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

29 September 2017
[26–17]

Supporting document 1

Safety Assessment Report (at Approval) – Application A1140

Food derived from Herbicide-tolerant Canola Line MS11

Executive summary

Background

A genetically modified (GM) canola line with OECD Unique Identifier BCS-BN012-7 (herein referred to as MS11) has been developed by Bayer CropScience. This canola line has been genetically modified to confer two novel agronomic traits – tolerance to the broad spectrum herbicide glufosinate ammonium (glufosinate) and expression of male sterility.

Tolerance to glufosinate is achieved through constitutive expression of phosphinothricin N-acetyltransferase (PAT) encoded by the *bar (bialaphos)* gene from *Streptomyces hygroscopicus*. The PAT protein has been assessed by FSANZ in 22 previous FSANZ applications, and globally is represented in six major crop species and over 30 approved GM single plant events.

Male sterility is conferred by the *barnase* gene from *Bacillus amyloliquefaciens* that is driven by a promoter specific to the tapetal cells of the developing anthers of MS11 canola. The Barnase (bacterial cytotoxic ribonuclease) protein causes RNA degradation, cell disruption, and cell death and hence leads to ablation of the tapetal cells that surround the pollen sac thereby preventing normal pollen formation. Hence MS11 is unable to either self-pollinate or pollinate other plants but the female reproductive parts of the flower remain functional. The Applicant's intention is to use the male sterile (MS) line in a hybrid breeding system in which MS11 (as the female parent line) is outcrossed with an agronomically-superior male line (the pollen donor) containing a protein (Barstar) which inhibits the Barnase protein, thus restoring fertility in the seed sown by the farmer. The plants germinating from this seed therefore show hybrid vigour as well as being able to self-pollinate and produce seed that is harvested for the food/feed market. MS11 will not, itself, be used as a food producing line.

MS11 also contains the *barstar* gene from *B. amyloliquefaciens*. The resulting Barstar protein is only weakly expressed and is not sufficient to override the effect of Barnase produced in the anther. However, it is sufficient to inhibit any Barnase that is inadvertently expressed in tissues other than the anther and which may adversely affect agronomic performance. Thus the presence of the *barstar* gene in MS11 assists in improving the quality of male-sterile lines identified during the selection phase.

In conducting a safety assessment of food derived from MS11, a number of criteria have been addressed including: a characterisation of the transferred gene sequences, their origin, function and stability in the canola genome; the changes at the level of DNA and protein in the whole food; compositional analyses; and evaluation of intended and unintended changes. This safety assessment report addresses only food safety and nutritional issues of the GM food *per se*. It therefore does not address:

- environmental risks related to the environmental release of GM plants used in food production
- the safety of animal feed, or animals fed with feed, derived from GM plants
- the safety of food derived from the non-GM (conventional) plant.

History of use

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

Canola seeds are processed into two major products, oil and meal. The oil is the major product for human consumption, being used directly for cooking and as an ingredient in a variety of manufactured food products including salad and cooking oil, margarine, shortening and a range of prepared foods such as mayonnaise, sandwich spreads, creamers and coffee whiteners. Canola oil is the third largest source of vegetable oil in the world after soybean oil and palm oil. Whole canola seeds are being used increasingly in products such as breads and there is potential for canola meal to be used as a source of protein isolate.

Molecular characterisation

MS11 was generated through *Agrobacterium*-mediated transformation with a single T-DNA containing three expression cassettes. Comprehensive molecular analyses indicate there is a single insertion site, in chromosome A03, comprising a single, complete copy of each of the *bar, barnase* and *barstar* genes with their regulatory elements. The introduced genes are stably inherited from one generation to the next. No plasmid backbone has been incorporated into the transgenic locus and no endogenous genes have been disrupted as a result of the transformation process.

Characterisation and safety assessment of new substances

Newly expressed proteins

MS11 expresses three novel proteins, PAT, Barnase and Barstar.

Mean levels of all proteins in the edible part (i.e. seed) were below or close to the Limit of Quantification (LOQ). The mean level of PAT protein was highest in whole plants sampled at the 3 – 5 leaf stage, while lowest values were obtained in root and seed samples. The Barnase protein was below the LOQ in all samples tested. Levels of Barstar in all tissues were either below the LOQ or very low; roots marginally appeared to have the highest mean level.

A number of studies were used to confirm the identity and physicochemical properties of the plant-derived PAT protein. These studies demonstrated that the protein conforms in size, amino acid sequence and activity to that expected, and does not exhibit any post-translational modification including glycosylation.

Very low yields of the plant-produced Barnase and Barstar proteins precluded their specific characterisation. However the weight of evidence, provided by a) translation of the known DNA sequences of the two genes introduced into MS11, b) the fact that the proteins function as predicted in the plant and c) the detailed characterisation of equivalent microbially-produced proteins is sufficient to confirm the identity of the proteins expressed in MS11.

For all three proteins, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens; digestibility studies suggest the proteins would be rapidly degraded in the gastro-intestinal tract following ingestion; and thermolability studies indicate the three proteins are functionally inactivated following heating.

Taken together, the evidence indicates that should PAT, Barnase or Barstar be present in the diet they are unlikely to be toxic or allergenic in humans.

Herbicide metabolites

For PAT, the metabolic profiles resulting from the novel protein x herbicide interaction have been established through a significant history of use. The glufosinate-tolerance trait is present in lines from 22 previous applications to FSANZ. There are no concerns that the spraying of MS11 with glufosinate ammonium would result in the production of metabolites that are not also produced in crops sprayed with the same herbicide and already used in the food supply.

Compositional analyses

Detailed compositional analyses were done to establish the nutritional adequacy of seed from MS11 and to characterise any unintended compositional changes. Seed samples were analysed for proximates, fibre, fatty acids, amino acids, minerals, vitamins, anti-nutrients (phytic acid, tannins and sinapine) and glucosinolates. In total, 87 analytes were considered of which 30 had negligible levels that precluded inclusion in a statistical analysis.

The levels of each analyte from glufosinate-sprayed and unsprayed MS11 were compared to levels in: a) the non-GM parental line, N90-740, b) six non-GM commercial reference lines and c) levels recorded in the literature.

Of the 57 analytes considered, only two – gluconapin and insoluble tannins - were significantly different in a comparison between unsprayed MS11 and the control. In both instances, the levels fell within the 95% tolerance interval generated from the reference lines.

In contrast to this, in the comparison of analytes between sprayed MS11 and the control, 31 analytes were significantly different - most being higher in MS11 than in the control. It was expected there would be little consistency in analyte levels between the seeds from sprayed and unsprayed MS11 because of the different pollen sources used to fertilise the sprayed MS11 plants. However, in all cases the levels in the sprayed MS11 seed were within the tolerance interval. In six instances the levels exceeded the literature range - but in four of these, the levels in the non-GM control were also higher than the literature range.

The conclusion is that seed from MS11, whether from unsprayed MS11 plants or plants sprayed with glufosinate is compositionally equivalent to seed from conventional canola varieties.

Conclusion

No potential public health and safety concerns have been identified in the assessment of MS11. On the basis of the data provided in the present Application, and other available information, food derived from MS11 is considered to be as safe for human consumption as food derived from conventional canola varieties.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ai/L	active ingredient per litre
ADI	Acceptable Daily Intake
ADF	acid detergent fibre
AP	alkaline phosphatase
ARfD	Acute reference dose
bar	Bialaphos resistance
Barnase	bacterial cytotoxic ribonuclease
BLAST	Basic Local Alignment Search Tool
BLAT	BLAST-like alignment tool
bp	base pairs
CoA	co-enzyme A
DIG	digoxigenin
Code	*Australia New Zealand Food Standards Code*
DMPT	demethylphosphinothricin
DNA	deoxyribonucleic acid
dw	dry weight
ELISA	enzyme linked immunosorbent assay
EMBOSS	European Molecular Biology Open Software Suite
FAO	Food & Agricultural Organization of the United Nations
FARRP	Food Allergy Research and Resource Program
FASTA	Fast Alignment Search Tool – All
FSANZ	Food Standards Australia New Zealand
fw	fresh weight
g	gram
GM	genetically modified
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
IC_{50}	half maximal inhibitory concentration
kb	kilo base
kDa	kilo Dalton
kg	kilogram
LB	Left Border of T-DNA (*Agrobacterium tumefaciens*)
LC	liquid chromatography
LOQ	limit of quantification
mg	milligram
MRL	Maximum residue limit
MS	male sterile OR mass spectrometry
MT	Million tonnes
NCBI	National Centre for Biotechnology Information
ND	not detectable
NDF	neutral detergent fibre
ng	nanogram
nos	nopaline synthase
NS	not significant
OECD	Organisation for Economic Co-operation and Development

OGTR	Office of the Gene Technology Regulator
ORF	open reading frame
PAT	Phosphinothricin acetyltransferase
PCR	polymerase chain reaction
PPT	phosphinothricin
PRESS	predicted residual sum of squares
PVDF	polyvinylidene fluoride
RB	Right Border of T-DNA (*Agrobacterium tumefaciens*)
RF	fertility restorer
RNA	ribonucleic acid
RNase	ribonuclease
Rubisco	ribulose bisphosphate carboxylase
SAS	Statistical Analysis Software
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
SSU	small sub-unit
T-DNA	transfer DNA
Ti	tumour inducing
U	Units (U/mg = measure of enzyme activity)
µg	microgram
µM	micro mole
UPLC	ultra performance liquid chromatography
USA	United States of America
UTR	untranslated region
UV	ultraviolet

1 Introduction

Bayer CropScience (Bayer) has submitted an application to FSANZ to vary Schedule 26 in the *Australia New Zealand Food Standards Code* (the Code) to include food from a new genetically modified (GM) canola (*Brassica napus*) line, MS11, with OECD Unique Identifier BCS-BN012-7 (herein referred to as MS11). This canola line has been genetically modified to confer two novel agronomic traits – tolerance to the broad spectrum herbicide glufosinate ammonium (glufosinate) and expression of male sterility.

Tolerance to glufosinate is achieved through constitutive expression of phosphinothricin N-acetyltransferase (PAT) encoded by the *bar (bialaphos resistance)* gene from *Streptomyces hygroscopicus*. 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). The PAT protein has been assessed by FSANZ in 22 previous FSANZ applications, and globally is represented in six major crop species and over 30 approved GM single plant events (CERA 2011).

Male sterility is conferred by the *barnase* gene from *Bacillus amyloliquefaciens* that is driven by a promoter specific to the tapetal cells of the developing anthers of MS11. The Barnase (bacterial cytotoxic ribonuclease) protein causes RNA degradation, cell disruption, and cell death and hence leads to ablation of the tapetal cells that surround the pollen sac. Since the tapetum produces substances that aid in development, or become components of, the outer pollen wall (Hartley 1989; Mariani et al. 1990), normal pollen formation is prevented and no detectable pollen grains are present at anther dehiscence. The Applicant's intention is to use the male sterile (MS) line in a hybrid breeding system in which MS11, as the female parent line, is crossed with a male line (RF) containing the fertility restoration, ribonuclease inhibitor, Barstar encoded by the *barstar* gene also derived *from B. amyloliquefaciens* (i.e. the MS 11 line is not intended by itself to be a food-producing line).

The resulting progeny co-express both genes but the *barstar* gene is dominant to the *barnase* gene and ribonuclease (RNase) activity is suppressed by the formation of RNase/RNase inhibitor complexes (Mariani et al. 1990; Mariani et al. 1992) i.e. the hybrids are fully fertile. In nature, canola reproduces predominantly through self-pollination but the resulting progeny lack the vigour of those produced by outcrossing. It is therefore advantageous for a primary seed breeder to maximise seed quality by preventing self-pollination and having a male sterile (female) parent that is then forced to outcross with an agronomically superior line. However, the resulting seed, which is planted by the farmer, needs to produce fertile plants that will go on to self-pollinate and produce seed that is harvested for the food/feed market. The Barnase/Barstar system allows both of the objectives to be met.

FSANZ has previously approved food derived from the intended fertility restorer line – RF3 – (as well as other MS lines) in Application A372 (FSANZ 2002) and therefore has assessed both the Barnase and Barstar proteins.

Glufosinate tolerance not only provides a trait useful for weed control but also contributes to the breeding system by being used to maintain the MS line. When the MS line is crossed with a non-GM counterpart there is a 1:1 segregation of male fertile/herbicide-sensitive plants and male sterile/herbicide-tolerant plants. The male fertile plants are eliminated by spraying with glufosinate leaving the seed-producing male sterile plants unharmed. Glufosinate tolerance is also used for selection of putative transformants during the transformation stage (see Section 3.1)

MS11 also contains the *barstar* gene, but in this instance it is driven by a weak constitutive promoter which means the Barstar protein is weakly expressed in all tissues, not just the tapetum. The level of expression in the tapetum is not sufficient to counteract the effect of Barnase, the gene of which has a strong, tapetum-specific promoter. However, it has been noted that sometimes endogenous enhancer elements in the host genome can, depending upon where the *barnase* gene has been integrated, result in low-level expression of Barnase in tissues other than the tapetum and this could potentially adversely affect agronomic performance. Therefore, the presence of Barstar in those tissues can decrease the between-transformant variability in expression of Barnase, thereby increasing the frequency of obtaining transformants with good agronomic performance i.e. the presence of Barstar in non-tapetum cells actually improves the likelihood of obtaining a high percentage of good-performing male sterile plants (Michiels et al. 2000).

It is the Applicant's intention to commercially cultivate canola varieties containing the MS11 event in all major canola-producing countries including Australia, where application will be made to the Office of the Gene Technology Regulator (OGTR) for a commercial release. Therefore, it is anticipated food products derived from canola containing this event will enter the Australian and New Zealand food supplies via local production and imports from major canola-producing countries.

2 History of use

2.1 Host organism

Canola (a trade name purported to be derived from 'Canadian oil low acid') is the name used for rapeseed (*Brassica napus, Brassica rapa* or *Brassica juncea*) crops that have less than 2% erucic acid (a fatty acid)[1] and less than 30 micromoles of glucosinolates per gram of seed solids (OECD 2001). Canola varieties were first developed in Canada in the 1950s, using traditional breeding techniques, in response to a demand for food-grade rapeseed products and animal feed with improved palatability. Rapeseed-derived products that do not meet the compositional standard cannot use the trademarked term, *canola*. In some countries, the term canola is not used and the low erucic acid crop is known generically as rapeseed.

Rapeseed is the second largest oilseed crop in the world behind soybean. In 2014, world production was 73.8 MT and the major oilseed rape producing countries were Canada (15.6 MT), China (14.8 MT) and India (7.8 MT); Australia ranked 6th at 3.8 MT (FAOSTAT3 2015). In the case of China and India, a significant amount of non-canola quality rapeseed, is included in the term 'rapeseed'. All of Australia's production is canola. New Zealand canola production was minor at approximately 3,000 T. In 2013, Canada was the largest exporter of canola seed (6.9 MT), while Australia was the second largest exporter at 3.8 MT. Australia's major export destinations in 2015–16 were Europe, Japan and Bangladesh (AOF 2016) and represented around 15% of the world's canola export trade.

In Australia, canola is the third largest broad acre crop behind wheat and barley and the growing area extends from south-western Western Australia to south-eastern Australia and northern New South Wales. GM canola now accounts for approximately 20% of the total Australian canola crop.

[1] Codex Standard for Named Vegetable Oils (CODEX-STAN 210-1999) - http://www.fao.org/docrep/004/y2774e/y2774e04.htm

Canola seeds are processed into two major products, oil and meal. Briefly, the processes involved in preparation of the oil and meal (CCC 2012) involve seed cleaning, seed pre-conditioning and flaking, seed cooking, pressing the flake to mechanically remove a portion of the oil, solvent extraction of the press-cake to remove the remainder of the oil, and desolventising and toasting of the meal. The oil is the major product for human consumption, being used in a variety of manufactured food products including salad and cooking oil, margarine, shortening and a range of prepared foods such as mayonnaise, sandwich spreads, creamers and coffee whiteners. The meal provides a good protein source in stock feed for a variety of animals, primarily pigs, poultry and dairy cattle (Bonnardeaux 2007). More recently, it has been identified as a potential alternative source of protein isolate for aquaculture[2] and human consumption (Campbell et al. 2016). Whole canola seeds are being used increasingly in products such as breads.

Another possible food product that can be derived from the canola plant is bee pollen (Bogdanov 2016). In the case of MS11, since pollen grains are not produced, this line would not be a source of bee pollen. However, the MS11 x RF3 hybrid would produce viable pollen.

The canola variety used as the recipient of the DNA insertion to create MS11 was *B. napus* N90-740. This is a non-GM, open-pollinated non-proprietary line that originated from Agriculture and Agri-Food Canada and has good transformation characteristics.

2.2 Donor organisms

2.2.1 *Bacillus amyloliquefaciens*

Bacillus amyloliquefaciens, a non-pathogenic soil bacterium, is the source of the *barnase* and *barstar* genes. It is closely related to *Bacillus subtilis* and, over the years, the classification of the organism has seen it as a subspecies of *B. subtilis*, a separate species (current classification) and as a member of the so-called *B. subtilis complex* (Priest et al. 1987; Environment Canada and Health Canada 2015).The following information about *B. amyloliquefaciens* has been sourced from a Report prepared by Environment Canada and Health Canada (2015).

- It is widely distributed in nature in a variety of habitats; it has been isolated from the inner tissues and rhizosphere of healthy plants and has a history of use in industrial fermentation and pest control e.g. strains have been approved for use as biocontrol agents against fungal disease in terrestrial plants

- There is no evidence that it causes infection or adverse effects in aquatic or terrestrial plants, vertebrates or invertebrates; Testing of pesticidal strains of *B. amyloliquefaciens* in models of human infection indicates a low pathogenic or toxic potential.

- It is used as a production micro-organism of enzymes (e.g. amylase, isoprene, protease, non-structural protein 3, ribonuclease, and phytases), biosurfactants, antibiotics and detergents which have industrial and commercial applications including cleaning, degreasing, and antibacterial applications. FSANZ has previously assessed *B. amyloliquefaciens* as a safe production organism for a number of food-grade enzymes (see Schedule 18[3] of the Code).

- It has been applied in a mixture with other bacterial species for water and waste water treatment to treat algal blooms, odours and sludge build-up.

[2] https://grdc.com.au/Media-Centre/Ground-Cover/Ground-Cover-Issue-108-Jan-Feb-2014/Canola-role-for-aquaculture
[3] https://www.legislation.gov.au/Series/F2015L00452

2.2.2 *Streptomyces hygroscopicus*

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

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).

2.2.3 Other organisms

Genetic elements from three other organisms not mentioned above (*Arabidopsis thaliana, Nicotiana tabacum* (tobacco); and *Agrobacterium tumefaciens*) have been used in the genetic modification of MS11 (refer to Table 1). These sequences are used to drive or terminate expression of the novel genetic material. None of the sources of these genetic elements is associated with toxic or allergenic responses in humans. The genetic elements derived from the plant pathogen *A. tumefaciens* are not pathogenic in themselves and do not cause pathogenic symptoms in MS11.

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

Studies submitted:

2008. Description of the MS11 transformation methodology. Unpublished Bayer Crop Science Report. Document no. M-307476-01-1.
2015. Description of vector pTCO113. Unpublished Bayer Crop Science report. Document no. M-182728-04-1.
2016. Detailed insert characterization and confirmation of the absence of vector backbone sequence in Brassica napus MS11. Unpublished Bayer Crop Science Report. Document no. M-547543-01-1.
2008. Full DNA sequence of event insert and integration site of *Brassica napus* transformation event MS11. Unpublished Bayer Crop Science Report. Document no. M-304805-01-1.
2016. Determination of additional flanking sequences and the corresponding insertion locus in *Brassica napus* MS11. Unpublished Bayer Crop Science Report. Document no. M-545355-01-1.
2016. MS11 *Brassica napus* - Identification of open reading frames (ORF) and homology search of sequences of more than 30 amino acids to known allergens and toxins. Unpublished Bayer Crop Science Report. Document no. M- 552421-01-1.
2016. Bioinformatics analysis of MS11 *Brassica napus* insertion locus. Unpublished Bayer Crop Science Report. Document no. M-307568-02-1.

2016. Structural stability analysis of *Brassica napus* MS11. Unpublished Bayer Crop Science Report. Document no. M-547544-01-1.
2016. MS11 *Brassica napus* - Inheritance of the insert over generations. Unpublished Bayer Crop Science report. Document no. M-545765-01-2.

3.1 Method used in the genetic modification

The transformation method was based on that of De Block et al. (1989). Embryogenic calli, initiated from hypocotyl segments of variety N90-740 seedlings germinated *in vitro*, were co-cultured with disarmed *A. tumefaciens* strain C58C1^{rif} (Van Larebeke et al. 1974) harbouring a binary vector system (Deblaere et al. 1987). The cloning vector, plasmid pTCO113 (Figure 1), contained the three genes of interest between right border (RB) and left border (LB) T-DNA sequences.

Following co-culture, the calli were then grown on medium containing carbenicillin, to inhibit the growth of excess *Agrobacterium*, and later transferred to a shoot induction medium containing glufosinate for selection of putative transformants. Healthy shoots were transferred to rooting medium and rooted shoots were transferred to soil in a glasshouse facility. MS11 was ultimately chosen as the lead event based on superior agronomic, phenotypic and molecular characteristics.

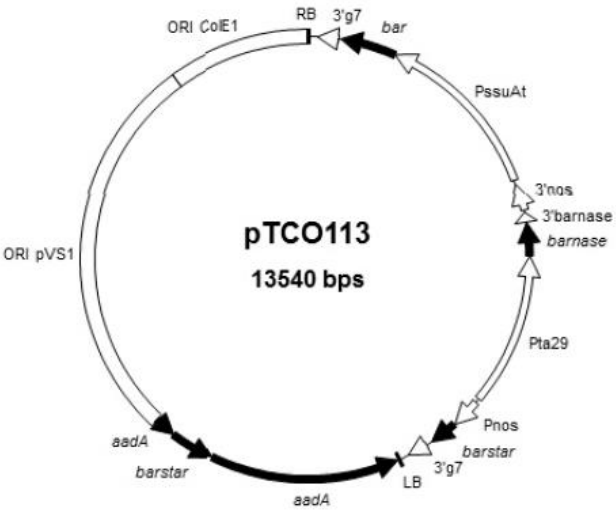

Figure 1: Genes and regulatory elements contained in plasmid pTCO113

3.2 Function and regulation of introduced gene fragments

Information on the genetic elements in the T-DNA used for transformation is summarised in Table 1. The complete plasmid is 13,540 bp comprising 7,675 bp vector backbone and 5,865 bp T-DNA (Figure 1). The T-DNA (Figure 2) comprises three cassettes located between a 25 bp LB and 25 bp RB. Intervening sequences, where present, have assisted with the cloning of the various components of each cassette.

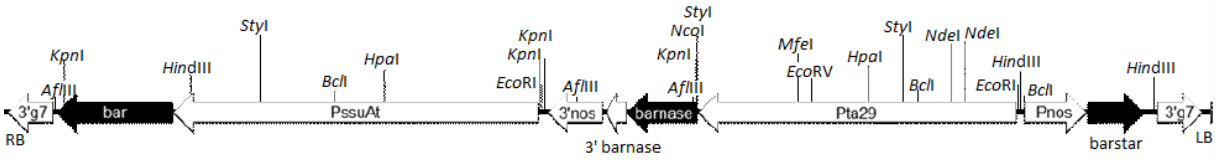

Figure 2: Design of pTCO113 T-DNA region (showing restriction sites and genetic elements detailed in Table 1)

Table 1: Description of the genetic elements contained in the T-DNA of pTCO113

Genetic element	Relative bp location on plasmid	Size (bp)	Source	Orient.	Description &Function	Reference
Right Border	1 - 25	25	[1]			
Intervening sequence	26 - 97	72				
***bar* cassette**						
3'g7	98-309	212	*Agrobacterium tumefaciens*	anti-clockwise	• Terminator region of the TL-DNA 7 gene of the octopine Ti plasmid • Directs polyadenylation of the *bar* gene	Dhaese et al (1983)
Intervening sequence	310 - 331	21				
bar	332 - 883	552	*S. hygroscopicus*	anti-clockwise	• Modified coding sequence of the *phosphinothricin acetyltransferase* gene	Thompson et al (1987)
SSUAt	884 - 2613	1730	*Arabidopsis thaliana*	anti-clockwise	• RUBISCO small sub-unit promoter region • Directs transcription of the *bar* gene	Krebbers et al (1988)
Intervening sequence	2614 - 2658	45				
barnase cassette						
3'nos	2659 - 2919	260	*Agrobacterium tumefaciens*	anti-clockwise	• Terminator region from the *nopaline synthase gene* • Directs polyadenylation of the *barnase* gene	Depicker et al (1982)
Intervening sequence	2920 - 2935	16				
3' barnase	2936 - 3033	98	*Bacillus amyloliquefaciens*	anti-clockwise	• Terminator region from the *barnase gene* • Directs polyadenylation of the *barnase* gene	Hartley (1988)
barnase	3034 - 3369	335	*Bacillus amyloliquefaciens*	anti-clockwise	• Coding sequence of the *barnase* gene	Hartley (1988)
Intervening sequence	3370 - 3371	2				
ta29	3372 - 4879	1508	*Nicotiana tabacum*	anti-clockwise	• ta29 promoter • Targets expression of *barnase* to the tapetum • Directs transcription of the *barnase* gene	Seurink et al (1990)
Intervening sequence	4880 - 4920	41				
***barstar* cassette**						
nos	4921 - 5214	294	*Agrobacterium tumefaciens*	clockwise	• Promoter region of the nopaline synthase gene • Directs transcription of the *barstar* gene	Depicker et al (1982)
Intervening sequence	5215 - 5216	2				
barstar	5217 - 5489	273	*Bacillus amyloliquefaciens*	clockwise	• Coding sequence of the *barstar* gene	Hartley (1988)
Intervening sequence	5490 - 5554	65				
3'g7'	5555 - 5766	212	*Agrobacterium tumefaciens*	clockwise	• Terminator region of the TL-DNA 7 gene of the octopine Ti plasmid • Directs polyadenylation of the *bar* gene	Dhaese et al (1983)
Intervening sequence	5767 - 5840	74				
Left Border	5841 - 5865	25				

3.2.1 *bar* cassette

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

Both the *bar* and *pat* genes, like other bacterial genes, have relatively high G:C content when compared to plant genes, and as a consequence the native microbial genes are inefficiently expressed in plants. In order to improve this expression, the codon usage pattern is modified but the resultant protein sequence is not altered (OECD 1999).

An *Nco*I site was created at the initiation codon of the *bar* gene and therefore the second codon – AGC (serine) has been modified to a GAC (aspartate) codon (Botterman et al. 1991).

Transcription of the *bar* gene is controlled by the promoter from the ribulose bisphosphate carboxylase (Rubisco) small sub-unit gene (from *A. thaliana*) which is active in all green tissue of the plant. This ensures there is strong transcription in the leaves and stems of the MS11 plant which would be the most exposed parts following spraying with glufosinate. Transcription termination of the *bar* gene is provided by the polyadenylation signal and 3' untranslated region (UTR) of the TL-DNA gene 7 (3'g7) of the *A. tumefaciens* octopine tumour-inducing (Ti) plasmid.

3.2.2 *barnase* cassette

The *barnase* gene is driven by the ta29 promoter from *N. tabacum* (tobacco) that restricts gene expression to the tapetum cells during anther development. The terminator sequence for the *barnase* gene is the 3' UTR of the *nopaline synthase (nos)* gene from the T-DNA of the Ti plasmid pTiT37 from *A. tumefaciens.*

3.2.3 *barstar* cassette

The *barstar* gene is controlled by the constitutive *nos* promoter from *A. tumefaciens*. The terminator for this gene is the same as used for the *bar* gene.

3.3 Breeding of MS11

Three breeding strategies (Figure 3) were followed for the development of the MS11 event and its introgression into various *B. napus* germplasm.

- T_0 hemizygous MS11 *B. napus* plants were cross-pollinated with non-genetically modified (non-GM) plants (N90-740 variety) to produce the T_1 generation. MS11 *B. napus* hemizygous plants from the T_1 generation were cross-pollinated with non-GM plants (N90-740 variety) to produce the T_2 generation. The process of crossing MS11 *B. napus* hemizygous plants with non-GM plants (N90-740 variety) was repeated to produce the T_3, T_4, and T_5 generations.

- MS11 *B. napus* hemizygous plants from the T_2 generation were cross-pollinated with non-GM plants (B144 variety) creating an F_1 generation. MS11 *B. napus* hemizygous plants from the F_1[*1] generation were backcrossed to non-GM plants (B144 variety) to produce a BC_1[*1] generation. The process of backcrossing MS11 *B. napus* hemizygous plants with non-GM B144 plants was repeated to produce the BC_2[*1], BC_3[*1], BC_4[*1], and BC_5[*1] generations.

- MS11 *B. napus* hemizygous plants from the T_2 generation were also cross-pollinated with non-GM plants (Ebony variety) creating a F_1^{*2} generation. MS11 *B. napus* hemizygous plants from the F_1 generation were backcrossed to non-GM plants (Ebony variety) to produce a BC_1^{*2} generation. The process of backcrossing MS11 *B. napus* hemizygous plants with non-GM plants (Ebony variety) was repeated to produce a BC_2^{*2} generation.

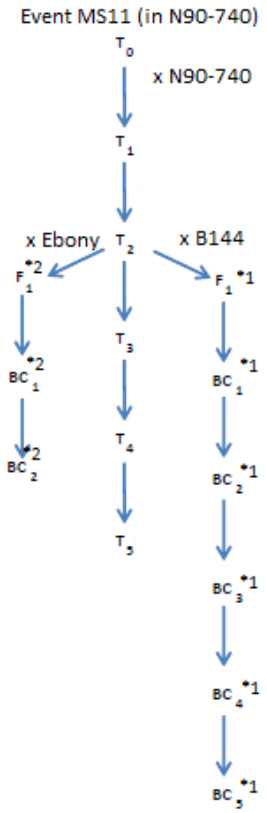

Figure 3: Breeding tree for MS11

The generations and controls used for various analyses described in this report are given in Table 2.

Table 2: MS11 generations used for various analyses

Analysis	MS11 generation(s) used	Control(s) used	Reference material
Molecular characterisation (Section 3.4.)	T_2, T_4	• Non-GM N90-740 • Non-GM N90-740 spiked with pTCO113	Plasmid pTCO113
Genetic stability (Section 3.5.1)	T_2, T_3, F_1^{*2}, BC_1^{*2}, BC_2^{*2}	• Non-GM N90-740 • Non-GM N90-740 spiked with pTCO113	N/A
Mendelian inheritance (Section 3.5.2)	T_3, T_4, T_5 BC_4^{*1}, and BC_5^{*1}	• N/A	N/A
Protein expression levels in plant parts (Section 4.1.3.2)	T_4	Non-GM N90-740	Microbially-produced PAT, Barnase and Barstar
Compositional analyses (Section 5)	T_4	Non-GM N90-740	Non-GM commercial lines 46A65; AC Elect; AC Excel; Peace; Spectrum; Westar.

3.4 Characterisation of the genetic modification in the plant

A range of analyses were undertaken to characterise the genetic modification in MS11. These analyses focussed on the nature of the insertion of the introduced genetic elements and whether any unintended re-arrangements may have occurred as a consequence of the transformation procedure.

3.4.1 Southern blot analysis: insert copy number, insert integrity and plasmid backbone

Genomic DNA isolation (Dellaporta et al. 1983) was performed from leaf samples from glasshouse-grown plants of generation T_2 (Figure 3 and Table 2) and from the unmodified parent (N90-740 negative control). Polymerase chain reaction (PCR) testing was used to confirm the identity of the samples.

Analysis of the DNA from both sources allowed determination of the number of insertion sites and the integrity of the inserted T-DNA, and testing for the presence or absence of plasmid vector backbone sequences by Southern blotting. For each characterisation, the DNA was first digested with appropriate restriction enzyme(s) before electrophoretic separation and visualisation following labelling with appropriate digoxigenin (DIG)-labelled probes. As a positive control, *Eco*RI -digested N90-740 DNA was spiked with *Eco*RI-digested reference plasmid (pTCO113) in an amount equivalent to equimolar (having previously ascertained that a $1/10^{th}$ equimolar concentration was sufficient enough to demonstrate the sensitivity of the method for detection of the target sequences).

3.4.1.1 Number of insertion sites and insert integrity

Pooled DNA from each test sample was digested with 11 restriction enzymes (sites shown in Figure 2*)* with recognition sites within the T-DNA. Following electrophoresis and Southern blotting, 10 probes were used to cover, between them, all elements of the T-DNA.

The N90-740 negative control DNA showed no hybridisation with any of the probes. For the positive control, two hybridising fragments of the predicted sizes were detected. For MS11 digests, fragments of predicted sizes were obtained. For four of the restriction enzyme digests of MS11, weak fragments additional to those predicted were obtained but could be explained, after further investigation, by incomplete digestion. The weight of evidence suggest the presence of a single insertion site in MS11 containing the *bar*, *barnase* and *barstar* cassettes as present in the pTCO113 plasmid.

3.4.1.2 Plasmid backbone

Pooled MS11 DNA from five plants (generation T_2) was digested with two restriction enzymes *Afl*III and *Nde*I, electrophoresed and probed with a) four DIG probes which together covered the entire plasmid backbone region and b) a single probe covering the T-DNA. Positive and negative controls were included.

The N90-740 positive control (digested with *Nde*I) spiked with *Eco*RI-digested pTCO113 gave the expected hybridisation fragments. MS11 incubated with the T-DNA probe showed the two expected fragments. No hybridisation fragments were detected in any other samples. This indicated that no backbone sequences are present in MS11.

As an additional check, the Applicant also tested for the presence of the *barstar* sequences present in the plasmid backbone (as opposed to those in the T-DNA). Using five primer combinations specific to the backbone *barstar* gene (plus/minus upstream and downstream sequences), as well as a primer pair targeting T-DNA sequences at the RB (serving as an internal control) PCR was done on MS11 DNA and DNA from negative and positive controls, and the products were electrophoresed and stained with ethidium bromide. No amplicons were obtained using negative controls or MS11 DNA as template with the five *barstar*-specific primers. The positive control (spiked with pTCO113) resulted in the expected fragments. This confirmed the absence of backbone *barstar* in MS11.

3.4.2 Insert sequence and insertion site

In an initial study, genomic DNA was extracted from leaf tissue obtained from glasshouse-grown, PCR-verified MS 11 plants (generation T_4). Six sets of oligonucleotide primers were used to amplify the insert and flanking regions as six individual overlapping fragments. The wild type locus was amplified in one fragment. The purified PCR fragments were shipped to the DNAVision[4] test site for sequence determination using a capillary genetic analyser. The obtained sequences were aligned with the pre-insertion locus and with plasmid pTCO113. All annotations and alignments were performed using Clone Manager[5] software.

In total, 6,753 bp of MS11 sequence were confirmed, containing 5,778 bp of intact sequence identical to sequence between the 3'g7 terminator of the *bar* cassette and the 3'g7 sequence of the *barstar* cassette of pTCO113 (see Table 1). This further confirmed that the insert consists of one copy of the three cassettes. At the 5' and 3' ends past this intact sequence, 419 bp and 556 bp respectively were sequenced. It was however, not possible to ascertain the origin/identity of the most proximal of these sequences since there were spans of sequence that were identical to both potential inserted sequences and potential pre-insertion site sequences i.e. it was not possible to say with certainty whether any RB or LB border sequences had been incorporated and how much of the intervening sequences between the LB and RB and the 5,778 bp insert had been incorporated. It is noted that the truncation of the border sequences is not uncommon for *Agrobacterium*-mediated transformation events (Tzfira et al. 2004; Kim et al. 2007) and therefore may have occurred in MS11.

Alignment of the MS11 sequence with the pre-insertion locus indicated that there was a 1 bp addition and a 40 bp target site deletion. A representation of the final transgenic locus in MS11 is given in Figure 4.

[4] DNAVision - http://www.dnavision.com/
[5] Sci-Ed Clone Manager software – http://www.scied.com/pr_cmpro.htm

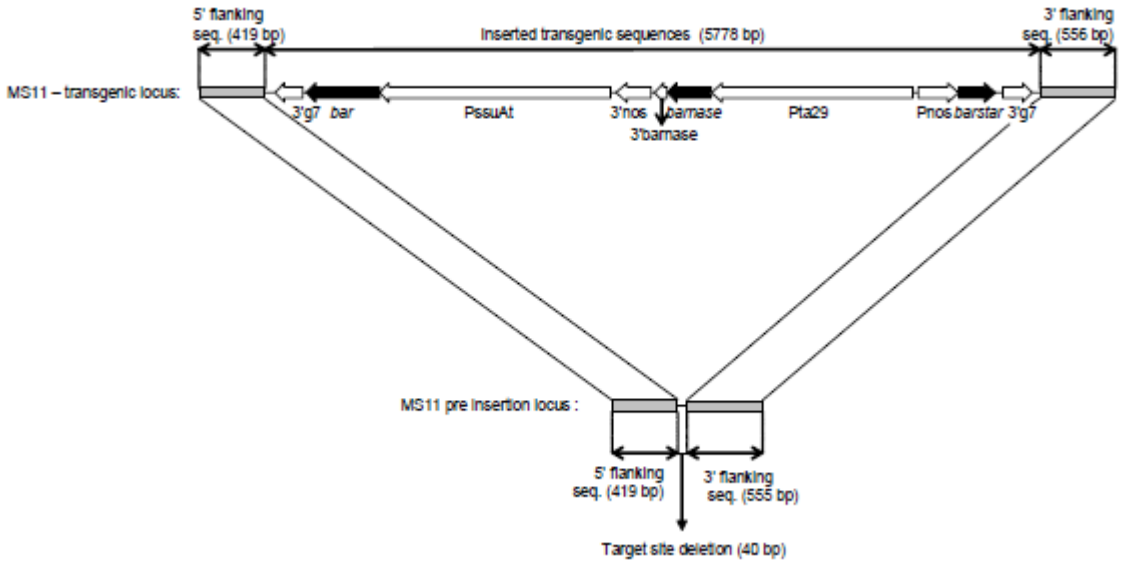

Figure 4: Representation of the map position of the T-DNA insert in the N90-740 genome

An additional study was undertaken to sequence at least 1Kb of each of the 5' and 3' flanking regions and obtain more information about the insertion locus. Genomic DNA was extracted from leaf tissue obtained from glasshouse-grown, PCR-verified MS 11 plants (generation T_2) and from verified non-GM N90-740 plants. Additional primers to those described previously were used to generate more fragments and Sanger sequencing was performed using the BigDye® Terminator Cycle Sequencing Kit[6]. The consensus sequences of the amplified fragments of the transgenic and insertion locus were analysed through Clone Manager software. The extended sequences were merged with the previously obtained sequences and validated by comparing with the consensus sequences of the amplified insertion locus fragment and with the original MS11 insertion locus sequence. Finally the extended transgenic sequence was compared with the extended insertion locus sequence.

The final MS11 transgenic locus consisted of 8209 bp, which included 1129 bp of 5' flanking sequence, 1302 bp of 3' flanking sequence and the same 5,778 bp of sequence from pTCO113 as previously identified. The flanking sequences were shown to correspond to sequences in the insertion locus of the N90-740 host and therefore to be of *B. napus* origin.

Bioinformatics analyses of the insertion locus was undertaken in order to ascertain a) whether there had been any disruption of endogenous genes as a result of the transformation procedure and b) the location of the insert in the genome. The full 8209 bp sequence was used as the query sequence.

A BLAST[7] (Basic Local Alignment Search Tool) analysis (Altschul et al. 1990), available on the NCBI (National Center for Biotechnology Information[8]) website, was used to compare the query sequence (or its translation) with sequences of known genes/proteins/expressed sequence tags. This search indicated the presence of one endogenous gene in the 3' flanking region of the insert but this was not interrupted. Sequence similarities to a number of coding sequences associated with genes from other species were also found in the 3' flanking region but, again, none of these sequences were disrupted.

[6] BigDye Sequencing Kit https://www.thermofisher.com/order/catalog/product/4337455
[7] BLAST - https://blast.ncbi.nlm.nih.gov/Blast.cgi
[8] NCBI - https://www.ncbi.nlm.nih.gov/

Similarities between the MS11 insertion locus and sequences in the unmodified host were identified using BLAT[9] (BLAST-like alignment tool) and a *B. napus* reference genome (Chalhoub et al. 2014) available on the Genoscope[10] website. The analysis indicated that the insert has gone into chromosome A03[11].

3.4.4 Open reading frame (ORF) analysis

The MS11 transgenic locus, containing the inserted DNA together with the additional 5' and 3' flanking sequences was used as query sequence. An ORF was defined as the region between two translation stop codons (TAA, TAG, or TGA) with a minimum size coding for 3 amino acids. The ORF search was performed using the GetORF search program from the European Molecular Biology Open Software Suite[12] (EMBOSS) tools (version 6.3.1, July 2010).

The search identified 554 ORFs (corresponding to 526 unique sequences). After elimination of duplicates, translated amino acid sequences of at least 30 amino acids length represented 107 unique sequences; ORFs shorter than 29 amino acids were not evaluated since a minimum 35% identity requires at least a match of 29 amino acids over 80 amino acids. The 35% identity is a recommended criterion for indicating potential allergenicity (Codex 2009). These 107 putative polypeptides were analysed using a bioinformatics strategy to determine similarity to known protein toxins or allergens (refer to Section 4.1.6)

3.5 Stability of the genetic change in MS11

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

3.5.1 Genetic stability

The genetic stability of event MS11 was evaluated by Southern blot analysis of genomic leaf DNA from verified plants from generations T_2, T_3, F_1*2, BC_1*2, and BC_2*2 (see Figure 3). Restriction enzyme digestion, with *Eco*RV, of samples from each generation were separated by agarose gel electrophoresis and blots were probed with the same DIG-labelled probe, specific to the whole T-DNA, as used in the plasmid backbone analysis (Section 3.4.1.2).). A negative control of DNA from non-GM N90-740 digested with *Eco*RV and a positive control of non-GM N90-740 digested with *Eco*R1 spiked with equimolar pTCO113 digested with *Eco*R1 were also included.

No hybridising bands were obtained for the negative control; the expected bands were obtained for the positive control (same sizes as obtained in all of the positive control Southern blots reported in Section 3.4.1.1). The same two hybridising fragments (same sizes as those obtained for the backbone analysis) were detected across all MS11 samples from the five generations thereby confirming stable integration of the inserted DNA in MS11.

[9] BLAT (BLAST-like alignment tool) is a pairwise sequence alignment algorithm that is designed to quickly find sequences with at least 95% similarity and with a length of at least 40 bases.
[10] Genoscope Centre National de Sequençage - http://www.genoscope.cns.fr/spip/spip.php?lang=en
[11] *B. napus* is an amphidiploid derived from a hybridisation event between *B. rapa* (A. genome, with 10 chromosomes) and *B. oleracea* (C genome, with 9 chromosomes).
[12] EMBOSS - http://emboss.sourceforge.net/

3.5.2 Phenotypic stability

Using PCR analysis, genomic DNA from leaf samples of MS11 generations T_3, T_4, T_5 BC_4*[1], and BC_5*[1] was tested for the absence or presence of the insert. According to Mendelian inheritance principles, MS11 with a single insert would be predicted to segregate at a 1:1 (positive:negative) ratio for progeny derived from non-GM and hemizygous parental plants.

Two types of PCR analysis were employed – a) event-specific analysis to determine the presence/absence of the insert and b) gene-specific analysis to determine the presence/absence of the *bar*, *barnase* and *barstar* genes.

The results of presence/absence of the insert are shown in Table 3. The chi-square (X^2) critical value at significance level α = 0.05 is 3.84 i.e. if the X^2 value is < 3.84 the observed ratio is not significantly different from the expected ratio. The X^2 values for all generations were less than 3.84 thereby indicating that the insert in each genetic background was inherited according to Mendelian principles and supporting the conclusion the MS11 event consists of a single insert integrated at a single chromosomal locus within the *B. napus* nuclear genome.

Table 3: Segregation of the T-DNA insert in MS11 over five generations

Generation/Background	Total plants	Ratio presence:absence		X^2	Probability[1]
		Observed	Expected		
T_3	84	42:42	42:42	0.000	NS
T_4	92	48:44	46:46	0.174	NS
T_5	95	39:56	47.5:47.5	3.042	NS
BC_4*[1]	89	43:46	44.5:44.5	0.101	NS
BC_5*[1]	98	51:47	49:49	0.163	NS

[1]NS = not significant -X^2<3.84 (5% confidence level)

The results from the gene-specific PCR analysis confirmed that the *bar*, *barstar*, and *barnase* genes are present for samples positive for the MS11 event and are absent for samples negative for the MS11event.

3.6 Conclusion

MS11 was generated through *Agrobacterium*-mediated transformation with a single T-DNA containing three expression cassettes. Comprehensive molecular analyses indicate there is a single insertion site, in chromosome A03, comprising a single, complete copy of each of the *bar, barnase* and *barstar* genes with their regulatory elements. The introduced genes are stably inherited from one generation to the next. No plasmid backbone has been incorporated into the transgenic locus and no endogenous genes have been disrupted as a result of the transformation process.

4 Characterisation and safety assessment of new substances

The main purpose of the characterisation is to describe the nature of any new substances and their phenotypic and biochemical effects on the organism in which they are expressed, particularly in the parts of the organism consumed as food. Typically, the main focus of the characterisation is on newly expressed (or potentially expressed) proteins, but other (non-protein) substances may need to be considered.

4.1 Newly expressed proteins

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

Two types of proteins were considered:

- The proteins expected to be directly produced as a result of the translation of the introduced genes. A number of different analyses were done to characterise these proteins and determine their *in planta* expression.

- Those that may be potentially translated as a result of the creation of ORFs during the transformation process (see Section 3.4.4).

4.1.1 PAT protein

Study submitted

2003. Description of the amino acid sequence of the PAT protein encoded from the *bar* gene. Unpublished Bayer Crop Science Report. Document no.M-084188-01-2.

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

The homologous polypeptide produced by both the *bar* and *pat* (from *S. viridochromogenes*) genes (see Section 3.2.1) is known as phosphinothricin acetyltransferase (PAT); it is an acetyl transferase with enzyme specificity for both L-PPT and demethylphosphinothricin (DMPT) in the acetylation reaction (Thompson et al., 1987). In the presence of acetyl-Coenzyme A (CoA), PAT catalyses the acetylation of the free amino group of L-PPT to N-acetyl-L-PPT, a herbicidally-inactive compound. The kinetics and substrate specificity of the PAT enzyme are well characterised; it has a high specificity for L-PPT and has been shown to have a very low affinity to related compounds and amino acids; even excess glutamate is unable to block the L-PPT-acetyltransferase reaction (Thompson et al. 1987). The proteins from the two different sources have a sequence identity of 85%

A translation of the DNA sequence of the *bar* gene in the MS11 insert gives a protein comprising 183 amino acids with a calculated molecular weight of 20.67 kDa. As noted in Section 3.2.1, the second codon – AGC (serine) has been modified to a GAC (aspartate) codon (Botterman et al. 1991).

4.1.2 Barnase and Barstar proteins

Studies submitted:

2004. The *barnase* and *barstar* gene products: Barnase and Barstar Description and Characterization. Unpublished Bayer CropScience Report. M-226785-01-1.2015.
2009. Barnase and barstar proteins - History of safe use. Unpublished Bayer Crop Science Report. Document no. M-355152-01-1.
2003. Description of the amino acid sequence of the Barnase protein. Unpublished Bayer Crop Science Report. Document no. M-232685-01-1.
2003. Description of the amino acid sequence of the Barstar protein. Unpublished Bayer Crop Science Report. Document no. M-232692-01-1.

A brief description of the function and effects of these two proteins has been given in Section 1. In the native host (*B. amyloliquefaciens*), Barnase is produced within the cell and then secreted to the outside where it is probably used to allow scavenging of nucleotides from RNAs found in the environment (Condon and Putzer 2002). In order to prevent its own RNA from being destroyed before the Barnase has been secreted, the host produces the inhibitor Barstar. One molecule of Barstar binds tightly to one molecule of Barnase thereby abolishing its ribonuclease activity (Hartley and Smeaton 1973). Barstar does this by mimicking the RNA substrate at the phosphate binding site of Barnase (Guillet et al. 1993).

Barnase and Barstar have been utilised in direct agronomic application to allow male sterility and restoration of fertility (see discussion in Section 1) and, for example, hybrid GM canola varieties engineered with this technology have been commercialised since 1996. In the fertility restorer line (RF3) that is intended to be crossed with MS11 (FSANZ 2002), the *barstar* gene has the same tapetum-specific promoter (ta29) as the *barnase* gene in MS11 thereby ensuring expression of the two proteins (and formation of a Barnase/Barstar complex) in the tapetum cells of the hybrid.

In MS11, Barstar has been used to inhibit inadvertent Barnase expression in tissues other than tapetal cells (see discussion in Section 1).

A translation of the DNA sequence of the *barnase* gene in the MS11 insert gives a protein comprising 111 amino acids with a calculated molecular weight of 12.51 kDa. The amino acid sequence of the plant-produced protein differs from the native barnase protein isolated from *Bacillus amyloliquefaciens* by the addition of a methionine (Asp718 site) in the start codon followed by substituting alanine and glutamine for valine and proline respectively.

A translation of the DNA sequence of the *barstar* gene in the MS11 insert gives a protein comprising 90 amino acids with a calculated molecular weight of 10.34 kDa.

4.1.3 Protein expression in the tissues of MS11

Since MS11 is not intended to be a stand-alone line from which food will be produced but will be crossed with the GM fertility restorer line RF3 (see Section 1) the Applicant carried out protein expression analyses on MS11, RF3, and MS11 x RF3 plants.

Studies submitted:

2015. MS11 *Brassica napus* - Summary of protein expression analyses of field samples grown in Canada and the USA during 2014. Unpublished Bayer Crop Science Report. Document no. M-549123-01-1.
2015. MS11 x RF3, MS11, and RF3 *Brassica napus* - Protein expression analyses of field samples grown in Canada and the USA during 2014. Unpublished Bayer Crop Science Report. Document no. M-542702-01-1.

Plants from PCR-verified seed of MS11 (generation T_4 – see Figure 3), RF3 and MS11 x RF3, together with the non-GM N90-740 parent were sampled from three locations across Canada and the USA[13] during the 2014 season.

Five blocks (replicates) of each entry (sample type) were established at each test site in a randomised complete block design. MS11 plants were either unsprayed or sprayed with Liberty® 280SL containing approximately 280 g ai/L glufosinate ammonium applied at the 2 – 4 leaf growth stage (BBCH12-14)[14]. Given that the T_4 plants had been obtained by crossing MS11 with the non-GM N90-740 (see Section 3.3) the resulting plants contained a 1:1 mix of positive:negative segregants. This meant that in the sprayed treatment, approximately half of the plants were killed by the glufosinate; to allow for this, planting density of these plants was double that of the unsprayed plants. Each entry was planted in 6-row plots of 5 metres in length.

Levels of each protein were determined for each sample type using a validated enzyme linked immunosorbent assay (ELISA) employing a commercially prepared kit as follows:

- Qualiplate™ Kit for LibertyLink PAT/*bar* (EnviroLogix)
- Barnase Plate Kit (EnviroLogix)
- Barstar Plate Kit (EnviroLogix)

A standard curve was prepared for each ELISA, and sample values were interpolated from this. SoftMax Pro® was used to calculate the concentration (ng/ml extract) of the proteins. These values were then converted to amounts per weight of tissue homogenised for each sample type and analyte, and ultimately expressed as µg protein per g dry weight (dw) tissue. Where protein levels in samples were close to the Limit of Quantification (LOQ) it was not possible to distinguish between positive segregants with low expression levels and negative segregants; these samples were excluded from calculation of the mean.

The results for each protein as expressed in MS 11 and MS11 x RF3 are presented in Table 4 (RF3 results not shown) and the following observations made:

- While there were differences in the means between sprayed and unsprayed plants these differences were not consistent either between plant parts or between MS11 and MS11 x RF3.
- In general, the means for all three proteins across all plant parts were slightly higher in the MS11 x RF3 plants than in the MS11 plants.
- For each protein, the trends of the means across sprayed/unsprayed and MS11/MS11 x RF3 were consistent i.e.:
  - The mean level of PAT protein was highest in whole plants sampled at the 3 – 5 leaf stage (BBCH13 – 15). Lowest values were obtained in root and seed samples.
  - The Barnase protein was below the LOQ in all samples except roots in the inflorescence development/flowering stage (BBCH57-65) in MS11 x RF3.
  - Levels of Barstar in all tissues were either below the LOQ or very low. Roots marginally appeared to have the highest mean level.
- Mean levels of all three proteins in the edible part (i.e. seed) were below the LOQ or very low.

[13] Norfolk Treherne – Manitoba, Canada; La Vallée-du-Richelieu – Quebec, Canada; Grant, Washington, USA.
[14] The canola growth stages are based on the Bayer, BASF, Ciba-Geigy and Hoechst (BBCH) Growth Stage Scale see e.g. CCC (2014)

Table 4: PAT, Barnase & Barstar protein content in various tissues of MS11 and MS11 x RF3, averaged across three sites

BBCH13-15 = 3 – 5 leaf stage; BBCH30-39 = stem elongation: BBCH57 – 65 = inflorescence emergence to flowering; ND = not determine

	Growth stage/plant part	LOQ (µg/g)	LOQ µg/g dw	MS11				MS11 x RF3			
				Unsprayed (µg/g dw)		Sprayed (µg/g dw)		Unsprayed (µg/g dw)		Sprayed (µg/g dw)	
				Mean	Range	Mean	Range	Mean	Range	Mean	Range
PAT	BBCH13-15/whole plant	0.05	0.05	22.02	14.9 – 30.8	35.4	7.32 – 74.4	41.16	16.03 – 75.9	52.09	20.42 – 83.02
PAT	BBCH30-39/whole plant	0.05	0.05	24.6	9.5 – 40.7	21.8	7.3 – 40.6	22.7	8.2 – 61.8	25.2	12.9 – 37.6
PAT	BBCH30-39/root	0.02	0.02	0.17	0.15 – 0.2	0.39	0.18 – 0.64	0.97	0.4 – 2.3	2.35	0.56 – 19.5
PAT	BBCH57-65/whole plant	0.05	0.05	18.9	7.9 – 24.5	14.8	6.1 – 27.5	22.5	8.5 – 42.1	28.0	12.3 – 59.3
PAT	BBCH57-65/root	0.02	0.02	0.17	<LOQ – 0.1	0.37	0.1 – 0.7	0.74	0.2 – 1.3	0.75	0.3 – 1.4
PAT	BBCH57-65/raceme	0.05	0.05	13.9	12.5 – 16.0	23.9	9.3 – 55.2	30.7	18.1 – 75.6	41.9	17.0 – 108.1
PAT	Maturity/seed	0.05	0.05	0.3	0.06 – 0.5	0.4	0.3 – 0.8	0.6	0.2 – 1.0	0.6	0.2 – 1.1
Barnase	BBCH13-15/whole plant	0.03	0.03	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barnase	BBCH30-39/whole plant	0.05	0.05	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barnase	BBCH30-39/root	0.25	0.25	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barnase	BBCH57-65/whole plant	0.05	0.05	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barnase	BBCH57-65/root	0.25	0.25	ND	<LOQ	ND	<LOQ	2.18	1.62 – 2.74	4.74	<LOQ – 4.74
Barnase	BBCH57-65/raceme	0.04	0.04	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barnase	Maturity/seed	0.05	0.05	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barstar	BBCH13-15/whole plant	0.05	0.05	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barstar	BBCH30-39/whole plant	0.03	0.03	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ
Barstar	BBCH30-39/root	0.05	0.05	0.4	0.2 – 1.0	0.5	0.2 – 1.0	0.5	0.4 – 0.6	0.7	0.3 – 0.9
Barstar	BBCH57-65/whole plant	0.03	0.03	ND	<LOQ	0.2	0.1 – 0.2	0.4	0.2 – 0.4	0.3	0.2 – 0.4
Barstar	BBCH57-65/root	0.05	0.05	0.4	0.3 – 0.4	0.3	0.2 – 0.5	0.3	0.1 – 0.6	0.4	0.2 – 0.8
Barstar	BBCH57-65/raceme	0.03	0.03	ND	<LOQ	0.6	0.4 – 0.9	0.5	0.2 – 1.7	0.7	0.2 – 3.6
Barstar	Maturity/seed	0.03	0.03	ND	<LOQ	ND	<LOQ	ND	<LOQ	ND	<LOQ

4.1.4 Characterisation of the proteins produced in MS11

In many cases, insufficient amounts of newly-expressed proteins can be obtained from the plant for safety evaluations (e.g. pepsin digestibility and heat stability studies, acute oral toxicity testing and use as a calibration standard for quantitative ELISA). A standard practice in such cases is to produce the proteins for these evaluations in a microbial system and directly confirm their equivalence to the plant-produced proteins in simultaneous analyses. Hence, individual studies will concurrently determine both the structural and functional characteristics of the plant-purified protein as well as its similarity to the equivalent microbially-produced protein.

4.1.4.1 PAT

Study submitted

2015. Characterization of plant produced PAT/*bar* protein purified from MS11 *Brassica napus* plants (batch 1520_PATbar(MS11)) and comparability with the recombinant protein batch 1215_PATbar. Unpublished Bayer CropScience Report. Document no. M-544805-01-1.

PAT protein was extracted from leaves of MS11 and purified by affinity chromatography (affinity resin contained rabbit polyclonal antibody raised against PAT). A negative control (if used) was prepared from non-GM N90-740. Depending on the test system, protein purified from *Escherichia coli* was characterised either as the stand-alone protein or as a spike to the non-GM N90-740 control.

A number of techniques were employed for analysis of the proteins from both sources as follows:

- **Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot analysis**
  The apparent molecular mass of the proteins was compared by SDS-PAGE and staining with Coomassie Blue. After electrophoresis, proteins were transferred to polyvinylidene fluoride (PVDF) membranes via electro-blotting and immuno-reactivity was assessed by incubation with a primary polyclonal rabbit anti-PAT antibody followed by goat anti-rabbit alkaline phosphate (AP)-conjugated secondary antibody.

- **Ultra Performance Liquid Chromatography** (UPLC) coupled with ultra violet (UV) detection and mass spectrometry (MS) was used for intact molecular weight determination, peptide mapping (of tryptic digests) and N-terminal sequencing (of tryptic digests) of the plant-derived protein. UPLC is a recent technique that builds on the principles of High Performance Liquid Chromatography (HPLC) to provide improved resolution and sensitivity in peptide analyses (Chawla and Ranjan 2016).

- **Glycostaining**. This technique tests for proteins that have been post-translationally modified by the addition of carbohydrate moieties (glycans) covalently linked to the polypeptide backbone. Glycosylation patterns have been implicated in contributing to possible allergenicity (Huby et al. 2000) since they may affect the susceptibility of a protein to processing and proteolysis and may introduce glycan peptides which are known to be highly cross-reactive epitopes. There has been the suggestion that transgenic proteins may have different glycosylation patterns from their native counterparts; while this had apparently been demonstrated in a transgenic pea (Prescott et al. 2005) it was later refuted (Lee et al. 2013).

Proteins produced in prokaryotes are not expected to be glycosylated and only a few specific endogenous proteins in *E. coli* have been confirmed to be glycosylated (Sherlock et al. 2006)

The analysis was done by using a commercial kit (Sigma® Glycoprotein Detection Kit) following SDS-PAGE. The kit is designed to selectively stain glycoproteins on a nitrocellulose membrane using a modification of the Periodic Acid-Schiff method. Staining of sugar moieties of glycoproteins yields magenta bands with a colorless background. The Schiff reagent stains vicinal diol groups found mainly on peripheral sugars and sialic acids and is used as a general glycoprotein stain.

- **Enzymatic activity**. The PAT activity assay is based on the ability of the PAT enzyme to acetylate PPT in the presence of acetyl CoA (see Section 4.1.1). The resulting reduced CoA reacts with 5,5-dithiobis-2-nitrobenzoic acid to yield a molar equivalent of free 5-thio-2-nitrobenzoic acid which can be quantified spectrophotometrically by its ability to absorb light at 412 nm. The difference of the CoA concentration at the end of the reaction and the initial amount of CoA at that start of the reaction is divided by the reaction time (15 minutes) to give the rate of reaction (µM/min or µmol*L^{-1}*min^{-1}). The rate of the reaction is then converted into specific activity, which is defined as the amount of substrate converted per unit of time per mg of enzyme (µmol*min^{-1}*mg^{-1} or U/mg).

SDS-PAGE and western blot

In the SDS gel, an intense band was observed at approx. 21 kDa for the MS11 and *E. coli* samples, demonstrating their equivalence as well as being in agreement with the calculated molecular weight of 20.67 kDa (Section 4.1.1). Similarly, the western blot showed an intense immuno-reactive band at the same approximate molecular weight that co-migrated in the MS11 and *E.coli* samples and increased in intensity with protein load. No band was observed for the negative control in either SDS-PAGE or western blots.

UPLC-UV-MS

The chemical average mass of the plant-derived PAT was measured to be 20.7 kDa, close to the calculated molecular weight of 20.67 kDa. Results from peptide mapping of a trypsin digest covered 100% of the theoretical amino acid sequence and therefore confirmed the identity of the plant-derived protein. The amino acid analyses from the N-terminus yielded a 5-amino acid sequence which matched the expected MDPER sequence.

Glycosylation analysis

A dark band was obtained for the positive control (horseradish peroxidase) while the PAT proteins from both plant and microbial sources and the negative control gave no dark bands. These results support the conclusion that neither microbially- nor plant-derived PAT proteins are glycosylated.

Enzymatic activity

The mean observed specific activity for the protein purified from *E. coli* was 18.727 U/mg and for the plant-purified protein was 13.894 U/mg. Therefore, both proteins were considered functionally equivalent.

Conclusion

These studies confirmed both the structural and functional characteristics of the plant-produced PAT protein and demonstrated equivalence between the plant- and microbially-derived PAT protein. Hence the *E.coli*-derived protein is a suitable surrogate for MS11-produced protein in safety assessment studies.

4.1.4.2 Barnase and Barstar

Studies submitted

2013. Characterization of the recombinant Barnase protein batch no 1205_Barnase. Unpublished Bayer Crop Science Report. Document no. M-467079- 01-1.
2016. Characterization of the recombinant Barnase protein batch 1518_Barnase. Unpublished Bayer Crop Science Report. Document no. M-551100-01-1.
2009. Certificate of analysis for the Barstar protein produced in *E.coli* batch no LB300909B. Unpublished Bayer Crop Science Report. Document no. M-433234-01-1.
2012. Supplementary characterization of the Barstar protein batch no LB300909B* produced in *Escherichia coli*. Unpublished Bayer Crop Science Report. Document no. M-433174-01-1.
2014. Characterization of the recombinant barstar protein batch 1340_Barstar. Unpublished Bayer Crop Science Report. Document no. M-495269- 01-1.
2012. Characterization of the barstar protein batch no MVW020911B produced in *Escherichia coli*. Unpublished Bayer Crop Science Report. Document no. M-433233-01-1.
2016. Characterization of the recombinant Barstar protein batch 1522_Barstar. Unpublished Bayer Crop Science Report. Document no. M-548907-01-1.
2016. Comparability of the Barstar protein from MS11 *Brassica napus* plants with the recombinant protein batch 1340_Barstar. Unpublished Bayer Crop Science Report. Document no. M-548891-01-1.

Results from the protein expression levels (Section 4.1.3) clearly show that neither the Barnase nor Barstar protein is expressed in significant enough amounts in MS11 to be used in safety evaluations. However, even repeated attempts by the Applicant to purify or further enrich either the Barnase or Barstar protein from MS11 (for purposes of characterisation), using immuno-affinity chromatography, were unsuccessful, as contaminants were also co-purified during this process. As such, both proteins in MS11 are classified as intractable (Bushey et al. 2014).

Therefore, a weight of evidence approach had to be used to assess the equivalence of the intractable plant proteins with the microbially-produced proteins and hence to indirectly characterise the plant-produced proteins. It is noted, however, that the sequence analysis of the MS11 insert confirmed the expected sequences of both the *barnase* and *barstar* genes. From these, the expected protein sequence was translated (see Section 4.1.2). It is also noted that, a) in the case of the Barnase protein, MS11 showed the expected male sterile phenotype and hence that a functional Barnase protein had been expressed and b) that, in the case of Barstar, a comparison of the use of the pTCO113 plasmid in transformation with an otherwise identical plasmid lacking the *nos-barstar* gene gave a higher percentage of good performing male-sterile canola plants (Michiels et al. 2000) i.e. that the Barstar expression brought about the desired phenotype and hence is a functional protein.

In addition, since Barstar was expressed at a low level in root tissue of MS11(see Table 4), it was possible to obtain a total protein extract from MS11 roots, that could be used in a western blot analysis for Barstar and compared directly with an extract taken from roots of non-GM N90-740 spiked with Barstar protein purified from *E. coli*. This analysis showed a clear unique band at approximately 10 kDa (which compared with the calculated molecular weight of 10.34 kDa) for samples from both sources.

In terms of the *E.coli*-produced proteins, the following analyses were done (performed on two separate batches of Barnase and four separate batches of Barstar):

- **SDS-PAGE and western blot** The antibodies used for the western blots were:

  - For Barnase – rabbit anti-Barnase primary antibody followed by goat anti-rabbit AP-conjugated or goat anti-rabbit horseradish peroxidase(HRP)-conjugated secondary antibody
  - For Barstar - rabbit anti-Barstar primary antibody followed by goat anti-rabbit AP-conjugated secondary antibody.

- Liquid Chromatography-UV-MS (**LC-UV-MS) or UPLC-UV-MS** were used for chemical average mass determination and peptide mapping of an endoproteinase Glu-C digest of each protein. Both automated **Edman degradation** and UPLC-UV-MS were used to determine as much of the N-terminal sequence as possible.

- **Glycosylation analysis** – using the same system as described in Section 4.1.4.1.

- **Enzymatic activity** The activity of the Barnase and Barstar proteins was assessed using the RNase Alert Lab Test Kit (Applied Biosystems). The fluorescent substrate used in this kit is a modified RNA oligonucleotide that emits green fluorescence when being cleaved by an RNase such as Barnase. If only Barnase is added to the substrate, a high fluorescence signal is measured. If additionally the specific Barnase inhibitor Barstar is added to the substrate, the Barnase forms a complex with Barstar and is not able to cleave the fluorescent substrate which therefore yields a lower fluorescence signal, showing that the RNase activity measured in the assay is specifically due to Barnase activity.

  For Barnase, the rate of formation of the fluorescent product, monitored over time, is converted to specific activity (µmol of product formed per min per mg of Barnase).

  For Barstar, the rates of Barnase titrated with different concentrations of Barstar is plotted against the concentration of Barstar, and this is then used to calculate an IC_{50} value. The IC_{50} is the half maximal inhibitory concentration; it is a measure of the effectiveness of a substance (Barstar) in inhibiting a specific biological biochemical function (Barnase activity). When comparing IC_{50} values, lower IC_{50} indicates higher activity of the inhibitor.

The results of the various analyses can be summarised as follows:

- The protein batches for Barnase all gave a single band at approximately 12 kDa for both SDS-PAGE and western blots. This was in agreement with the calculated molecular weight of 12.51 kDa.

- The protein batches for Barstar all gave an intense band at approximately 9-10 kDa for both SDS-PAGE and western blots. This was in agreement with the calculated molecular weight of 10.34 kDa. A less intense band at approximately 20 kDa is considered to be a dimer.

- No glycosylation signal was observed for either protein.

- A 100% coverage of both proteins was obtained from peptide mapping of the digest of each protein and gave the expected sequence. N-terminal sequencing further confirmed the identity of each protein. Intact molecular mass determination gave a chemical average mass of 12.51 kDa for Barnase and 10.34 kDa for Barstar.

- Under the conditions of the various studies used to generate the information, the RNase activity of Barnase was shown to be reduced by 40% in the presence of Barstar. This confirmed the activity of both proteins.

Conclusion

Due to the low levels of Barnase and Barstar in MS11, purified proteins of sufficient quantity and quality could not be extracted from the MS11 plant to be able to directly characterise them. However the weight of evidence, provided by a) translation of the known DNA sequences of the three introduced genes present in MS11, b) the fact that the proteins function as predicted in the plant and c) the detailed characterisation of equivalent microbially-produced proteins, is sufficient to confirm the identity of the proteins expressed in MS11.The *E.coli*-derived proteins are considered to be suitable surrogates for the MS11-produced proteins in safety assessment studies.

4.1.5 Safety of the introduced proteins

4.1.5.1 Potential toxicity of PAT, Barnase and Barstar

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

The assessment focuses on:

- whether the novel proteins have a prior history of safe human consumption, or are sufficiently similar to proteins that have been safely consumed in food;
- amino acid sequence similarity with known protein toxins and anti-nutrients;
- structural properties of the novel proteins including whether they are resistant to heat or processing.

An oral toxicity study is only deemed necessary if the results of biochemical, bioinformatic, digestibility or stability studies indicate further investigation of potential toxicity is warranted.

Studies submitted

Bioinformatics

2016. PAT/*bar* protein - Amino acid sequence homology search with known allergens and known toxins. Unpublished Bayer Crop Science Report. Document no. M-084359-10-1.

2016. Barnase protein - Amino acid sequence homology search with known allergens and known toxins. Unpublished Bayer Crop Science Report. Document no. M-552256-01-1.

2016. Barstar protein - Amino acid sequence homology search with known allergens and known toxins. Unpublished Bayer Crop Science Report. Document no. M-552416-01-1.

Thermolability

2016. The effect of temperature on PAT/*bar* as assessed by the PAT quantitative activity assay. Unpublished Bayer Crop Science Report. Document no. M-554703-01-1.

2014. The effect of temperature on microbially produced Barnase assessed by the barnase quantitative activity assay. Unpublished Bayer Crop Science Report. Document no. M-490632-01-1.
2012. The heat stability of microbially produced Barnase assessed by SDS-PAGE and western blot analyses. Unpublished Bayer Crop Science Report. Document no. M-440532-01-1.
2013. The effect of temperature on microbially produced Barnase assessed by ELISA. Unpublished Bayer Crop Science Report. Document no. M-475710-01-1.
2014. The effect of temperature on microbially-produced Barstar assessed by the Barstar quantitative activity assay. Unpublished Bayer Crop Science Report. Document no. M-490635-01-1.
2012. The heat stability of microbially produced Barstar assessed by SDS-PAGE and western blot analyses. Unpublished Bayer Crop Science Report. Document no. M-433396-01-1.
2014. The effect of temperature on microbially produced Barstar assessed by ELISA. Unpublished Bayer Crop Science Report. Document no. M-479248-01-1.

History of human consumption

S. hygroscopicus (source of the *bar* gene) and *S. viridochromogenes* (source of the *pat* gene) 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 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 (Hérouet et al. 2005).

Both Barnase and Barstar are produced by *B. amyloliquefaciens* which is widespread in the environment (see Section 2.2.1) and homologues of the two proteins occur in a range of other bacterial species. Hence humans have had an indirect safe history of exposure to the two proteins. While both proteins are also expressed in approved GM canola lines that have been used for food production for over 15 years, it is noted that neither protein has been expressed in edible parts of these canola plants.

Amino acid sequence similarity to known protein toxins

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

Similarity searches were done for the PAT, Barnase and Barstar proteins using the Fast Alignment Search Tool – All (FASTA) algorithm (Pearson and Lipman 1988) version 35.04 and BLOSUM50[15] scoring matrix against known protein sequences present in a) a non-redundant general protein sequence database, containing over 81 million sequences, compiled from several publicly available databases by the NCBI and b) a Bayer toxin-specific database (Toxin_BCS_prot_16_1) built from the NCBI database, but with additions from the UniProt animal toxin annotation program, and Bt sequences belonging to the parasporin family, then filtered using certain key words and removal of redundancies; this contained 24,496 sequences.

Comparisons between highly homologous proteins yield *E*-values[16] approaching zero, indicating the very low probability that such matches would occur by chance. A larger *E*-value indicates a lower degree of similarity. Typically, alignments between two sequences will need to have an *E*-value of $1e^{-5}$ (1×10^{-5}) or smaller to be considered to have significant homology. However, any conclusion about the significance of a homology needs to be tempered by an understanding of its biological relevance.

[15] The BLOSUM series of matrices tabulate the frequency with which different substitutions occur in conserved blocks of protein sequences and are effective in identifying distant relationships. BLOSOM50 is the default for the FASTA programmes (Pearson 2013).
[16] The Statistics of Sequence Similarity Scores - https://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html

The results of the overall homology searches with the general protein database showed there were matches for all three proteins. For PAT, these matches were with acetyltransferase sequences from various origins. For Barnase and Barstar, the matches were with proteins from their respective ribonuclease families. No biologically relevant identities were found for any of the three proteins with any toxic proteins from the Bayer toxin database.

Heat stability

The thermolability of a protein provides an indication of the stability of the protein under cooking/processing conditions.
For all three proteins, heat stability was determined by measuring enzymatic activity (as described in Sections 4.1.4.1 and 4.1.4.2) after each protein was pre-incubated at 4^{o} C, 25^{o} C, 37^{o} C, 55^{o} C, 75^{o} C, 95^{o} C for 30 min. The purified enzyme preparations used in the analysis were derived from microbially-expressed proteins.

For PAT, activity remained maximal after incubation at 4^{o} C, 25^{o} C, and 37^{o} C. After incubation at temperatures at or above 55^{o} C there was a complete loss of activity.

For Barnase, activity remained maximal after incubation at 4^{o} C, 25^{o} C, and 37^{o} C. After treatment at 55°C and 75°C, activity fell approximately 20% and 70% respectively. There was no residual specific activity detected after treatment at 95°C.

For Barstar, activity remained maximal after incubation at 4^{o} C, 25^{o} C, 37^{o} C and 55^{o} C. The IC_{50} for Barstar treated at 75°C increased 1.66-fold over the IC_{50} measured at 4°C. The IC_{50} for Barstar treated at 95°C was not able to be calculated, because there was no difference in the observed rates of Barnase inactivation by Barstar at any of the concentrations of the 95°C-treated Barstar. This indicated that Barstar sample treated at 95°C had no activity as an inhibitor of Barnase.

Heat stability was also determined for the Barnase and Barstar proteins by SDS-PAGE/western blotting and ELISA. Results from both analyses aligned with the activity results; Barnase begins to become unstable at 55^{o} C and Barstar is not stable at 95^{o} C.

The main food use for canola is oil. The temperatures required to completely inactivate all three proteins are within those normally applied during the cooking/conditioning of canola seed prior to oil extraction. Cooking serves to thermally rupture oil cells which have survived flaking, reduce oil viscosity thereby promoting coalescence of oil droplets, increase the diffusion rate of prepared oil cake, and denature hydrolytic enzymes. The cooking cycle usually lasts 15 to 20 minutes and the temperatures usually range between 80 and 105°C, with an optimum of about 88°C (OECD 2011).

4.1.5.2 Potential allergenicity of PAT, Barnase and Barstar

The potential allergenicity of the three novel proteins was evaluated using an integrated, step-wise, case-by-case approach relying on various criteria used in combination. This is because no single criterion is sufficiently predictive of either allergenicity or non-allergenicity (see e.g. Thomas et al. 2009). The assessment focuses on:

- the source of the novel protein;
- any significant amino acid sequence similarity between the novel protein and known allergens;
- the structural properties of the novel protein, including susceptibility to digestion, heat stability and/or enzymatic treatment; and

- specific serum screening if the novel protein is derived from a source known to be allergenic or has amino acid sequence similarity to a known allergen.

Applying this approach systematically provides reasonable evidence about the potential of a novel protein to act as an allergen.

Studies submitted

Bioinformatics

2016. PAT/*bar* protein - Amino acid sequence homology search with known allergens and known toxins. Unpublished Bayer Crop Science Report. Document no. M-084359-10-1.

2016. Barnase protein - Amino acid sequence homology search with known allergens and known toxins. Unpublished Bayer Crop Science Report. Document no. M-552256-01-1.

2016. Barstar protein - Amino acid sequence homology search with known allergens and known toxins. Unpublished Bayer Crop Science Report. Document no. M-552416-01-1.

***In vitro* digestibility**

2009. PAT/*bar* protein: in vitro digestibility study in human simulated gastric fluid. Unpublished Bayer Crop Science Report. Document no. M-217195-04-1.

2012. Barnase protein - In vitro digestibility study in human simulated gastric fluid at pH 1.2. Unpublished Bayer Crop Science Report. Document no. M-430109-01-1.

2012. Barstar protein in vitro digestibility study in human simulated gastric fluid at pH 1.2. Unpublished Bayer Crop Science Report. Document no. M-429793-01-1.

2016. PAT/*bar* protein: in vitro digestibility study in human simulated intestinal fluid - Report amendment no 1 of final report. Unpublished Bayer Crop Science Report. Document no. M-20879304-1.

2016. Barnase protein - In vitro digestibility study in human simulated intestinal fluid – Report amendment no.1 of final study. Unpublished Bayer Crop Science Report. Document no. M-430112-01-1.

2016. Barstar protein - In vitro digestibility study in human simulated intestinal fluid - Report amendment no 1 of final report. Unpublished Bayer Crop Science Report. Document no. M-429800-02-1.

2014. Recombinant Barnase/Barstar complex protein: In vitro digestibility study in human simulated gastric fluid at pH 1.2. Unpublished Bayer Crop Science Report. Document no. M-476903-01-1.

2014. Recombinant Barnase/Barstar complex protein: In vitro digestibility study in human simulated intestinal fluid. Unpublished Bayer Crop Science Report. Document no. M-476904-01-1.

Sources of the proteins

The PAT protein is derived from *S. hygroscopicus* and the Barnase and Barstar proteins are derived from *B. amyloliquefaciens* (see Section 2.2). No endogenous proteins from either of these species are listed in the Food Allergy Research and Resource Program (FARRP) Allergen Protein Database[17] (accessed 19 January 2017 and containing 2,035 entries). As indicated in Section 2.2, *B. amyloliquefaciens* has been used as a safe production organism for a number of food-grade enzymes.

Similarity to known allergens

Bioinformatic analysis provides part of a weight of evidence approach for assessing potential allergenicity of novel proteins introduced to GM plants (Thomas et al. 2005; Goodman 2006). It is a method for comparing the amino acid sequence of the introduced protein with sequences of known allergens in order to indicate potential cross-reactivity between allergenic proteins and the introduced protein. As with the bioinformatic analysis that looked at similarities of the novel proteins with known protein toxins (refer to Section 4.1.5.1), the generation of an *E* value provides an important indicator of significance of matches (Pearson 2000; Baxevanis 2005).

[17] AllergenOnline - http://www.allergenonline.org/

To evaluate the similarity to known allergens of the three proteins an epitope search was carried out to identify any short sequences of amino acids that might represent an isolated shared allergenic epitope.

The PAT, Barnase and Barstar sequences were compared with all known putative allergen and celiac sequences residing in a reference allergen database, (FARRP version 16, released on 27 January 2016 – containing 1,956 non-redundant entries) using the FASTA (version 35.04) algorithm and BLOSUM50 scoring matrix (refer to Section 4.1.5.1).

No biologically relevant alignment for any of the three proteins met or exceeded the Codex Alimentarius (Codex 2009) FASTA alignment threshold for potential allergenicity (35% identity over 80 amino acids) and no alignments of eight or more consecutive identical amino acids were found between any of the proteins and known allergens in the database. For CRTI, a search was also done to compare every possible peptide of eight contiguous amino acids (Metcalfe et al. 1996) with the sequences in the FARRP database. No alignments were found. It was concluded that PAT, Barnase and Barstar do not contain any cross-reactive IgE binding epitopes with known allergens.

Glycosylation search

N-glycosylated proteins are glycosylated on an asparagine residue and commonly contain an asparagine-X-serine/threonine sequence (N-X~(P)-[S/T), where X~(P) indicates any amino acid except proline (Orlando and Yang 1998). Although rare, the sequence asparagine-X-cysteine (N-X-C) can also be an N-glycosylation site (Miletich and Broze Jr. 1990). Using a bioinformatics approach (FASTA and BLOSOM50 scoring matrix), these sequences in the PAT, Barnase and Barstar proteins were searched for potential N-glycosylation sites. None were identified in any of the three proteins. This confirmed the findings from the glycostaining analyses described in Sections 4.1.4.1 and 4.1.4.2.

Heat stability – see Section 4.1.5.1

In vitro digestibility

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

For the PAT, Barnase and Barstar proteins (all purified from a bacterial system), analyses using simulated gastric fluid – SGF (containing pepsin) (U.S. Pharmacopeia 1990) and simulated intestinal fluid – SIF (containing pancreatin, a mixture of enzymes including amylase, trypsin, lipase, ribonuclease and protease) (U.S. Pharmacopeia 1990) were done. The SIF study by itself may not be entirely informative because ordinarily an ingested protein would first be exposed to pepsin-mediated hydrolysis in the acidic environment of the stomach before being subject to further digestion in the small intestine.

a) Simulated gastric fluid (SGF)

A pepsin digestibility assay (Thomas et al. 2004) was conducted. The three proteins were incubated in SGF at 37º for 0, 0.5, 1 (except PAT), 2, 5, 10, 20, 30 and 60 minutes and then inactivated by addition of $NaHCO_3$. The samples were then run on SDS-PAGE. Proteins were visualised by Coomassie blue staining of the resulting gels. Western blotting of the SDS gels was also performed using an appropriate rabbit polyclonal primary antibody and a HRP-conjugated goat anti-rabbit polyclonal secondary antibody.

Following exposure to SGF for 30 seconds, the earliest time point sampled during the digestion, more than 90% of the PAT protein had disappeared and all of the Barnase and Barstar proteins had degraded - as assessed by both SDS-PAGE and western immunoblot analysis.

b) Simulated intestinal fluid (SIF)

The temperatures, sampling times and SDS-PAGE/western blot analyses for the SIF procedure were the same as for the SGF procedure (with the exception that there was no 1 s incubation for any of the proteins). Reactions were terminated by addition of distilled water.

Following exposure to SIF for 30 seconds, the earliest time point sampled during the digestion, more than 90% of the PAT protein had disappeared as assessed by both SDS-PAGE and western immunoblot analysis. For Barnase, there was only slight change to the protein after 60 min digestion. For Barstar, more than 90% of the protein was digested within 10 min.

c) Conclusion

These results support the conclusion that the PAT, Barnase and Barstar proteins are readily digested by pepsin under simulated gastric conditions and that PAT and Barstar are readily digested under typical mammalian intestinal conditions. Overall, the three proteins would be degraded in a mammalian digestive system.

4.1.5.3 Safety of the Barnase/Barstar complex

Since MS11 will be crossed with the fertility restorer line RF3, the Applicant also provided safety data (thermolability and digestibility studies) on the Barnase/Barstar complex protein (produced in *E. coli*). However, it is noted that in the MS11 x RF3 hybrid neither of the two proteins is expressed in seed (Table 4), the only part of the plant used for human food. Since the main food produced from canola seed is refined oil, which is unlikely to contain any protein, the overall likelihood of food for human consumption containing either protein is negligible.

Studies submitted

2014. The effect of temperature on microbially-produced Barnase/Barstar complex protein assessed by the Barnase quantitative activity assay and the Barstar quantitative activity assay. Unpublished Bayer Crop Science Report. Document no. M-492536-01-1.
2013. The effect of temperature on microbially produced Barnase/Barstar protein - Complex assessed by SDS-PAGE and western blot. Unpublished Bayer Crop Science Report. Document no. M-549535-01-1.
2014. The effect of temperature on microbially produced Barnase/Barstar protein complex assessed by ELISA. Unpublished Bayer Crop Science Report. Document no. M-477906-01-1.

Heat stability was determined by measuring enzymatic activity (as described in Sections 4.1.4.1 and 4.1.4.2) after pre-incubation of the microbially-produced complex at 4º C, 25º C, 37º C, 55º C, 75º C, 95º C for 30 min (same conditions as for the heat stability testing of the individual proteins – Section 4.1.5.1).

The residual Barnase specific activity and the residual Barstar IC_{50} of each temperature-treated sample were measured by the Barnase quantitative activity assay and the Barstar quantitative activity assay, respectively.

As expected from the heat stability results described in Section 4.1.5.1, the Barnase/Barstar complex showed no residual Barnase or Barstar activity at any of the incubation temperatures.

Heat stability was also determined by SDS-PAGE/western blotting and ELISA. Results from both analyses suggest that the Barnase/Barstar protein complex is degraded and forms oligomers upon heating at 75°C and above.

In vitro digestibility was analysed using both SGF and SIF as described in Section 4.1.5.2. In SIF, there was no degradation of the complex within 60 min. However, the full length recombinant Barnase/Barstar complex protein was degraded very rapidly, within 30 seconds of incubation with SGF, and by 5 min the small molecular weight residual fragments (2.5 to 3.5 kDa) were fully degraded. This indicated the Barnase/Barstar complex would be readily digested in a mammalian digestive system.

4.1.6 Bioinformatic analyses of potential ORFs created by the transformation procedure

Study submitted:

2016. MS11 *Brassica napus* - Identification of open reading frames (ORF) and homology search of sequences of more than 30 amino acids to known allergens and toxins. Unpublished Bayer Crop Science Report. Document no. M- 552421-01-1.

Bioinformatics analyses were performed to assess the similarity to known allergens and toxins of the putative polypeptides encoded by the 107 identified ORFs in the insert and flanking regions of the MS11 transgenic locus (see Section 3.4.4).

The bioinformatics analyses were carried out by comparing the sequences of the 107 ORFs with sequences present in the same databases as described for the toxin and allergen searches in Sections 4.1.5.1 and 4.1.5.2 respectively, and using the same FASTA algorithm and BLOSOM50 scoring matrix.

No biologically relevant identities were found between any of the query sequences and any toxic proteins from the NCBI non-redundant database.

For the allergen search, no identity matches of greater than 35% over 80 residues were observed for any of the ORFs, and no eight contiguous identical amino acid matches were observed for any ORF i.e. none of the ORFs contain any cross-reactive IgE binding epitopes with known allergens.

It is concluded that in the event an unexpected translation product were to be derived from any of the 107 ORFs detected in the MS11 insert and flanking regions, these putative polypeptides are not expected to possess functional cross-reactivity with known allergenic proteins or be toxic.

4.1.7 Conclusion

MS11 expresses three novel proteins, PAT, Barnase and Barstar.

Mean levels of all proteins in the edible part (i.e. seed) were below or close to the LOQ. The mean level of PAT protein was highest in whole plants sampled at the 3 – 5 leaf stage, while lowest values were obtained in root and seed samples. The Barnase protein was below the LOQ in all samples tested. Levels of Barstar in all tissues were either below the LOQ or very low; roots marginally appeared to have the highest mean level.

A number of studies were used to confirm the identity and physicochemical properties of the plant-derived PAT protein. These studies demonstrated that the protein conforms in size, amino acid sequence and activity to that expected, and does not exhibit any post-translational modification including glycosylation.

Very low yields of the plant-produced Barnase and Barstar proteins precluded their specific characterisation. However the weight of evidence, provided by a) translation of the known DNA sequences of the two genes introduced into MS11, b) the fact that the proteins function as predicted and c) the detailed characterisation of equivalent microbially-produced proteins is sufficient to confirm the identity of the proteins expressed in MS11.

For all three proteins, bioinformatic studies confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens; digestibility studies suggest the proteins would be rapidly degraded in the gastro-intestinal tract following ingestion; and thermolability studies indicate the three proteins are functionally inactivated following heating.

Taken together, the evidence indicates that should PAT, Barnase or Barstar be present in the diet they are unlikely to be toxic or allergenic in humans.

4.2 Herbicide metabolites

For GM foods derived from crops that are herbicide tolerant, there are two issues that require consideration. The first is dealt with in this safety assessment and involves assessment of any novel metabolites that are produced after the herbicide is applied, to determine whether these are present in the final food 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 consideration, 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[18] (MRL). Where necessary, an MRL may have to be set.

In the case of PAT, the metabolic profiles resulting from the novel protein x herbicide interaction have been established through a significant history of use. The glufosinate-tolerance trait is present in lines from 22 previous applications to FSANZ. The enzyme activity of 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-tolerant plants and was previously considered in detail by FSANZ in cotton line LL25 (FSANZ 2006). There are no concerns that the spraying of MS11 with glufosinate ammonium would result in the production of metabolites that are not also produced in crops sprayed with the same herbicide and already used in the food supply.

[18] http://www.foodstandards.gov.au/consumer/chemicals/maxresidue/Pages/default.aspx

5 Compositional analyses

The main purpose of compositional analyses is to determine if, as a result of the genetic modification, any unexpected changes have occurred to the food. These changes could take the form of alterations in the composition of the plant and its tissues and thus its nutritional adequacy. Compositional analyses can also be important for evaluating the intended effect where there has been a deliberate change to the composition of the food.

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

5.1 Key components

Canola oil is the primary food product used for human consumption. The key components to be analysed for a comparison between transgenic and conventional canola are proximates, amino acids, fatty acids (seed or oil), vitamins E and K, glucosinolates, tannins, sinapine and phytic acid (OECD 2011).

5.2 Study design and conduct for key components

Studies submitted:

2016. MS11 *B. napus* - Composition analysis of field samples grown in Canada and the USA during 2014. Unpublished Bayer Crop Science Report. Document no. M-549080-01-1.
2016. MS11 x RF3 and MS11 *B. napus* – Field production in Canada and the USA during 2014. Unpublished Bayer Crop Science Report. Document no. M-549076-01-1.
2016. MS11 *B. napus* - Processing of Grain and Analysis of Resultant Fractions, 2015. Unpublished Bayer Crop Science Report. Document no. M-552078-01-1.

Verified seed lots of MS11 (generation T_4), and non-GM N90-740 were used for plantings at 10 sites[19] across Canada and the USA in 2014.

Planting and crop maintenance were done according to local agronomic practices at each site. For MS11, plants were either left unsprayed with glufosinate or sprayed with Liberty® 280SL containing approximately 280 g ai/L glufosinate ammonium applied at the 2 – 4 leaf growth stage (BBCH12-14). Additionally, a total of six non-GM registered commercial lines[20] were also grown as reference lines, with three different lines being grown at each site, in order to generate a tolerance interval for each analyte and hence to aid in the determination of the normal variation found in canola analyte levels. These reference lines are representative of, and adapted to, the growing conditions of the major *B. napus* growing regions in Canada and the USA.

[19] Hoodoo – Saskatchewan, Canada; Sturgeon - Alberta, Canada; North Norfolk – Manitoba, Canada; MacDonald – Manitoba, Canada; Whitewater – Manitoba, Canada; Corman Park – Saskatchewan, Canada; Grand Forks – North Dakota, USA; Case – North Dakota, USA; Jerome, Idaho, USA; Grant – Washington, USA
[20] Reference lines - 46A65; AC Elect; AC Excel; Peace; Spectrum; Westar.

Each trial site comprised six entries:

- A – N90-740
- B – MS11 unsprayed with glufosinate
- C – MS11 sprayed with glufosinate
- F, G, H, I, J, K – three of the six reference lines.

Each entry was replicated four times (24 plots in total) in a randomised complete block design. Each entry was planted in a minimum of six-row plots of 5-6 meters in length. The seeding density for the MS11 plots sprayed with glufosinate ammonium was double that of the other plots because approximately 50 % of seedlings (negative segregants) were not tolerant to glufosinate-ammonium application. Therefore, all plots resulted in approximately the same plant density after glufosinate-ammonium spraying. A further complication in experimental design was that normal reproduction in canola occurs primarily through self-pollination (70%) but in order for the male-sterile MS11 plants to produce seed there would need to be 100% pollination from adjacent fertile canola. Ideally, for consistency in analytical results, the adjacent plants would be the N90-740 parental line and, indeed, in the unsprayed MS11 plots this is the case. However, in the sprayed MS11 plants, any adjacent non-GM N90-740 plants are destroyed by spraying so that fertilisation is achieved by pollen from neighbouring plots that comprise N90-740 as well as up to three different non-GM commercial reference lines. Therefore there is expected to be considerably less consistency between the results for seed analytes from the sprayed and unsprayed MS11.

Seed was harvested at physiological maturity; each sample consisted of approximately 0.3 kg of seed, with the exception of some plots with lower yields.

Methods of composition analysis were based on internationally recognised procedures (e.g. those of the Association of Official Analytical Chemists), methods specified by the manufacturer of the equipment used for analysis, or other published methods.

5.3 Analyses of key components in seed

The replicated sites were analysed both separately (by-site analysis) and combined across all sites (combined-site analysis) using a mixed model analysis of variance. Descriptive statistics (mean and standard deviation (SD) were generated and are presented in Tables 5 – 10 which represent results from combined-site analyses. Pairwise comparisons (t-test) were made between the non-GM control (A) and the herbicide unsprayed (B) and sprayed (C) lines containing event MS11 (i.e. A vs B; A vs C). In assessing the significance of any difference between least square means, a P-value of 0.05 was used (i.e. a P-value of ≥0.05 was not significant).

Any statistically significant differences between MS11 and the N90-740 control have been compared to the 95% tolerance interval (i.e. 95% confidence that the interval contains 99% of the values expressed in the commercial lines) compiled from the results of the six commercial reference lines combined across all sites, to assess whether the differences are likely to be biologically meaningful. Additionally, the results for MS11 and N90-740 have been compared to a combined literature range (where available) for each analyte, compiled from published literature for commercially available canola[21]. It is noted, however, that information in the published literature is limited and is unlikely to provide a broad reflection of the natural diversity that occurs within canola (Harrigan et al. 2010). Therefore, even if means fall outside the published range, this is unlikely to raise a concern.

[21] Published literature for canola includes Wang et al (1999), Pritchard et al (2000), Szmigielska et al (2000), Marwede et al (2004), Barthet & Daun (2005), Brand et al (2007), Seberry et al (2007), Spragg & Mailer (2007), OECD (2011).

Seed samples were analysed for proximates, fibre, fatty acids, amino acids, minerals, vitamins, anti-nutrients (phytic acid, tannins and sinapine) and glucosinolates. In total, 86 analyte levels were measured, and carbohydrate was calculated rather than being measured i.e. there were 87 analytes considered. A total of 30 analytes had more than 33% of the observations below the assay limit of quantitation (LOQ) and were excluded from the statistical analysis. The data for 57 analytes were therefore analysed.

Data were transformed into Statistical Analysis Software[22] (SAS) data sets and analysed using SAS® software (SAS, version 9.3). The SAS GLM procedure was applied to all data (test, control and reference) to detect potential outliers in the dataset by screening studentised PRESS residuals[23]. None of the values was considered to be an outlier.

Bad weather (flooding) at one site (Case – North Dakota, USA) caused significant damage to the plots and this site was therefore removed from the statistical consideration.

5.3.1 Proximates and fibre

The results are given in Table 5 and show there was no significant difference between the means for unsprayed MS11 and those for the control for any analyte measured. For sprayed MS11, levels of protein, fat, acid detergent fibre (ADF) and neutral detergent fibre (NDF) were significantly lower than those in the control and the levels of ash and moisture were significantly higher than those in the control. However, the means for all entries, for all proximates and fibre, were within the range of the reference varieties and the tolerance intervals.

Table 5: Mean percentage ±SD of proximates and fibre in seed of MS11 and the N90-740 control collected from nine locations

Analyte[1]	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B[3]	p-value A vs C	95% Tolerance interval	Combined literature range
Protein (%)	28.4±2.3	29.1±2.5	29.6±1.3[4]	0.213	**0.006**	21.0, 35.6	17.4 – 44.3
Fat (%)	37.1±4.6	36.5±4.3	33.8±4.7[4]	0.261	**<0.001**	26.4, 49.7	24.0 – 49.5
Ash (%)	4.87±0.86	5.00±0.84	5.47±0.91[4]	0.316	**0.001**	1.36, 7.99	3.36 – 6.02
Carbohydrate (%)[2]	29.7±2.3	29.4±1.8	31.2±3.3	0.994	0.078	20.7,37.3	
ADF (%)	21.3±1.9	21.5±1.7	20.1±2.7[4]	0.979	**0.023**	14.0, 25.4	11.6 – 26.7
NDF (%)	25.0±1.9	24.6±1.9	23.4±2.5[4]	0.413	**0.011**	17.1, 29.4	16.49 – 34.72
Moisture (% fw)	9.73±2.78	9.49±2.28	11.54±4.20[4]	0.922	**0.01**	1.32, 17.25	3.17 – 10.0

[1] Analyte means determined on a dry weight (dw) basis (except for moisture)
[2] Carbohydrate determined by calculation
[3] p-values indicating significant differences are bolded and underlined.
[4] mauve shading represents MS11 means that are significantly lower than the N90-740 means while orange shading represents MS11 means that are significantly higher.

[22] SAS website - http://www.sas.com/technologies/analytics/statistics/stat/index.html
[23] A PRESS (predicted residual sum of squares) statistic provides a comparison of the predicted marginal mean and the observed mean when the predicted value is calculated without the deleted observation in question (Schabenberger 2004).

5.3.2 Fatty acids

The levels of 33 fatty acids were measured, and of these the following 22 contained more entries below the LOQ than could be considered in a statistical analysis: caproic (6:0), caprylic (8:0), capric (10:0), lauric (12:0), myristic (14:0), myristoleic (14:1), pentadecanoic (15:0), pentadecenoic (15:1), heptadecanoic (17:0), heptadecenoic (17:1), gamma linolenic (18:3), octadecatetraenoic (18:4), eicosadienoic (20:2), eicosatrienoic (20:3), homogamma linolenic (20:3), arachidonic n3 (20:4), arachidonic n6 (20:4), erucic (22:1), docosadienoic (22:2), docosapentaenoic n3 (22:5); docosapentaenoic n6 (22:5), and docosahexaenoic (22:6).

The results for the remaining 11 fatty acids (Table 6) show there was no significant difference between the means for unsprayed MS11 and those for the control for any fatty acid measured. For sprayed MS11, levels of palmitic, stearic, arachidic, behenic, lignoceric and nervonic acids were significantly higher than those in the control and the levels of oleic acid was significantly lower than that in the control. However, the means for all entries, for all fatty acids, were within the tolerance interval of the reference varieties.

The means for oleic acid in both MS11 and N90-740 were higher than the combined literature range, while the means for eicosenoic acid in both MS11 and N90-740 were lower than the combined literature range. The means for all other analytes for both MS11 and N90-470 were within the literature range.

Table 6: Mean percentage ±SD composition, relative to total fat of major fatty acids in seed from MS11 and the N90-740 control collected from nine locations

Fatty acid % total	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	95% Tolerance interval	Combined literature range
Palmitic (16:0)	4.17±0.18	4.21±0.20	4.34±0.20[4]	0.357	**0.013**	3.34, 5.22	2.7 – 7.0
Palmitoleic (16:1)	0.226±0.018	0.215±0.018	0.235±0.021	0.067	0.175	0.179, 0.302	ND[2] – 0.6
Stearic (18:0)	2.16±0.25	2.22±0.28	2.27±0.34[2]	0.236	**0.009**	1.40, 2.72	0.8 – 3.0
Oleic (18:1)	63.1±2.0	63.5±2.0	61.6±2.5[2]	0.433	**<0.001**	52.2, 69.3	8.0 – 60.0
Linoleic (18:2)	18.4±1.0	18.2±1.0	18.8±1.3	0.543	0.207	13.9, 26.6	15.0 – 30.0
α-Linolenic (18:3)	9.05±1.40	8.70±1.41	9.55±1.49	0.085	0.053	4.06, 14.37	5.0 – 13.0
Arachidic (20:0)	0.731±0.068	0.757±0.080	0.782±0.103[2]	0.097	**<0.001**	0.436, 0.936	ND[3] – 3.0
Eicosenoic (20:1)	1.34±0.10	1.38±0.11	1.42±0.12	0.486	0.071	0.11, 2.95	3.0 – 15.0
Behenic (22:0)	0.408±0.042	0.425±0.049	0.452±0.060[2]	0.076	**<0.001**	0.183, 0.547	ND[3] – 2.0
Lignoceric (24:0)	0.198±0.039	0.209±0.043	0.234±0.055[2]	0.132	**<0.001**	0.075, 0.314	ND[3] – 2.0
Nervonic (24:1)	0.195±0.043	0.191±0.050	0.221±0.052[2]	0.896	**0.018**	0.057, 0.338	ND[3] – 3.0

[1] p-values indicating significant differences are bolded and underlined.
[2] mauve shading represents MS11 means that are significantly lower than the N90-740 means while orange shading represents MS11 means that are significantly higher.
[3] ND = not detectable

5.3.3 Amino acids

Levels of 18 amino acids were measured. Since asparagine and glutamine are converted to aspartate and glutamate respectively during the analysis, levels for aspartate include both aspartate and asparagine, while glutamate levels include both glutamate and glutamine. Results of the analysis are given in Table 7.
There was no significant difference between the means for unsprayed MS11 and those for the control for any amino acid measured. For sprayed MS11, levels of alanine, cysteine, lysine, methionine, phenylalanine, proline, serine and tryptophan were significantly higher than those in the control. However, the means for all entries, for all amino acids, were within the tolerance interval of the reference varieties.

The means for cysteine, phenylalanine and tryptophan in both MS11 and N90-740 were higher than the combined literature range. The means for all other analytes for both MS11 and N90-470 were within the literature range.

Table 7: Mean ±SD amino acid composition (% dw) in seed from MS11 and the N90-740 control collected from nine locations

Amino acid % dw	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	95% Tolerance interval	Combined literature range
Alanine	1.27±0.10	1.28±0.09	1.32±0.06[2]	0.984	**0.005**	0.96, 1.53	0.71 – 1.38
Arginine	1.96±0.17	1.98±0.17	2.01±0.11	0.745	0.064	1.38, 2.43	0.93 – 2.46
Aspartate	2.25±0.23	2.25±0.22	2.29±0.15	0.822	0.227	1.46, 2.67	1.20 – 2.03
Cysteine	0.627±0.064	0.639±0.059	0.644±0.056[2]	0.428	**0.039**	0.474, 0.910	0.32 – 0.52
Glutamate	4.91±0.46	4.99±0.41	5.03±0.30[2]	0.567	0.057	3.57 – 6.28	3.23 – 4.71
Glycine	1.43±0.12	1.44±0.10	1.45±0.08[2]	0.774	0.115	1.06, 1.75	0.82 – 2.22
Histidine	0.697±0.058	0.721±0.064	0.720±0.048[2]	0.237	0.061	0.522, 0.910	0.41 – 0.82
Isoleucine	1.19±0.10	1.20±0.09	1.21±0.06	0.729	0.239	0.84, 1.46	0.62 – 1.02
Leucine	2.06±0.17	2.08±0.16	2.10±0.10	0.770	0.107	1.51, 2.49	1.07 – 1.99
Lysine	1.52±0.11	1.57±0.11	1.59±0.09[2]	0.120	**0.002**	1.20, 1.98	0.96 – 1.85
Methionine	0.505±0.037	0.511±0.040	0.522±0.035[2]	0.757	**0.021**	0.379, 0.657	0.27 – 0.52
Phenylalanine	1.19±0.10	1.20±0.10	1.22±0.06[2]	0.704	**0.044**	0.87, 1.44	0.64 – 1.07
Proline	1.73±0.21	1.77±0.21	1.83±0.26[2]	0.362	**<0.001**	1.22, 2.32	0.85 – 3.74
Serine	1.22±0.10	1.23±0.09	1.25±0.06[2]	0.685	**0.043**	0.92, 1.47	0.69 – 1.55
Threonine	1.22±0.09	1.23±0.08	1.24±0.06	0.946	0.205	0.93, 1.47	0.74 – 1.30

Amino acid % dw	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	95% Tolerance interval	Combined literature range
Tryptophan	0.416±0.033	0.420±0.036	0.428±0.028[2]	0.718	**0.050**	0.283, 0.534	0.20 – 0.37
Tyrosine	0.899±0.067	0.901±0.064	0.911±0.040	0.933	0.187	0.669, 1.072	0.51 – 0.92
Valine	1.45±0.12	1.46±0.11	1.47±0.07	0.996	0.126	1.06, 1.77	0.8 – 1.55

[1] p-values indicating significant differences are bolded and underlined.
[2] orange shading represents MS11 means that are significantly higher than the N90-740 means.

5.3.4 Minerals

Levels of nine minerals were measured. Sodium was below the LOQ in many samples and was excluded from the analysis. The means for the remaining minerals are given in Table 8. There was no significant difference between the means for unsprayed MS11 and those for the control for any mineral measured. For sprayed MS11, levels of copper, magnesium, phosphorus, potassium and zinc were significantly higher than those in the control. However, the means for all entries, for all minerals, were within both the tolerance interval of the reference varieties and the combined literature range.

Table 8: Mean ±SD mineral composition (mg/kg dw) in seed from MS11 and the N90-740 control collected from nine locations

Mineral (mg/kg dw)	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	95% Tolerance interval	Combined literature range
Calcium	4885±1102	4651±718	5227±1361	0.360	0.078	2359, 7211	3610 – 7280
Copper	4.33±0.79	4.29±0.94	4.68±0.91[2]	0.966	**<0.001**	1.57, 6.27	1.57 – 5.39
Iron	131.6±85.1	119.6±75.8	158.5±95.6	0.731	0.264	0, 461.9	ND – 900.59
Magnesium	3659±438	3654±413	3938±560[2]	0.126	**<0.001**	2126, 5060	2770 –4270
Manganese	39.4±7.5	38.1±6.7	38.4±5.4	0.675	0.148	20.3, 60.5	33.95 – 65.20
Phosphorus	7947±1450	8219±1540	8674±1253[2]	0.450	**0.003**	2316, 13044	5400 – 8900
Potassium	8577±1374	8865±1424	9516±1667[2]	0.144	**<0.001**	3672, 12,636	7020 – 10,200
Zinc	48.4±8.8	50.2±9.6	54.1±9.3[2]	0.211	**<0.001**	21.8, 70.3	ND – 122.362

[1] p-values indicating significant differences are bolded and underlined.
[2] orange shading represents MS11 means that are significantly higher than the N90-740 means.

5.3.5 Vitamins

Levels of three tocopherols and Vitamin K were measured. β-tocopherol was below the LOQ in many samples and was excluded from the analysis. The means for the remaining vitamins are given in Table 9. There was no significant difference between the means for unsprayed MS11 and those for the control for any vitamin measured. For sprayed MS11, levels of γ-tocopherol and Vitamin K were significantly higher than those in the control.

However, the means for all entries, for all minerals, were within the tolerance interval of the reference varieties.

Table 9: Mean ±SD vitamin composition (mg/kg dw) in seed from MS11 and the N90-740 control collected from 9 locations

Vitamin (mg/kg.dw)	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	95% Tolerance interval	Combined literature range
α-tocopherol	94.8±11.6	93.5±7.4	91.2±10.9	0.356	0.436	48.0, 154.9	71.1 – 108.4
γ-tocopherol	171±29	165±24	153±29[2]	0.316	**0.028**	44 - 326	
Vitamin K	1.297±0.410	1.286±0.359	1.702±0.604[2]	0.891	**0.002**	0.168 - 2.140	

[1] p-values indicating significant differences are bolded and underlined.
[2] orange shading represents MS11 means that are significantly higher than the N90-740 means.

5.3.6 Anti-nutrients

Levels of 10 glucosinolates (plus total glucosinolates), phytic acid, sinapine and three categories of tannins were measured. Six of the glucosinolates (epi-progoitrin, glucoalyssin, glucobrassicanapin, gluconastutiin, glucoraphanin, neoglucobrassicin) were below the LOQ in many samples and were excluded from the analysis.

The results (Table 10) can be summarised as follows:

- No significant differences were observed between the means of N90-740 and unsprayed MS11 or those of N90-740 and sprayed MS11 for 4-hydroxyglucobrassicin, glucobrassicin, phytic acid, sinapine, soluble tannins or total condensed tannins.
- Means for gluconapin were significantly higher than the N90-740 control in both sprayed and unsprayed MS11.
- Means for progoitrin and total glucosinolates in sprayed MS11 were significantly higher than those in N90-740.
- Mean for insoluble tannins was significantly higher in unsprayed MS11 than in N90-740 (noting that while the reported arithmetic mean in Table 10 for unsprayed MS11 is actually lower than the sprayed MS11 mean which was not significant, the least square mean derived from the statistical analysis gave 0.404, 0.467 and 0.447 for N90-740, unsprayed MS11 and sprayed MS11 respectively).
- For all anti-nutrients, the means for both sprayed and unsprayed MS11 fell within the tolerance interval of the reference varieties and, where available, within the literature range.

Table 10: Mean ±SD anti-nutrient levels in seed from MS11 and the N90-740 control collected from nine locations

Anti-nutrient	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	95% Tolerance interval	Combined literature range
4-Hydroxygluco-brassicin (µmol/g dw)	3.93±0.92	4.08±1.03	4.19±1.07	0.520	0.284	1.96 , 7.28	
Glucobrassicin (µmol/g dw)	0.391±0.222	0.425±0.204	0.479±0.219	0.516	0.177	0, 1.364	
Gluconapin (µmol/g dw)	2.09±0.85	2.69±1.11	2.99±1.16[2]	**0.013**	**<0.001**	0, 5.25	
Progoitrin (µmol/g dw)	5.48±2.75	5.88±2.86	7.38±3.51[2]	0.601	**0.001**	0, 13.67	
Total glucosinolates (µmol/g dw)	12.3±4.1	13.7±4.2	15.8±4.8[2]	0.179	**<0.001**	0.673, 26.6	1 - 28
Phytic acid (%dw)	2.11±0.43	2.18±0.46	2.29±0.30	0.698	0.051	0.31, 3.78	
Sinapine (%dw)	0.717±0.060	0.738±0.052	0.701±0.087[2]	0.526	0.739	0.337, 1.022	0.7 – 1.1
Insoluble tannins (%dw)	0.403±0.095	0.455±0.110	0.458±0.125[2]	**0.027**	0.109	0, 0.749	
Soluble tannins (%dw)	0.099±0.038	0.100±0.047	0.134±0.081[2]	0.778	0.051	0, 0.194	
Total condensed tannins (%dw)	0.503±0.121	0.554±0.147	0.591±0.189	0.097	0.060	0, 0.923	

[1] p-values indicating significant differences are bolded and underlined.
[2] orange shading represents MS11 means that are significantly higher than the N90-740 means

5.3.7 Summary of analyses of key components

A summary of the statistically significant differences in the mean analyte levels found between seed of unsprayed and sprayed MS11 and the N90-740 control is provided in Table 11.

Table 11: Summary of analyte levels found in seed of MS11 that are significantly (p<0.05) different from those found in seed of the control N90-740

Analyte	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	Within tolerance interval?	Within literature range?
Protein (%dw)	28.4±2.3	29.1±2.5	29.6±1.3[4]	0.213	**0.006**	yes	yes
Fat (%dw)	37.1±4.6	36.5±4.3	33.8±4.7[4]	0.261	**<0.001**	yes	yes
Ash (%dw)	4.87±0.86	5.00±0.84	5.47±0.91[4]	0.316	**0.001**	yes	yes
ADF (%dw)	21.3±1.9	21.5±1.7	20.1±2.7[4]	0.979	**0.023**	yes	yes
NDF (%dw)	25.0±1.9	24.6±1.9	23.4±2.5[4]	0.413	**0.011**	yes	yes
Moisture (% fw)	9.73±2.78	9.49±2.28	11.54±4.20[4]	0.922	**0.01**	yes	no

Analyte	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	Within tolerance interval?	Within literature range?
Palmitic acid (%total)	4.17±0.18	4.21±0.20	4.34±0.20[4]	0.357	**0.013**	yes	yes
Stearic acid (% total)	2.16±0.25	2.22±0.28	2.27±0.34[2]	0.236	**0.009**	yes	yes
Oleic acid (% total)	63.1±2.0	63.5±2.0	61.6±2.5[2]	0.433	**<0.001**	yes	no – control also higher
Arachidic acid (% total)	0.731±0.068	0.757±0.080	0.782±0.103[2]	0.097	**<0.001**	yes	yes
Behenic acid (% total)	0.408±0.042	0.425±0.049	0.452±0.060[2]	0.076	**<0.001**	yes	yes
Lignoceric acid (% total)	0.198±0.039	0.209±0.043	0.234±0.055[2]	0.132	**<0.001**	yes	no
Nervonic acid (% total)	0.195±0.043	0.191±0.050	0.221±0.052[2]	0.896	**0.018**	yes	yes
Alanine (% dw)	1.27±0.10	1.28±0.09	1.32±0.06[2]	0.984	**0.005**	yes	yes
Cysteine (% dw)	0.627±0.064	0.639±0.059	0.644±0.056[2]	0.428	**0.039**	yes	no – control also higher
Lysine (% dw)	1.52±0.11	1.57±0.11	1.59±0.09[2]	0.120	**0.002**	yes	yes
Methionine (% dw)	0.505±0.037	0.511±0.040	0.522±0.035[2]	0.757	**0.021**	yes	yes
Phenylalanine (% dw)	1.19±0.10	1.20±0.10	1.22±0.06[2]	0.704	**0.044**	yes	no – control also higher
Proline (% dw)	1.73±0.21	1.77±0.21	1.83±0.26[2]	0.362	**<0.001**	yes	yes
Serine (% dw)	1.22±0.10	1.23±0.09	1.25±0.06[2]	0.685	**0.043**	yes	yes
Tryptophan (% dw)	0.416±0.033	0.420±0.036	0.428±0.028[2]	0.718	**0.050**	yes	no – control also higher
Copper (mg/kg dw)	4.33±0.79	4.29±0.94	4.68±0.91[2]	0.966	**<0.001**	yes	yes
Magnesium (mg/kg dw)	3659±438	3654±413	3938±560[2]	0.126	**<0.001**	yes	yes
Phosphorus (mg/kg dw)	7947±1450	8219±1540	8674±1253[2]	0.450	**0.003**	yes	yes
Potassium (mg/kg dw)	8577±1374	8865±1424	9516±1667[2]	0.144	**<0.001**	yes	yes
Zinc (mg/kg dw)	48.4±8.8	50.2±9.6	54.1±9.3[2]	0.211	**<0.001**	yes	yes
γ-tocopherol (mg/kg dw)	171±29	165±24	153±29[2]	0.316	**0.028**	yes	Not available
Vitamin K (mg/kg dw)	1.297±0.410	1.286±0.359	1.702±0.604[2]	0.891	**0.002**	yes	Not available
Gluconapin (µmol/g dw)	2.09±0.85	2.69±1.11	2.99±1.16[2]	**0.013**	**<0.001**	yes	Not available
Progoitrin (µmol/g dw)	5.48±2.75	5.88±2.86	7.38±3.51[2]	0.601	**0.001**	yes	Not available
Total glucosinolates	12.3±4.1	13.7±4.2	15.8±4.8[2]	0.179	**<0.001**	yes	yes

Analyte	N90-740 (A)	MS11 unsprayed (B)	MS11 sprayed (C)	p-value A vs B	p-value[1] A vs C	Within tolerance interval?	Within literature range?
(µmol/g dw)							
Insoluble tannins (%dw)	0.403±0.095	0.455±0.110	0.458±0.125	**0.027**	0.109	yes	Not available

[1] p-values indicating significant differences are bolded and underlined.
[2] mauve shading represents MS11 means that are significantly lower than the N90-740 means while orange shading represents MS11 means that are significantly higher.

5.4 Conclusions of the compositional analyses

Detailed compositional analyses were done to establish the nutritional adequacy of seed from MS11 and to characterise any unintended compositional changes. Seed samples were analysed for proximates, fibre, fatty acids, amino acids, minerals, vitamins, anti-nutrients (phytic acid, tannins and sinapine) and glucosinolates. In total, 87 analytes were considered of which 30 had negligible levels that precluded inclusion in a statistical analysis.

The levels of each analyte from glufosinate-sprayed and unsprayed MS11 were compared to levels in: a) the non-GM parental line, N90-740, b) six non-GM commercial reference lines and c) levels recorded in the literature.

Of the 57 analytes considered, only two – gluconapin and insoluble tannins - were significantly different in a comparison between unsprayed MS11 and the control. In both instances, the levels fell within the 95% tolerance interval generated from the reference lines.

In contrast to this, in the comparison of analytes between sprayed MS11 and the control, 31 analytes were significantly different - most being higher in MS11 than in the control. As discussed in Section 6.2, it was expected there would be little consistency in analyte levels between the seeds from sprayed and unsprayed MS11 because of the different pollen sources used to fertilise the sprayed MS11 plants. However, in all cases the levels in the sprayed MS11 seed were within the tolerance interval. In six instances the levels exceeded the literature range - but in four of these, the levels in the non-GM control were also higher than the literature range.

In conclusion, seed from MS11, whether from unsprayed MS11 plants or plants sprayed with glufosinate is compositionally equivalent to seed from conventional canola varieties.

6 Nutritional impact

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

Where a GM food has been shown to be compositionally equivalent to conventional varieties already in the food supply, as is the case with MS11, this indicates the food will be as nutritionally adequate as those varieties. Therefore, dietary modelling is not required and feeding studies using target livestock species are unlikely to contribute any further useful information (see e.g. OECD 2003; Bartholomaeus et al. 2013; Herman and Ekmay 2014).

7 References[24]

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215(3):403–410

AOF (2016) Australian oilseeds. Australian Oilseeds Federation. http://www.australianoilseeds.com/

Astwood JD, Fuchs RL (1996) Allergenicity of foods derived from transgenic plants. Monographs in Allergy 32 (Highlights in Food Allergy):105–120

Barthet VJ, Daun JK (2005) Effect of sprouting on the quality and composition of canola seed and oil. Journal of the American Oil Chemists' Society 82:511–517

Bartholomaeus A, Parrott W, Bondy G, Walker K (2013) The use of whole food animal studies in the safety assessment of genetically modified crops: Limitations and recommendations. Critical Reviews in Toxicology 43(S2):1–24

Baxevanis AD (2005) Assessing Pairwise Sequence Similarity: BLAST and FASTA. Ch 11 In: Baxevanis AD, Ouellette BFF (eds) Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. John Wiley & Sons, Inc., p. 295–324

Bogdanov S (2016) Pollen: production, nutrition and health: a review. Bee Product Science. www.bee-hexagon.net/files/file/fileE/Health/PollenBook2Review.pdf

Bonnardeaux J (2007) Uses for canola meal. Department of Agriculture and Food, Government of Western Australia,

Botterman J, Gossele V, Thoen C, Lauwereys M (1991) Characterization of phosphinothricin acetyltrasnferase and C-terminal enzymatically active fusion proteins. Gene 102(1):33–37

Bradbury JF (1986) Guide to plant pathogenic bacteria. Cambridge News Ltd, Aberystwyth

Brand TS, Smith N, Hoffman LC (2007) Anti-nutritional factors in canola produced in the Western and Southern Cape areas of South Africa. South African Journal of Animal Science 37(1):45–50

Bushey DF, Bannon GA, Delaney BF, Graser G, Hefford M, Jiang X, Lee TC, Madduri KM, Pariza M, Privalle LS, Ranjan R, Saab-Rincon G, Schafer BW, Thelen JJ, Zhang JXQ, Harper MS (2014) Characteristics and safety assessment of intractable proteins in genetically modified crops. Regulatory Toxicology and Pharmacology 69:154–170

Campbell L, Rempel CB, Wanasundara JPD (2016) Canola/rapeseed protein: future opportunities and directions - workshop proceedings of IRC 2015. Plants 5(2):7 pp

CCC (2012) Steps in oil and meal processing. Canola Council of Canada. http://www.canolacouncil.org/oil-and-meal/what-is-canola/how-canola-is-processed/steps-in-oil-and-meal-processing/

CCC (2014) Growth stages of the canola plant. Canola Council of Canada. http://www.canolacouncil.org/canola-encyclopedia/crop-development/growth-stages/

CERA (2011) A review of the environmental safety of the PAT protein. Environmental Biosafety Research 10:73–101. DOI: 10.1051/ebr/2012004.

Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Correa M, Da Silva C, Just J, Falentin C, and others (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science 345(6199):950–953. doi: 10.1126/science.1253435.

Chawla G, Ranjan C (2016) Principle, instrumentation, and application of UPLC: a novel technique of liquid chromatography. Open Chemistry Journal 3:1–16

Codex (2009) Foods derived from modern biotechnology, 2nd edition. Codex Alimentarius Commission, Joint FAO/WHO Food Standards Programme, Food and Agriculture Organization of the United Nations, Rome. http://www.codexalimentarius.org/standards/thematic-publications/

[24] All website references were current as at 5 July 2017

Condon C, Putzer H (2002) The phylogenetic distribution of bacterial ribonucleases. Nucleic Acids Res 30(24):5339–5346

De Block M, De Brouwer D, Tenning P (1989) Transformation of *Brassica napus* and *Brassica oleracea* using *Agrobacterium tumefaciens* and the expression of the *bar* and *neo* genes in the transgenic plants. Plant Physiol 91:694–701

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

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

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: Version II. Plant Molecular Biology Reporter 1(4):19–21

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

Dhaese P, de Greve H, Gielen J, Seurinck J, van Montagu M, Schell J (1983) Identification of sequences involved in the polyadenylation of higher plant nuclear transcripts using *Agrobacterium* T-DNA genes as models. The EMBO Journal 2(3):419–426

Environment Canada and Health Canada (2015) Final screening assessment for 'DSL *Bacillus licheniformis/subtilis* group'. http://publications.gc.ca/site/archivee-archived.html?url=http://publications.gc.ca/collections/collection_2016/eccc/En14-228-2015-eng.pdf

FAOSTAT3 (2015) Online database of the Food and Agriculture Organization of the United Nations. http://faostat3.fao.org/browse/T/TP/E

FSANZ (2002) Application A372 - Food derived from herbicide-tolerant canola lines Topas 19/2 and T45 and herbicide-tolerant and pollination controlled canola lines Ms1, Ms8, Rf1, Rf2 and Rf3. Report prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/code/applications/pages/applicationa372glufo948.aspx

FSANZ (2006) Application A533 - Food derived from herbicide-tolerant cotton line LL25. Report prepared by Food Standards Australia New Zealand. http://www.foodstandards.gov.au/code/applications/pages/applicationa533foodd2440.aspx

Goodman RE (2006) Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. Molecular Nutrition and Food Research 50:655–660

Guillet V, Lapthorn A, Hartley RW, Mauguen Y (1993) Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar. Structure 1(3):165–176

Harrigan GG, Glenn KC, Ridley WP (2010) Assessing the natural variability in crop composition. Regulatory Toxicology and Pharmacology 58:S13–S20; doi:10.1016/j.yrtph.2010.08.023

Hartley RW (1988) Barnase and barstar. Expression of its cloned inhibitor permits expression of a cloned ribonuclease. Journal of Molecular Biology 202(4):913–915

Hartley RW (1989) Barnase and barstar: two small proteins to fold and fit together. Trends in Biochemical Sciences 14(11):450–454. DOI: http://dx.doi.org/10.1016/0968-0004(89)90104-7.

Hartley RW, Smeaton JR (1973) On the reaction between the extracellular ribonuclease of *Bacillus amyloliquefaciens* (Barnase) and its intracellular inhibitor (Barstar). The Journal of Biological Chemistry 248(16):5624–5626

Herman RA, Ekmay R (2014) Do whole-food animal feeding studies have any value in the safety assessment of GM crops? Regulatory Toxicology and Pharmacology 68:171–174

Herman RA, Woolhiser MM, Ladics GS, Korjagin VA, Schafer BW, Storer NP, Green SB, Kan L (2007) Stability of a set of allergens and non-allergens in simulated gastric fluid. International Journal of Food Sciences and Nutrition 58:125–141

Hérouet C, Esdaile DJ, Mallyon BA, Debruyne E, Schulz A, Currier T, Hendrickx K, van der Klis R-J, 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

Huby RDJ, Dearman RJ, Kimber I (2000) Why are some proteins allergens? Toxicological Sciences 55:235–246. https://doi.org/10.1093/toxsci/55.2.235

Kim S-I, Veena, Gelvin SB (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. The Plant Journal 51:779–791

Kimber I, Kerkvliet NI, Taylor SL, Astwood JD, Sarlo K, Dearman RJ (1999) Toxicology of protein allergenicity: prediction and characterization. Toxicological Sciences 48:157–162

Krebbers E, Seurinck J, Herdies L, Cashmore AR, Timko MP (1988) Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. Plant Molecular Biology 11:745–749

Kützner HJ (1981) The Family Streptomycetaceae. In: Starr MP, Stolp H, Trüper HG, Ballows A, and Schlegel HG (eds) The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria. Springer Verlag, Berlin, p. 2028–2090

Lee R-Y, Reiner D, Dekan G, Moore AE, Higgins TJV (2013) Genetically modified a-amylase inhibitor peas are not specifically allergenic in mice. PLoS ONE 8 (online access)(1):e52972

Mariani C, De Beuckeleer M, Truettner J, Leemans J, Goldberg RB (1990) Induction of male sterility in plants by a chimaeric ribonuclease gene. Nature 347:737–741. DOI: 10.1038/347737a0.

Mariani C, Gossele V, De Beuckeleer M, De Block M, Goldberg RB, De Greef W, Leemans J (1992) A chimaeric ribonuclease-inhibitor gene restores fertility to male sterile plants. Nature 357(6377):384–387. DOI: 10.1038/357384a0.

Marwede V, Schierholt C, Möllers C, Becker HC (2004) Genotype x environment interactions and heritability of tocopherol contents in canola. Crop Science 44:728–731

Metcalfe DD, Astwood JD, Townsend R, Sampson HA, Taylor SL, Fuchs RL (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. Critical Reviews in Food Science and Nutrition 36 Suppl:S165–S186

Michiels, F., Botterman, J., Cornelissen, M (2000) Method to obtain male sterile plants. Patent US 6025546 Patent Application PCT/EP96,00722

Miletich JP, Broze Jr. GJ (1990) b Protein C is not glycosylated at asparagine 329. The Journal of Biological Chemistry 265:11397–11404

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

OECD (1999) Consensus document on general information concerning the genes and their enzymes that confer tolerance to phosphinothricin herbicide. ENV/JM/MONO(99)13. https://www.oecd.org/env/ehs/biotrack/46815628.pdf

OECD (2001) Consensus document on key nutrients and key toxicants in low erucic acid rapeseed (canola). ENV/JM/MONO(2001)13. Organisation for Economic Co-operation and Development, Paris. http://www.oecd.org/science/biotrack/consensus-document-for-work-on-safety-novel-and-foods-feeds-plants.htm

OECD (2003) Considerations for the safety assessment of animal feedstuffs derived from genetically modified plants. ENV/JM/MONO(2003)10. Organisation for Economic Co-operation and Development, Paris. http://www.oecd.org/officialdocuments/displaydocumentpdf?cote=ENV/JM/MONO(2003)10&doclanguage=en

OECD (2011) Revised consensus document on compositional considerations for new varieties of low erucic acid rapeseed (canola): key food and feed nutrients, anti-nutrients and toxicants. ENV/JM/MONO(2011)55. Organization for Economic Co-operation and Development, Paris. http://www.oecd.org/document/22/0,3746,en_2649_34385_46808854_1_1_1_1,00.html

Orlando R, Yang Y (1998) Analysis of Glycoproteins. Ch 9 In: Larsen BS, McEwen CN (eds) Mass Spectrometry of Biological Materials. 2nd ed, Marcel Dekker, p. 216–246

Pearson WR (2000) Flexible sequence similarity searching with the FASTA3 program package. Ch 10 In: Misener S, Krawetz SA (eds) Methods in Molecular Biology, Volume 132: Bioinformatics Methods and Protocols. Human Press Inc., Totowa, NJ, p. 185–219

Pearson WR (2013) Selecting the right similarity-scoring matrix. Current Protocols in Bioinformatics 43:3.5.1–3.5.9

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proceedings of the National Academy of Sciences 85(8):2444–2448

Prescott VA, Campbell PM, Moore A, Mattes J, Rothenberg ME, Foster PS, Higgins TJV, Hogan SP (2005) Transgenic expression of bean a-amylase inhibitor in peas results in altered structure and immunogenicity. J Agric Food Chem 53:9023–9030

Priest FG, Goodfellow M, Shute LA, Berkeley RCW (1987) *Bacillus amyloliquefaciens* sp. nov., nom. rev. International Journal of Systematic Bacteriology 37(1):69–71

Pritchard FM, Eagles HA, Norton RM, Salisbury PA, Nicolas M (2000) Environmental effects on seed composition of Victorian canola. Australian Journal of Experimental Agriculture 40:679–685

Schabenberger O (2004) Mixed model influence diagnostics. Paper 189-29. Proceedings of the 29th Annual SAS Users Group International Conference (SUGI 29), May 9 - 12, 2004, Montreal. http://www2.sas.com/proceedings/sugi29/189-29.pdf

Seberry DE, Mailer RJ, Parker PA (2007) Quality of Australian Canola 2006. Volume No.13. Australian Oilseeds Federation, and NSW Department of Primary Industries. http://www.australianoilseeds.com/commodity_groups/canola_association_of_australia/quality_of_australian_canola

Seurinck J, Truettner J, Goldberg RB (1990) The nucleotide sequence of an anther-specific gene. Nucleic Acids Res 18(11):3403

Sherlock O, Dobrindt U, Jensen JB, Munk Vejborg R, Klemm P (2006) Glycosylation of the self-recognizing *Escherichia col*i Ag43 autotransporter protein. Journal of Bacteriology 188(5):1798–1807

Spragg J, Mailer R (2007) Project Code: 1B-103-0506: Canola Meal Value Chain Quality Improvement. A final report prepared for Australian Oilseeds Federation and Pork Cooperative Research Centre, JCS Solutions Pty Ltd, Victoria,

Szmigielska AM, Schoenau JJ, Levers V (2000) Determination of glucosinolates in canola seed using anion exchange membrane extraction combined with the high-pressure liquid chromatography detection. J Agric Food Chem 48:4487–4491

Thomas K, Bannon G, Hefle S, Herouet C, Holsapple M, Ladics G, MacIntosh S, Privalle L (2005) In silico methods for evaluating human allergenicity to novel proteins: International Bioinformatics Workshop Meeting Report 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, 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

Thomas K, Aalbers M, Bannon GA, Bartels M, Dearman RJ, Esdaile DJ, Fu T-J, Glatt CM, Hadfield N, Hatzos C, Hefle SL, Heylings JR, Goodman RE, Henry B, Herouet C, Holsapple M, Ladics GS, Landry TD, MacIntosh SC, Rice EA, Privalle LS, Steiner HY, Teshima R, Van Ree R, Woolhiser M, Zawodny J (2004) A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. Regulatory Toxicology and Pharmacology 39:87–98

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

Tzfira T, Li J, Lacroix B, Citosky V (2004) *Agrobacterium* T-DNA integration: molecules and models. Trends in Genetics 20(8):375–383

U.S.Pharmacopeia (1990) United States Pharmacopeia, Volume 22. United States Pharmacopeia Convention, Inc, Rockville, MD

Van Larebeke N, Engler G, Holsters S, Van Den Elsacker S, Zaenen I, Schilperoort RA, Schell J (1974) Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. Nature 252:169–170. DOI: 10.1038/252169a0 .

Wang J, Lydiate DJ, Parkin IAP, Falentin C, Delourme R, Carion PWC, King G (2011) Integration of linkage maps for the amphidiploid *Brassica napus* and comparative mapping with *Arabidopsis* and *Brassica rapa*. BMC Genomics 12:101. DOI: 10.1186/1471-2164-12-101

Wang Y, McAllister TA, Pickard MD, Xu Z, Rode LM, Cheng K-J (1999) Effect of micronizing full fat canola seed on amino acid disappearance in the gastrointestinal tract of dairy cows. Journal of Dairy Science 82:537–544

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