.57 (0.50 - 1.41)
Total Gossypol	1.04 (0.037) (0.84 - 1.24)	0.97 (0.037) (0.74 - 1.10)	0.066 (0.017) (-0.021 - 0.23)	0.031, 0.10	<0.001	0.064, 1.76 (0.56 - 1.61)

<sup>1</sup>dw = dry weight; FA = fatty acid.

<sup>2</sup>MON 88701 was sprayed with dicamba and glufosinate.

<sup>3</sup>Mean (S.E.) = least-square mean (standard error).

<sup>4</sup>Control refers to the non-biotechnology derived, conventional control (Coker 130).

<sup>5</sup>With 95% confidence, interval contains 99% of the values expressed in the population of commercial substances. Negative limits set to zero.

Considering the data from the eight individual sites in the trial, statistically significant differences in the mean level of dihydrosterculic acid occurred at one site; free gossypol at two sites; and total gossypol at three sites. Assessment of the reproducibility of the significant differences observed in the combined-site analysis therefore indicates a lack of consistency across all trial sites. All individual-site mean values for these anti-nutrients were within the 99% tolerance interval established from the conventional commercial reference varieties grown concurrently with MON88701 and Coker 130 in the same trial. Overall, the observed differences in anti-nutrient values between MON88701 and the conventional control were not considered to be meaningful from a nutritional perspective because they were generally small in magnitude, were variable across trial sites, and the mean levels in MON88701 in all cases were within the natural range of variability of existing commercial cotton varieties.

#### 5.4 Conclusion from compositional analysis

Detailed compositional analyses on delinted cottonseed from herbicide-treated MON88701 (glufosinate and dicamba), the control Coker 130 and nine commercial varieties were conducted on plants grown under normal agricultural conditions at eight trial sites in cotton growing regions of the U.S.. The analyses included proximates (protein, fat, ash, moisture, carbohydrates by calculation), fibre components (ADF, NDF, CF), fatty acids, amino acids, micronutrients (minerals and  $\alpha$ -tocopherol) and anti-nutrients (gossypol and cyclopropenoid fatty acids). The levels of these key constituents in MON88701 were compared to those in the non-GM control and to the ranges found in other commercial non-GM cotton varieties grown concurrently in the same trial, or as reported in the literature.

For the combined-site analysis, statistically significant differences were found in a number of individual analytes between MON88701 and the control, however, with the exception of two analytes (ash and calcium), in each case the differences occurred at only some sites, were small in magnitude, and were within tolerance intervals. In addition, all nutrient mean values were within the range of natural variability for commercial cottonseed. The composition of cotton can vary significantly with the site and agricultural conditions, and the differences reported here most likely reflect normal biological variability.

Overall, the compositional data are consistent with the conclusion that there are no biologically significant differences in the levels of key constituents in MON88701 cottonseed when compared with conventional cotton cultivars already available in agricultural markets.

## 6. Nutritional impact

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

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

MON88701 cotton is the result of a simple genetic modification to confer dual herbicide tolerance, with no intention to significantly alter nutritional parameters in the food. The compositional analyses have demonstrated that the genetic modification has not altered the nutritional adequacy of MON88701 as a source of food when compared with that of conventional cotton varieties. The introduction of foods derived from MON88701 cotton into the food supply is therefore expected to have negligible nutritional impact.

The Applicant submitted one feeding study with cotton MON88701, which has been evaluated by FSANZ.

### Study submitted:

Evaluation of Cottonseed Meal from MON88701 as a Feed Ingredient for Channel Catfish. Completion date: March, 2012. Testing facility: Smithers Viscient, Snow Camp, North Carolina 27349. Monsanto Study No: SE-2011-0191 (unpublished).

The feeding study assessed the growth performance and general health of channel catfish fed diets containing cottonseed meal from either MON88701 or non-GM cotton for approximately eight weeks. The results of this study, which did not show any significant differences between meal from MON88701 compared to non-GM cotton, were consistent with the findings from the compositional analyses (Section 5).

## References

- Abu-Qarn, M., Eichler, J. and Sharon, N. (2008) Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Current Opinion in Structural Biology* 18:544-550.
- Astwood, J.D. and Fuchs, R.L. (1996) Allergenicity of foods derived from transgenic plants. *Highlights in food allergy. Monographs in Allergy*, 32. 105-120.
- Barker, R.F., Idler, K.B., Thompson, D.V. and Kemp J.D. (1983). Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. *Plant Molecular Biology* 2:335-350.
- Baxeavanis, A.D. (2005) Assessing Pairwise Sequence Similarity: BLAST and FASTA. In: Baxeavanis, A.D. and Ouellette, B.F.F. eds. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. Chapter 11. John Wiley & Sons, Inc., pp. 295-324.

- Bevan, M., Barnes, W.M. and Chilton, M.D. (1983). Structure and transcription of the nopaline synthase gene region of T-DNA. *Nucleic Acids Research* 11:369-385.
- Bradbury, J.F. (1986) *Guide to Plant Pathogenic Bacteria*. Cambridge News Ltd, Aberystwyth, p190-197.
- Cairo, R.C., Silva, L.R., deAndrade, C.F., deAndrade Barberino, M.G., Bandeira, A.C., Santos, K.P. and Diniz-Santos, D.R. (2008). Bacterial contamination in milk kitchens in pediatric hospitals in Salvador, Brazil. *Brazilian Journal of Infectious Diseases* 12(3):217-221.
- Cherry, J.P. (1983). Cottonseed oil. *Journal of the American Oil Chemists' Society* 60:360-367.
- Codex Alimentarius (2009). Foods derived from modern biotechnology. Codex Alimentarius Commission, Joint FAO/WHO Foods Standard Programme, Food and Agriculture Organisation of the United Nations, Rome, Italy.
- Deblaere, R., Reynaerts, A., Höfte, H., Hernalsteens, J.-P., Leemans, J. and van Montagu, M. (1987) Vectors for cloning in plant cells. *Methods in Enzymology* 153:277-292.
- Delaney, B., Astwood, J.D., Cunny, H., Eichen Conn, R., Herouet-Guicheney, C., MacIntosh, S., Meyer, L.S., Privalle, L., Gao, Y., Mattsson, J., Levine, M. and ILSI. (2008) Evaluation of protein safety in the context of agricultural biotechnology. *Food and Chemical Toxicology* 46:S71-S97.
- Denton, M., Todd, N.J., Kerr, K.G., Hawkey, P.M. & Littlewood, J.M. (1998). Molecular epidemiology of *Stenotrophomonas maltophilia* isolated from clinical specimens from patients with cystic fibrosis and associated environmental samples. *Journal of Clinical Microbiology* 36:1953-1958.
- Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M. (1982) Nopaline synthase: transcript mapping and DNA sequence. *Journal of Molecular and Applied Genetics* 1(6):561-573.
- D'Ordine, R.L., Rydel, T.J., Storek, M.J., Sturman, E.J., Moshiri, F., Bartlett, R.K., Brown, G.R., Eilers, R.J., Dart, C., Qi, Y., Flasiński, S. and Franklin, S.J. (2009). Dicamba monooxygenase: Structural insights into a dynamic Rieske oxygenase that catalyzes an exocyclic monooxygenation. *Journal of Molecular Biology* 392:481-497.
- Dumitru, R., Jiang, W.Z., Weeks, D.P. and Wilson, M.A. (2009). Crystal structure of dicamba monooxygenase: A Rieske nonheme oxygenase that catalyzes oxidative demethylation. *Journal of Molecular Biology* 392:498-510.
- Duncan, D.R. (2010) Cotton Transformation. Pages 65-77 in *Cotton: Biotechnological Advances*, vol65. U.B. Zehr (ed.). Springer-Verlag, Berlin, Germany.
- Duncan, D.R. and Ye, G. (2011) Methods for inducing cotton embryogenic callus. Patent 7,947,869, U.S. Patent Office, Washington, D.C.
- EFSA. (2008) Safety and nutritional assessment of GM plants and derived food and feed: The role of animal feeding trials. *Food and Chemical Toxicology* 46:S1-S70.  
[http://www.efsa.europa.eu/EFSA/efsa\\_locale-1178620753812\\_1211902590265.htm](http://www.efsa.europa.eu/EFSA/efsa_locale-1178620753812_1211902590265.htm). Accessed 5 April 2010.
- EFSA. (2011) Conclusion on the peer review of the pesticide risk assessment of the active substance dicamba. *The EFSA Journal* 9(1):1965 [52 pp.] Online: [www.efsa.europa.eu/efsajournal.htm](http://www.efsa.europa.eu/efsajournal.htm)
- FARRP (2011). Allergen database. University of Nebraska, Food Allergy Research and Resource Program, Lincoln, Nebraska. <http://www.allergenonline.org/databasebrowse.shtml> [accessed 14 February 2011].

- Fling, M.E., Kopf, J. and Richards, C. (1985) Nucleotide sequence of the transposon Tn7 gene encoding an aminoglycoside-modifying enzyme, 3''(9)-O-nucleotidyltransferase. *Nucleic Acids Research* 13(19):7095-7106.
- Fraley, R.T., Rogers, S.G., Horsch, R.B., Sanders, P. R., Flick, J.S., Adams, S.P., Bittner, M.L., Brand, L.A., Fink, C.L., Fry, J.S., Galluppi, G.R., Goldberg, S.B., Hoffmann, N.L. and Woo, S.C. (1983). Expression of bacterial genes in plant cells. *Proceedings of the National Academy of Sciences of the United States of America* 80:4803-4807.
- FSANZ (2002) *Application A436 - Oil and Linters Produced from Insect-Protected Cotton Containing Event 15985*. Report Prepared by Food Standard Australia New Zealand. <http://www.foodstandards.gov.au/foodstandards/applications/applicationa436insectprotectedcottonevent15985/index.cfm>.
- FSANZ (2010) *Application A1028 - Food Derived from Insect-Protected and Herbicide-Tolerant Cotton Line T304-40: Safety Assessment Report*. Report prepared by Food Standards Australia New Zealand. <http://www.foodstandards.gov.au/foodstandards/applications/applicationa1028oild4457.cfm>.
- Giza, P.E. and Huang, R.C.C. (1989). A self-inducing runaway-replication plasmid expression system utilizing the Rop protein. *Gene* 78:73-84.
- Goodman, R.E. (2006) Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. *Molecular Nutrition and Food Research* 50:655-660.
- Hajdukiewicz, P., Svab, Z. and Maliga, P. (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Molecular Biology* 25:989-994.
- Herman, P.L., Behrens, M., Chakraborty, B.M., Chrastil, B.M., Barycki J. and Weeks, D.P. (2005). A three-component dicamba O-demethylase from *Pseudomonas maltophilia*, strain DI-6: Gene isolation, characterisation, and heterologous expression. *Journal of Biological Chemistry* 280:24759-24767.
- H rouet, C., Esdaile, D.J., Mallyon, B.A., Debruyne, E., Schulz, A., Currier, T., Hendrickx, K., van der Klis, R.-J. and Rouan, D. (2005) Safety evaluation of the phosphinothricin acetyltransferase proteins encoded by the *pat* and *bar* sequences that confer tolerance to glufosinate-ammonium herbicide in transgenic plants. *Regulatory Toxicology and Pharmacology* 41:134-149.
- Herrmann, K.M. (1995). The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. *Plant Cell* 7:907-919.
- ILSI (2011) *International Life Sciences Institute Crop Composition Database Version 3.0* available online at [http://www.cropcomposition.org/cgi-perl/search\\_orc.cgi](http://www.cropcomposition.org/cgi-perl/search_orc.cgi).
- John, M.E. (1996). Structural characterisation of genes corresponding to cotton fiber mRNA, E6: Reduced E6 protein in transgenic plants by antisense gene. *Plant Molecular Biology* 30:297-306.
- Joshi, C.P., Zhou, H., Huang, X. and Chiang, V.L. (1997) Context sequences of translation initiation codon in plants. *Plant Molecular Biology* 35:993-1001.
- Kay, R., Chan, A., Daly, M. and McPherson, J. (1987). Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science* 236:1299-1302.
- Kimber, I., Kerkvliet, N.I., Taylor, S.L., Astwood, J.D., Sarlo, K. and Dearman, R.J. (1999) Toxicology of protein allergenicity: prediction and characterization. *Toxicol.Sci* 48(2):157-162.
- Klee, H.J., Muskopf, Y.M. and Gasser C.S. (1987). Cloning of an *Arabidopsis thaliana* gene encoding 5-enolpyruvylshikimate-3-phosphate synthase: Sequence analysis and manipulation to obtain glyphosate-tolerant plants. *Molecular and General Genetics* 210:437-442.



- Krueger, J.P., Butz, R.G., Atallah, Y.H. and Cork, D.J. (1989) Isolation and identification of microorganisms for the degradation of dicamba. *J Agric Food Chem* 37:534-538.
- Kützner, H.J. (1981) The Family Streptomycetaceae. In: Starr, M.P., Stolp, H., Trüper, H.G., Ballows, A., and Schlegel, H.G. eds. *The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*. Springer Verlag, Berlin, pp. 2028-2090.
- Ladics, G.S., Bannon, G.A., Silvanovich, A. and Cressman, R.F. (2007). Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens. *Molecular Nutrition and Food Research* 51:985-998.
- Looney, W.J. (2009) *Stenotrophomonas maltophilia*: an emerging opportunist human pathogen. *The Lancet Infectious Diseases* 9(5):312-323.
- Maiti, I.B. and Shepherd R.J. (1998). Isolation and expression analysis of peanut chlorotic streak caulimovirus (PCISV) full-length transcript (FLt) promoter in transgenic plants. *Biochemical and Biophysical Research Communications* 244:440-444.
- Metcalfe, D.D., Astwood, J.D., Townsend, R., Sampson, H.A., Taylor, S.L. and Fuchs, R.L. (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Crit Rev Food Sci Nutr* 36 Suppl:S165-S186.
- Miletich, J.P. and Broze Jr., G.J. (1990) b Protein C is not glycosylated at asparagine 329. *The Journal of Biological Chemistry* 265:11397-11404.
- Molony, L.A. and Hassall, G.H.D. (2008) *Signposts for Australian Agriculture - The Australian Cotton Industry*. National Land and Water Resources Audit, Australian Government.  
<http://nlwra.gov.au/files/products/national-land-and-water-resources-audit/pn21908/pn21908.pdf>. Accessed 29 March 2010.
- Murakami, T., Anzai, H., Imai, S., Satoh, A., Nagaoka, K. and Thompson, C.J. (1986) The bialaphos biosynthetic genes of *Streptomyces hygroscopicus*: Molecular cloning and characterization of the gene cluster. *Molecular and General Genetics* 205:42-50.
- Niepel, M. and Gallie D.R. (1999). Identification and characterisation of the functional elements within the tobacco etch virus 5' leader required for cap-independent translation. *Journal of Virology* 73:9080-9088.
- Odell, J.T., Nagy, F. and Chua, N.H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* 313(6005):810-812.
- OECD (2003) *Considerations for the Safety Assessment of Animal Feedstuffs Derived from Genetically Modified Plants*. Series on the Safety of Novel Foods and Feeds, No. 9. Organisation for Economic Cooperation and Development, Paris.  
[http://www.oilis.oecd.org/oilis/2003doc.nsf/LinkTo/NT0000426A/\\$FILE/JT00147696.PDF](http://www.oilis.oecd.org/oilis/2003doc.nsf/LinkTo/NT0000426A/$FILE/JT00147696.PDF). Accessed 22 February 2010.
- OECD (2004) *Consensus Document on Compositional Considerations for New Varieties of Cotton (Gossypium hirsutum and Gossypium barbadense: Key Food and Feed Nutrients and Anti-Nutrients*. Report No. ENV/JM/MONO(2004)16, Organisation for Economic Co-operation and Development.  
[http://www.oilis.oecd.org/oilis/2004doc.nsf/LinkTo/NT0000480E/\\$FILE/JT00168142.PDF](http://www.oilis.oecd.org/oilis/2004doc.nsf/LinkTo/NT0000480E/$FILE/JT00168142.PDF). Accessed 22 February 2010.
- OGTR (2008a) *Risk Assessment and Risk Management Plan for DIR 087: Limited and Controlled Release of Cotton Genetically Modified for Insect Resistance and Herbicide Tolerance*. Document prepared by the Office of the Gene Technology Regulator, Australia.  
<http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/dir087-2008>. Accessed 9 February 2010.
- OGTR (2008b) *The Biology of Gossypium hirsutum L. and Gossypium barbadense L. (cotton)*. Document prepared by the Office of the Gene Technology Regulator, Australia.

[http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/cotton-3/\\$FILE/biologycotton08.pdf](http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/cotton-3/$FILE/biologycotton08.pdf). Accessed 9 February 2010.

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

Palleroni, N.J. & Bradbury, J.F. (1993). *Stenotrophomonas*, a new bacterial genus for *Xanthomonas maltophilia* (Hugh 1980) Swings et al. 1983. *International Journal of Systematic Bacteriology* 43:606-609.

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

Quereshi, A., Mooney, L., Denton, M. and Kerr, K.G. (2005). *Stenotrophomonas maltophilia* in salad. *Emerging infectious diseases* 11:1157-1158.

Rensing, S.A. and Maier, U.G. (1994). Phylogenetic analysis of the stress-70 protein family. *Journal of Molecular Evolution* 39:80-86.

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

Ryan, R.P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M.B., Berg, G., van der Lelie, D. and Dow, J.M. (2009) The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. *Nature Reviews Microbiology* 7:514-525.

Sacchetti, R., De Luca, G. and Zanetti, F. (2009). Control of *Pseudomonas aeruginosa* and *Stenotrophomas maltophilia* contamination of microfiltered water dispensers with peracetic acid and hydrogen peroxide. *International Journal of Food Microbiology* 132:162-166.

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

Stalker, D.M., Thomas, C.M. and Helinski, D.R. (1981). Nucleotide sequence of the region of the origin of replication of the broad host range plasmid RK2. *Molecular and General Genetics* 181:8-12.

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

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

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

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



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

Wang, X.Z., Li, B., Herman, P.L. and Weeks, D.P. (1997). A three-component enzyme system catalyzes the O-demethylation of the herbicide dicamba in *Pseudomonas maltophilia* DI-6. *Applied and Environmental Microbiology* 63:1623-1626.

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

Winter, J., Wright, R., Duck, N., Gasser, C., Fraley, R. and Shah, D. (1988). The Inhibition of petunia hsp70 mRNA processing during CdCl<sub>2</sub> stress. *Molecular and General Genetics* 211:315-319.

Zambryski, P., Depicker, A., Kruger, K. and Goodman H.M. 1982. Tumor induction by *Agrobacterium tumefaciens*: Analysis of the boundaries of T-DNA. *Journal of Molecular and Applied Genetics* 1:361-370.