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

30 October 2023
267-23

Supporting document 1

Safety assessment – Application A1276

Food derived from herbicide-tolerant soybean line MON94313

Executive summary

Background

Application A1276 seeks approval for the sale and use of food derived from soybean line MON94313 that has been genetically modified (GM) for tolerance to the herbicides glufosinate, dicamba, 2,4-dichlorophenoxyacetic acid (2,4-D) and mesotrione.

Tolerance to these herbicides in MON94313 is conferred thorough expression of the:

- *dmo* gene from *Stenotrophomonas maltophilia*, which encodes a dicamba mono-oxygenase (DMO) protein and provides tolerance to dicamba
- *pat* gene from *Streptomyces viridochromogenes*, which encodes a phosphinothricin-N-acetyltransferase (PAT) protein and provides tolerance to glufosinate
- *ft_t.1* gene that is a modified version of the *RdpA* gene from *Sphingobium herbicidovorans*, which encodes a 2,4-D and FOPs dioxygenase protein (FT_T.1). This protein provides tolerance to 2,4-D
- *TDO* gene from *Oryza sativa*, which encodes a triketone dioxygenase (TDO) protein and provides tolerance to mesotrione.

FSANZ has previously assessed the PAT and DMO proteins, as well as the FT_T protein, which is 98% identical to FT_T.1. This is the first time FSANZ has assessed the TDO protein.

This safety assessment addresses food safety and nutritional issues associated with GM food. It therefore does not address:

- risks related to the environmental release of GM plants used in food production
- risks to animals that may consume feed derived from GM plants
- the safety of food derived from the non-GM (conventional) plant.

History of use

Soybean has a long history of safe use in the food supply. Soybean oil is widely used as cooking oil and as an ingredient in a wide range of manufactured products. Soybean grains are also used to make soy milk, soy sauce, soy lecithin and meat substitutes such as tofu and tempeh.

Molecular characterisation

The genes encoding PAT (*pat*), DMO (*dmo*), FT_T.1 (*ft_t.1*) and TDO (*TDO*) were introduced into soybean line MON94313 via *Agrobacterium*-mediated transformation. Detailed molecular analyses indicate a single copy of each of the four gene cassettes is present at a single insertion site in the MON94313 genome. There are no extraneous plasmid sequences or antibiotic resistance genes present in this line.

The introduced genetic elements were shown by molecular techniques and phenotypic analyses to be present within a single locus and stably inherited across multiple generations.

Characterisation and safety assessment of new substances

All four novel proteins (PAT, DMO, FT_T.1 and TDO) are expressed throughout the plant (except TDO which was not detected in root tissue). Expression levels are low in grain.

A range of characterisation studies confirmed the identity of the plant-expressed FT_T.1 and TDO proteins and their equivalence with the corresponding proteins produced in a bacterial expression system. The plant-derived and bacterially-derived proteins had the expected molecular weight, amino acid sequence, immunoreactivity, lack of glycosylation and enzyme activity. Characterisation studies confirmed that the PAT and DMO proteins were identical to proteins previously assessed by FSANZ.

All proteins are rapidly degraded and heat inactivated, based on studies submitted with this application and/or conclusions from previous assessments. Bioinformatics studies on all four proteins confirmed the lack of any significant amino acid sequence similarity to known protein toxins or allergens. Taken together, the evidence supports the conclusion that PAT, DMO, FT_T.1 and TDO are not toxic or allergenic to humans.

Herbicide metabolites

For PAT and DMO, the metabolic profiles resulting from the novel protein/herbicide interaction have been established through a significant history of use. There are no concerns that the spraying of soybean line MON94313 with glufosinate, dicamba, 2,4-D or mesotrione would result in the production of metabolites that are not also produced in non-GM crops sprayed with the same herbicides and already used in the food supply.

Compositional analyses

Detailed compositional analyses were performed on MON94313. Statistically significant differences in mean values were found between grain from MON94313 and the control for 7 of the 49 analytes evaluated, however these differences were within the range established for existing commercial non-GM soybean varieties. Overall, the compositional data support the conclusion that there are no biologically significant differences in the levels of key constituents in grain from MON94313 compared to non-GM soybean varieties available on the market.

Conclusion

No potential public health and safety concerns have been identified in the assessment of herbicide-tolerant soybean line MON94313. On the basis of the data provided in the present application and other available information, food derived from MON94313 is considered to be as safe for human consumption as food derived from non-GM soybean varieties.

Table of contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
µg	Microgram(s)
2,4-D	2,4-dichlorophenoxyacetic acid
aa	Amino acid
AFSI	Agriculture and Food Systems Institute
BBCH	Biologische Bundesanstalt, Bundessortment und Chemische Industrie
BLAST	Basic Local Alignment Search Tool
bp	Base pair
CCI	Confidential Commercial Information
COMPARE	COMprehensive Protein Allergen Resource
CTP	Chloroplast transit peptide
DCSA	3,6-dichlorosalicylic acid
DMO	Dicamba mono-oxygenase
DNA	Deoxyribonucleic acid
DW/dw	Dry weight
ECL	Enhanced chemiluminescence
EFSA	European Food Safety Authority
ELISA	Enzyme Linked Immunosorbent Assay
FAO	Food and Agriculture Organization of the United Nations
FSANZ	Food Standards Australia New Zealand
GM	Genetically modified
h	Hours
HPLC	High performance liquid chromatography
HPPD	4-hydroxyphenylpyruvate dioxygenase
ILSI	International Life Sciences Institute
kDa	Kilodalton
kg	Kilogram
LB	Left border

Abbreviation	Definition
LOD	Limit of detection
LOQ	Limit of quantification
mg	Milligram(s)
MT	Million tonnes
NCBI	National Centre for Biotechnology Information
ng	Nanogram(s)
NGS	Next Generation Sequencing
OECD	Organisation for Economic Cooperation and Development
OGTR	Office of the Gene Technology Regulator
ORF	Open reading frame
PAT	Phosphinothricin-N-acetyl transferase
PCR	Polymerase chain reaction
ppm	Parts per million
RB	Right border
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	Standard error
T-DNA	Transfer DNA
TDO	Triketone deoxygenase
USDA	United States Department of Agriculture

1 Introduction

FSANZ has received an application from Bayer CropScience Proprietary Limited to vary Schedule 26 in the Australia New Zealand Food Standards Code (the Code). The variation is to include food derived from the genetically modified (GM) soybean line MON94313, with the OECD Unique Identifier MON-94313-8. This soybean line is tolerant to the herbicides dicamba, glufosinate, 2,4-D, and mesotrione.

Tolerance to glufosinate is achieved with the expression of the phosphinothricin-N-acetyltransferase (PAT) protein, encoded by the *pat* gene from the bacterium *Streptomyces viridochromogenes*. The PAT protein has been assessed by FSANZ in numerous previous applications.

Tolerance to dicamba is achieved with the expression of the dicamba mono-oxygenase (DMO) protein, encoded by the *dmo* gene from the bacterium *Stenotrophomonas maltophilia*. The DMO protein has been assessed by FSANZ in 5 previous applications[1].

Tolerance to 2,4-D is conferred by the expression of the FT_T.1 protein, encoded by a modified version of the *R-2,4-dichlorophenoxypropionate dioxygenase (Rdpa*) gene from the bacterium *Sphingobium herbicidovorans*. The FT_T.1 protein is 98% identical in amino acid sequence to FT_T which was assessed in Application A1192 (FSANZ 2020). Both proteins share the same mechanism of action.

Mesotrione tolerance is conferred by expression of the triketone dioxygenase (TDO) protein, encoded by the *TDO* gene from *Oryza sativa* (rice). This is the first time FSANZ has assessed the TDO protein.

If approved, food derived from soybean line MON94313 may enter the Australian and New Zealand food supply as imported food products.

2 History of use

2.1 Host organism

The host organism is soybean (*Glycine Max* (L.)), from the family Leguminosae. The non-GM soybean variety A3555 was used as the parental variety for the genetic modification described in this application and served as the conventional control for the purposes of comparative assessment with MON94313.

Soybean has a long history of safe human and animal consumption, having first been cultivated in northern China as early as 5000 years ago (Liu 2004). The commodity is the leading oilseed crop in the world, with total global production reaching 371.7 MT[2] in 2021 (FAOSTAT 2023).

Soybean production in Australia is comparatively minor, totalling 0.040 MT in 2021 (FAOSTAT 2023), while New Zealand has no commercial soybean cultivation. Australia and New Zealand are net importers of soybeans, with 2325.03 tonnes and 3044.08 tonnes imported respectively in 2021 (FAOSTAT 2023).

Figure 1 shows the major soybean producing countries in the world. In many of these

[1] Applications A1063 (MON87798 soybean); A1080 (MON88701 cotton); A1118 (MON87419 corn); A1192 (MON87429 corn); A1216 (MON94100 canola)
[2] Million tonnes

countries, herbicide-tolerant GM soybean varieties are a major contributor to total production – for example, in 2022, 95% of all soybeans planted in the US were GM[3]. No GM soybean lines are currently grown commercially in Australia[4].

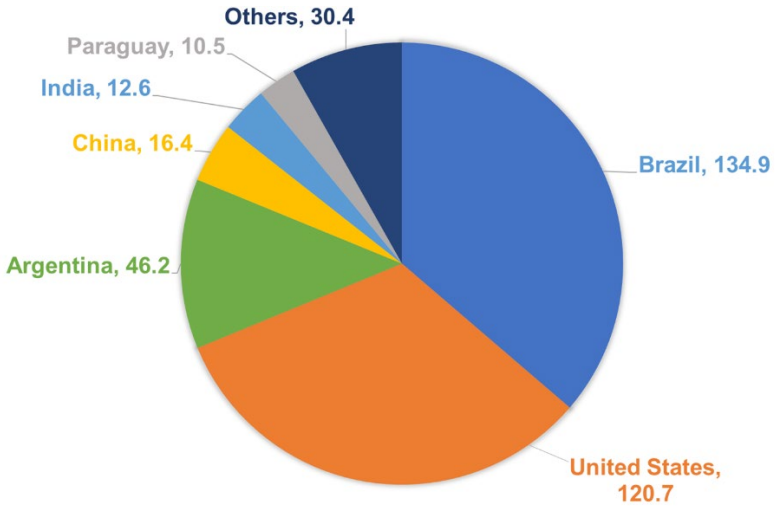


Figure 1: *Major soybean producing countries in 2021 (in MT). Data obtained from FAOSTAT (2023)*

Whole soybeans are used to produce soy milk, tofu, and soy sprouts, as well as fermented foods including miso, soy sauce, natto and tempeh. Soybeans may also be eaten with minimal processing, for example in the Japanese food edamame, in which immature soybeans are boiled whole in their pods and served with salt. Soybean grains are processed into two major products: oil and meal. Soybean oil is the second most consumed vegetable oil worldwide and accounts for 29% of global vegetable oil consumption (American Soybean Association 2023). It is used in a variety of manufactured foods, including cooking oil, shortening, margarine, salad dressings, frozen desserts and confectionery products. Soybean meal is a good source of protein and is primarily processed into livestock feed (pet and poultry food) and protein products such as soy flour, concentrates and isolates.

Soybeans are also a rich source of a number of bioactive phytochemicals, such as isoflavones and tocopherols (vitamin E) (Liu 2004), and are used as a source of these compounds for dietary supplements. Unprocessed (raw) soybean grain products are not suitable for food uses, due to the presence of anti-nutrients, such as phytic acid and lectins (OECD 2012). The heat applied during processing inactivates these anti-nutrients.

2.2 Donor organisms

2.2.1 Streptomyces viridochromogenes

The source of the *pat* gene is the bacterium *Streptomyces viridochromogenes.* This Gram-positive, spore-forming species is widespread in the environment and is not pathogenic to humans or animals. The *pat* gene produces a protein that is structurally and functionally equivalent to the protein encoded by the *bar* gene from the closely related species *S. hygroscopicus* (Wehrmann et al. 1996). While *S. viridochromogenes* and *S. hygroscopicus* themselves do not have a history of use in food, the *pat* and *bar* genes have been used to confer glufosinate tolerance in food-producing crops for almost three decades with no toxicity

[3] For more information please see USDA Economic Research Service: https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-u-s/
[4] Information on approved commercial releases of GM crops in Australia can be found on the website of the Office of the Gene Technology Regulator: https://www.ogtr.gov.au/

or allergenicity concerns (ILSI 2016).

2.2.2 Stenotrophomonas maltophilia

The *dmo* gene is derived from the bacterium *Stenotrophomonas maltophilia* strain DI-6, isolated from soil at a dicamba manufacturing plant (Krueger et al. 1989). *S. maltophilia* is an aerobic, Gram-negative bacterium commonly present in aquatic environments and soil, as well as often being associated with plant roots (Ryan et al. 2009). *S. maltophilia* is widespread in a variety of foods, ranging from to drinking water (Munsch-Alatossava and Alatossava 2006; Todaro et al. 2011; Gomes et al. 2018).

S. maltophilia can act as an opportunistic human pathogen under favourable conditions: instances of infection generally occur in immunocompromised individuals in hospital settings (Mukherjee and Roy 2016; Lira et al. 2017). Despite these instances, *S. maltophilia* is not known to possess specific virulence genes and is found in healthy individuals without causing illness (Lira et al. 2017).

The *dmo* gene in MON94313 has been manipulated through standard DNA cloning methods subsequent to its isolation, meaning that extraneous material from *S. maltophilia* would not have been transferred to MON94313. In addition, *S. maltophilia* has a history of safe use as the donor organism of the *dmo* gene and been assessed in previous applications to FSANZ[5].

2.2.3 Sphingobium herbicidovorans

The *ft_t.1* gene is a modified version of the *RdpA* gene from the common soil bacterium *Sphingobium herbicidovorans*. *S. herbicidovorans* is a strictly aerobic, Gram-negative bacterium commonly isolated from natural environments such as soil and water (Chaudhary et al. 2017). Members of the genus *Sphinogbium* have also been isolated from food including corn (Rijavec et al., 2007), rice (Videira et al., 2010), papaya (Thomas et al., 2007), tomato (Enya et al., 2007) and fermented milk (Bauer et al., 2009).

Given the widespread distribution of *Sphingobium* species in nature, it is likely that humans and animals have been exposed to these species without any known safety concerns. *S. herbicidovorans* has been previously assessed as a donor organism in two applications to FSANZ[6].

2.2.4 Oryza sativa

The *TDO* gene is a version of the HIS1 gene from *Oryza sativa* (rice) that has been codon-optimised for expression in soybean. The encoded protein, TDO, is a triketone deoxygenase. In certain rice cultivars, this group of enzymes confers natural tolerance to mesotrione and other β-triketone herbicides (Maeda et al. 2019).

Rice has a long and extensive history of safe human consumption, and serves as a primary food source for more than one third of the world's population (Khush 1997; OECD 2016).

2.2.5 Other organisms

Genetic elements from several other organisms have been used in the genetic modification of MON94313 (refer to Table 1 and Appendix 2). These genetic elements are non-coding sequences that are used to regulate the expression of the inserted genes.

[5] Applications A1063 (MON87798 soybean); A1080 (MON88701 cotton); A1118 (MON87419 corn); A1192 (MON87429 corn); A1216 (MON94100 canola)
[6] Applications A1042 (DAS-40278-9 corn) and A1192 (MON87429 corn)

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.

3.1 Transformation method

To create the MON94313 soybean line, the conventional soybean variety A3555 was transformed using the binary plasmid vector PV-GMHT529103. This plasmid contains two separate transfer DNAs (T-DNAs): T-DNA I, which contains the four trait expression cassettes, and T-DNA II, which contains the selectable marker cassettes *splA* and *aadA* (Figure 2). While both T-DNAs were inserted into the soybean genome during transformation, plants that contained only T-DNA I, and not T-DNA II, were isolated for further development. The transformation and development methodology is outlined in the flowchart in Appendix 1 and summarised below.

Transformation of A3555 was achieved by co-culturing meristem explants with *Agrobacterium tumefaciens* containing the PV-GMHT529103 plasmid. The meristem explants were then placed on selective media containing spectinomycin, carbenicillin, cefotaxime and timentin. Spectinomycin inhibits the growth of untransformed plant cells, while carbenicillin, cefotaxime and timentin suppress the growth of excess *Agrobacterium*.

Putative transformants (R0) with normal phenotypes were selected and screened for the presence of the T-DNA I insert unlinked to T-DNA II insert, and the absence of the vector backbone. Selected R0 plants were transferred to soil and self-pollinated to produce R1 seed, then evaluated phenotypically for the presence of *splA* (which produces a wrinkled seed phenotype) and by polymerase chain reaction (PCR) to identify the presence of the *aadA* gene. Traditional breeding, segregation and screening were used to eliminate plants containing the T-DNA II insert. R1 plants that were homozygous for the T-DNA I insert were selected for further development. Following the evaluation of insert integrity, gene expression, phenotypic characteristics and agronomic performance, soybean line MON94313 was selected.

3.2 Detailed description of inserted DNA

Soybean line MON94313 contains the T-DNA I from the PV-GMHT529103 plasmid (Figure 2) and includes the *dmo*, *pat*, *ft_t.1* and *TDO* expression cassettes. Information on these four expression cassettes is summarised in Table 1. Additional detail, including the expression cassettes contained in the T-DNA II region from PV-GMHT529103 used as selectable markers during transformation, and intervening sequences used to assist with cloning, mapping and sequence analysis, can be found in Appendix 2.

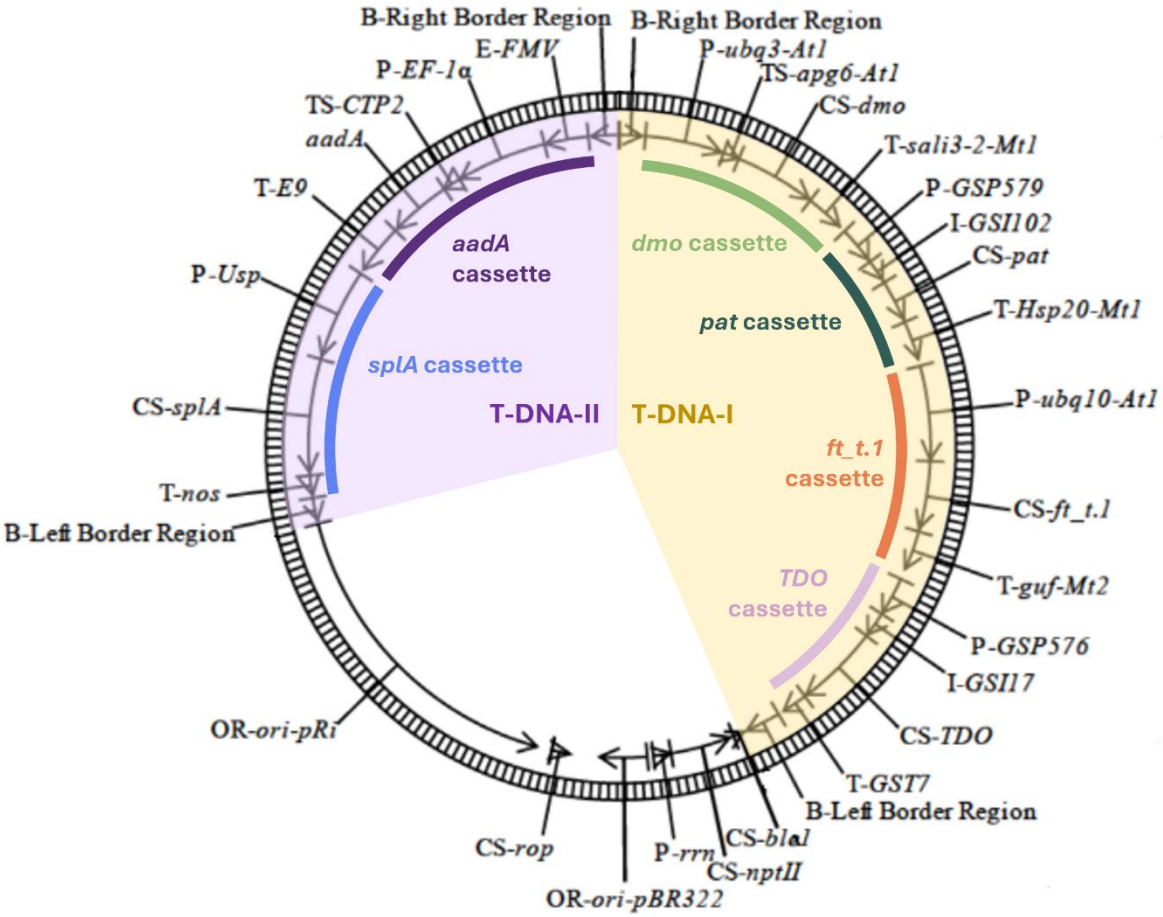


Figure 2: *Map of plasmid PV-GMHT529103 (24,549 bp). The T-DNA I region comprising the dmo, pat, ft_t.1 and TDO expression cassettes is shaded yellow. The T-DNA II region comprising the splA and aadA selectable marker cassettes, which does not form part of the insert in MON94313, is shaded purple.*

Table 1: Expression cassettes contained in the MON94313 insert

	Promoter / Regulatory / Targeting sequences	Coding sequence	Terminator(s)	Function
***dmo* cassette**	Promoter, leader and intron from the polyubiquitin (*ubq3*) gene from *Arabidopsis thaliana* Targeting sequence of the *APG6* gene from *Arabidopsis thaliana* (directs transport of the DMO protein to the chloroplast)	Codon-optimised dicamba mono-oxygenase (*dmo*) gene from *Stenotrophomonas maltophilia*	3′-UTR sequence from an aluminium-induced Sali3-2 protein from *Medicago trunculata* (barrel medic)	Confers tolerance to dicamba
***pat* cassette**	Promoter, 5′-UTR and intron sequences developed from multiple sequences from *Arabidopsis thaliana*	Codon-optimised phosphinothricin N-acetyltransferase (*pat*) gene from *Streptomyces viridochromogenes*	3′-UTR sequence from a putative heat shock protein gene (*Hsp20*) from *Medicago trunculata* (barrel medic)	Confers tolerance to glufosinate
***ft_t.1* cassette**	Promoter, leader and intron from the polyubiquitin (*ubq10*) gene from *Arabidopsis thaliana*	Modified version of the R-2,4 dichlorophenoxypropionate dioxygenase (*RdpA*) gene from *Sphingobium herbicidovorans*	3′-UTR sequence from an expressed gene of unknown function from *Medicago trunculata* (barrel medic)	Confers tolerance to 2,4-D
***TDO* cassette**	Promoter, 5′-UTR and intron sequences developed from multiple sequences from *Arabidopsis thaliana*	Codon-optimised triketone dioxygenase (*TDO*) gene from *Oryza sativa* (rice)	3′-UTR sequence developed from multiple 3′-UTR sequences from *Zea mays* (maize)	Confers tolerance to mesotrione

3.3 Development of the soybean line from the original transformant

A breeding program was undertaken for the purpose of:

- obtaining generations suitable for analysing the characteristics of soybean line MON94313; and
- ensuring that the MON94313 event is incorporated into elite lines for commercialisation.

Table 2 indicates the specific generations and controls used in the characterisation of MON94313.

Table 2: MON94313 generations used for various analyses

Analysis	Section	Generation(s) used	Comparators
Number of integration sites	Section 3.4.1	R3	A3555
Absence of backbone and other sequences	Section 3.4.2	R3	A3555
Insert integrity and site of integration	Section 3.4.3	R3	A3555
Genetic stability	Section 3.4.4.1	R3, R4, R5, R6, R7	A3555
Mendelian inheritance	Section 3.4.4.2	F2, F3, F4	N/A
Expression of phenotype over several generations	Section 3.4.4.2	R3, R4, R5, R6, R7	A3555
Compositional analysis	Section 5	R5	A3555

3.4 Characterisation of the inserted DNA and site(s) of insertion

A range of analyses were undertaken to characterise the genetic modification in MON94313. These analyses focused on the nature and stability of the insertion and whether any unintended re-arrangements or products may have occurred as a consequence of the transformation procedure.

Genomic DNA from MON94313 (R3 generation) and from the conventional control (A3555) was sequenced using next generation sequencing (NGS). This method generates ~150 bp short sequence reads which are randomly distributed throughout the genome and in sufficient number to ensure the genomes are covered comprehensively. In addition, the transformation plasmid PV-GMHT529103 was sequenced to serve as a reference. To assess the sensitivity of the NGS method, plasmid DNA was spiked and sequenced. Sufficient sequence reads were obtained to cover the entire genomes of MON94313 and the control, with a depth of coverage ≥75x and an adequate level of sensitivity[7].

3.4.1 Number of integration site(s)

NGS reads from MON94313 (R3) that mapped to the intended T-DNA I insert were analysed and two unique insert-flank junction sites were identified. Each comprised the inserted T-DNA I border sequence joined to a flanking sequence in the soybean genome. This indicates that a single copy of the T-DNA I insert has been integrated into the genome of MON94313

[7] The NGS method was sufficiently sensitive to detect 100% of the spiked plasmid when present at 1/10th of a copy per genome equivalent or greater. Additionally, reads were mapped to a single copy of an endogenous gene and the depth of coverage for MON94313 and the controls was comprehensive.

(Figure 3). As expected, no junction sequences were detected in the A3555 control.

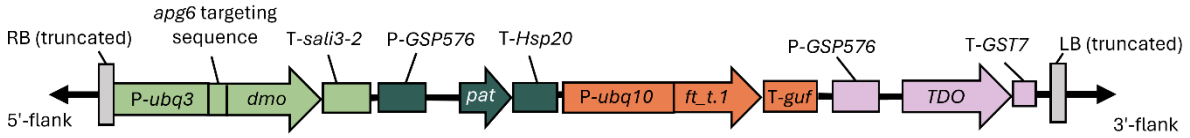


Figure 3*: T-DNA I insert (10,196 bp) present in MON94313.*

3.4.2 Absence of backbone and other sequences

Alignment of NGS reads from MON94313 (R3) and the PV-GMHT529103 transformation plasmid confirmed that the MON94313 genome does not contain any sequences from T-DNA II or the PV-GMHT529103 backbone, including antibiotic resistance genes.

When NGS reads from the A3555 control were aligned with the transformation plasmid sequence, a small number of reads mapped to parts of the T-DNA I and plasmid backbone. This low number of reads does not indicate the presence of backbone sequence in A3555 and is likely due to the presence of environmental bacteria in the genomic DNA prepared for NGS, as has been previously described (Yang et al. 2013; Zastrow-Hayes et al. 2015).

3.4.3 Insert integrity and site of integration

Locus-specific PCR and DNA sequence analysis of seed-derived genomic DNA from MON94313 showed that a single copy of T-DNA I from PV-GMHT529103 was integrated into the host genome and the organisation of the genetic elements within the insert is as expected. Both border regions in the MON94313 insert had small terminal truncations relative to these regions in PV-GMHT529103. These changes would not have a functional impact on the expression of the inserted cassettes. No deletions, insertions, mutations or rearrangements of the expression cassettes were detected. As expected, no T-DNA II elements were present. These results were fully consistent with the NGS dataset.

To examine the T-DNA I insertion site, PCR primers flanking the insertion site were used to amplify genomic DNA from the A3555 conventional control. Comparing the sequence of the PCR product to the sequence generated from the flanking regions of MON94313 identified a 40 base deletion of the soybean genomic DNA that occurred during T-DNA integration. All other flanking sequences in MON94313 were identical to those in the control. Such changes during T-DNA insertion are common during *Agrobacterium*-mediated plant transformation due to double-stranded break repair mechanisms (Salomon and Puchta 1998; Anderson et al. 2016) and would not affect the expression of the *dmo*, *pat*, *ft_t.1*, or *TDO* genes.

3.4.4 Stability of the genetic changes in soybean line MON94313

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 events) over successive generations. Phenotypic stability refers to the expressed trait remaining unchanged over successive generations.

3.4.4.1 Genetic stability

NGS was used to show the genetic stability of the inserted DNA in MON94313 (R3) by evaluating seed-derived DNA from four additional breeding generations of MON94313 (R4, R5, R6, R7). Control genomic DNA was isolated from the non-GM parental line A3555. The analysis showed that the two insert-flank junction sequences present in the R3 generation (section 3.4.1)

were identical in each of the four additional generations. No other junction sequences were present. The consistency of these results across the breeding generations tested demonstrates that the single T-DNA I insert is stably integrated in MON94313.

3.4.4.2 Phenotypic stability

Mendelian inheritance

Since the inserted DNA resides at a single locus within the MON94313 genome, the inserted DNA would be expected to be inherited according to Mendelian inheritance principles. To produce the generations for segregation analysis, homozygous positive R3 plants containing the T-DNA I insert were crossed with a proprietary elite line lacking the insert, producing hemizygous F1 seed. The F1 plants were self-pollinated to produce F2 seed. The inheritance of the T-DNA I insert was assessed by PCR in the F2 generation, as well as in the subsequent F3 and F4 generations. A chi-square (χ^2) analysis was undertaken to confirm the segregation and stability of the insert.

According to Mendelian principles, the predicted segregation ratio in all generations was 1:2:1 (homozygous positive: hemizygous positive: homozygous negative). The results presented in Table 3 demonstrate the expected segregation ratio for each generation. This confirms that the inserted DNA is present at a single locus in the MON94313 genome and is inherited predictably according to Mendelian inheritance rules.

Table 3: Segregation of T-DNA I in three generations of MON94313

Generation	Expected segregation ratio	Observed number of plants (expected number)				Statistical analysis	
		Homozygous Positive	Hemizygous positive	Homozygous Negative	Total	χ^2	*P-value*
F2	1:2:1	68 (65.5)	131 (131)	63 (65.5)	262	0.19	0.909
F3	1:2:1	134 (131.75)	251 (263.5)	142 (131.75)	527	1.43	0.489
F4	1:2:1	76 (72.75)	144 (145.5)	71 (72.75)	291	0.20	0.904

Expressed phenotype over several generations

The expression of the DMO, PAT, FT_T.1 and TDO proteins in five generations of MON94313 (R3, R4, R5, R6 and R7) was examined. Western blot analysis was conducted on seed tissue from each generation, with seed tissue from the conventional line A3555 used as a negative control, and *E. coli*-produced versions of the proteins used as positive controls. In all five breeding generations, the DMO, PAT, FT_T.1 and TDO proteins migrated indistinguishably from the corresponding positive controls on the same Western blot. Some additional lower molecular weight bands in the PAT, FT_T.1, and TDO blots. The applicant attributed these observations to a combination of effects of the soy matrix and nonspecific antibody binding. None of the proteins were detected in the tissue from the conventional control. These data support the conclusion that the DMO, PAT, FT_T.1 and TDO proteins are stably expressed over multiple generations.

3.4.5 Open reading frame (ORF) analysis

A bioinformatic analysis of the MON94313 insert, as well as the flanking DNA regions, was undertaken to identify any novel open reading frames (ORFs) which had been created in

MON94313 as a result of the insertion of T-DNA I, and whether any of these putative peptides have the potential for allergenicity or toxicity.

Sequences spanning the right and left insert-flank junctions of MON94313 were translated *in silico* from stop codon to stop codon (TGA, TAG, TAA) in all six reading frames[8]. Similarly, the entire MON94313 insert DNA was translated in all six reading frames. A total of 8 ORFs were identified that corresponded to putative peptides of eight amino acids (aa) or greater in length from the insert-flank junction sequences. Along with the 6 putative peptides translated from the insert sequence, these were investigated further to determine whether their amino acid sequences showed similarity with known allergen and toxin peptide sequences in established databases.

These analyses are theoretical only, as it is highly unlikely that any of the identified ORFs other than the intended DMO, PAT, FT_T.1 and TDO proteins would be expressed *in planta*.

3.4.5.1 Bioinformatic analysis for potential allergenicity

The 8 theoretical ORFs present in the 5′ and 3′ insert-flank junction sequences and the 6 putative peptides translated from the MON94313 insert sequence were compared to known allergenic proteins listed in the COMprehensive Protein Allergen REsource (COMPARE[9]) database, from the Health and Environmental Science Institute. At the date of the search, there were 2,463 sequences in the allergen database (AD_2022). Sequences were also compared to the GenBank all protein database (PRT_2022), downloaded from the National Centre for Biotechnology Information (NCBI[10]), which contained 184,933,782 sequences at the date of download.

Three types of analyses were performed for this comparison:

(a) A FASTA search algorithm (v36.3.5d) (Pearson and Lipman, 1988) was used to identify alignments between the query sequences and the COMPARE database, using a BLOSUM50 scoring matrix, which identifies blocks of residues with at least 50% sequence identity (Henikoff and Henikoff 1992). Only matches with E-scores of $\leq 1\times10^{-5}$ were considered.
(b) 80-mer sliding window search – a FASTA alignment was performed comparing all contiguous 80 amino acids to the database entries. Only matches of greater than 35% similarity over ≥ 80 amino acids were considered.
(c) 8-mer exact match search – an in-house algorithm was used to identify whether an 8 amino acid peptide match existed between the query sequences and sequences within the allergen database. Only matches of 100% similarity over 8 amino acids were considered.

The alignment of the 8 putative peptides present in the 5' and 3' insert-flank junctions with database sequences did not identify any matches. One sequence – from the translation of frame 4 of the MON94313 insert – produced an 8 contiguous aa match (FCRATLSL) to an allergen in the COMPARE database. This allergen, AKJ77987.1, is described as “Allergen Tri a 43: unknown from *Triticum aestivum*” (wheat). However, there is no appropriate upstream methionine available to serve as a start codon for the peptide containing this match, meaning that the possibility of it being translated *in planta* is remote.

In addition to the allergen database results, five of the six translated insert reading frames yielded alignments with E-scores $\leq 1\times10^{-5}$ when queried against the all protein PRT_2022

[8] Evaluation of sequences stop-to-stop codon is a more conservative approach compared to the evaluation of start-to-stop codon sequences.
[9] http://comparedatabase.org/database/
[10] https://www.ncbi.nlm.nih.gov/protein/

database with a FASTA search. All five of these alignments represented self-identifications of the four novel proteins in MON94313 (DMO, PAT, FT_T.1 and TDO).

Given these results, the risk of allergenic proteins with relevance to human safety being produced by the ORFs in MON94313 is negligible.

3.4.5.2 Bioinformatic analysis for potential toxicity

The putative peptides encoded by the junction and insert sequences were compared *in silico* to the UniProt toxin protein database (TOX_2022). This database is a subset of sequences derived from the Swiss-Prot protein database[11], curated to remove likely non-toxin proteins, and contained 8,131 sequences at the date of analysis. A FASTA algorithm with a BLOSUM50 scoring matrix was used to query the putative peptides against the toxin database. No alignments with an E-score $\leq 1\times10^{-5}$ were identified, indicating that there was no significant homology between the putative peptides and any known protein toxins.

3.5 Conclusion

The data provided by the applicant showed that an integration event has occurred at a single locus in the soybean genome. Sequencing data confirmed that the *dmo*, *pat*, *ft_t.1* and *TDO* expression cassettes are present with the expected sequence and organisation in the genome of MON94313. No plasmid backbone sequences, including antibiotic resistance genes, from the transforming PV-GMHT529103 plasmid are present. The T-DNA II from PV-GMHT529103, which was inserted in the initial transformation, is not present in MON94313. The introduced DNA was shown to be stably inherited and expressed across several breeding generations of MON94313. None of the new ORFs created by the insertion raise any allergenicity or toxicity concerns.

[11] https://www.uniprot.org/

4 Characterisation and safety assessment of novel substances

In considering the safety of novel proteins it is important to understand that a large and diverse range of proteins are ingested as part of the normal human diet without any adverse effects. Only a small number of dietary proteins have the potential to impair health, because of anti-nutrient properties or triggering of allergies in some consumers (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-nutrient or allergenic effects.

To effectively identify any potential hazards, 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 is required. It is also important to determine if the newly expressed protein is expressed in the plant as expected, including whether any post-translational modifications have occurred.

Four novel proteins are expressed in MON94313, each of which confer herbicide tolerance: DMO, which confers tolerance to dicamba; PAT, which confers tolerance to glufosinate; FT_T.1, which confers tolerance to 2,4-D, and TDO, which confers tolerance to mesotrione. In considering the safety of newly expressed substances 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.

4.1 DMO

Tolerance to dicamba in MON94313 is conferred by the expression of the enzyme dicamba mono-oxygenase (DMO). Wildtype DMO was initially purified from the *S. maltophilia* strain DI-6, which was isolated from soil at a dicamba manufacturing plant (Krueger et al. 1989). DMO prevents the build-up of toxic levels of dicamba (3,6-dichloro-2-methoxy benzoic acid) by catalysing its demethylation to form the non-herbicidal compound 3,6-dichlorosalicylic acid (DCSA) (Behrens et al. 2007).

In *S. maltophilia*, DMO is the final component of a three-component enzyme system, in which electrons move from NADH through a reductase and ferredoxin, before being passed to DMO. Ferredoxin in plant chloroplasts closely resembles the ferredoxin component of this bacterial enzyme system (Behrens et al. 2007). Therefore, DMO in transgenic plants is typically expressed along with an N-terminal chloroplast transit peptide (CTP) to allow targeting of DMO to the chloroplast and co-localisation with reduced ferredoxin as an electron source.

In MON94313, the DMO expression cassette includes an additional CTP coding sequence from *Arabidopsis thaliana* upstream of the *dmo* gene. This results in the expression of a 408 amino acid precursor protein consisting of a 68 amino acid CTP (APG6) fused to the N-terminus of DMO. N-terminal sequencing analysis (Section 4.1.1) indicates that processing of the MON94313-produced DMO precursor protein results in a single 340 amino acid isoform of the mature DMO protein, with no residues from the transit peptide remaining at the N-terminus. After cleavage of the CTP, the mature protein has an apparent molecular weight of 38.4 kDa.

The mature DMO protein expressed in MON94313 protein differs from the wildtype *S. maltophilia* DMO by an additional leucine at position 2 (Herman et al. 2005). Alternatively-processed DMO proteins, some of which contained isoforms with parts of the transit peptide remaining at the N-terminus, have been previously assessed by FSANZ in the following

applications:

- A1063 (MON87708 soybean; FSANZ 2012)
- A1080 (MON88701 cotton; FSANZ 2013a)
- A1118 (MON87419 corn; FSANZ 2016)
- A1192 (MON87429 corn; FSANZ 2020)
- A1216 (MON94100 canola; FSANZ 2021).

A comparison of the alternatively-processed DMO proteins in these different plant lines with the wildtype bacterial DMO protein and the DMO expressed in MON94313 is shown in Figure 4.

The alignment of the amino acid sequences of the different DMO proteins shows they are highly similar. As can be observed in Figure 4, the sequence of DMO expressed in MON94313 is identical to one of the isoforms of DMO found in MON87429 corn (FSANZ 2020). Each DMO protein has the same function and catalyses the same enzymatic reaction. The small differences in sequence were not expected to result in changes in overall structure, immunoreactivity, enzyme activity or substrate specificity (D'Ordine et al. 2009; Dumitru et al. 2009).

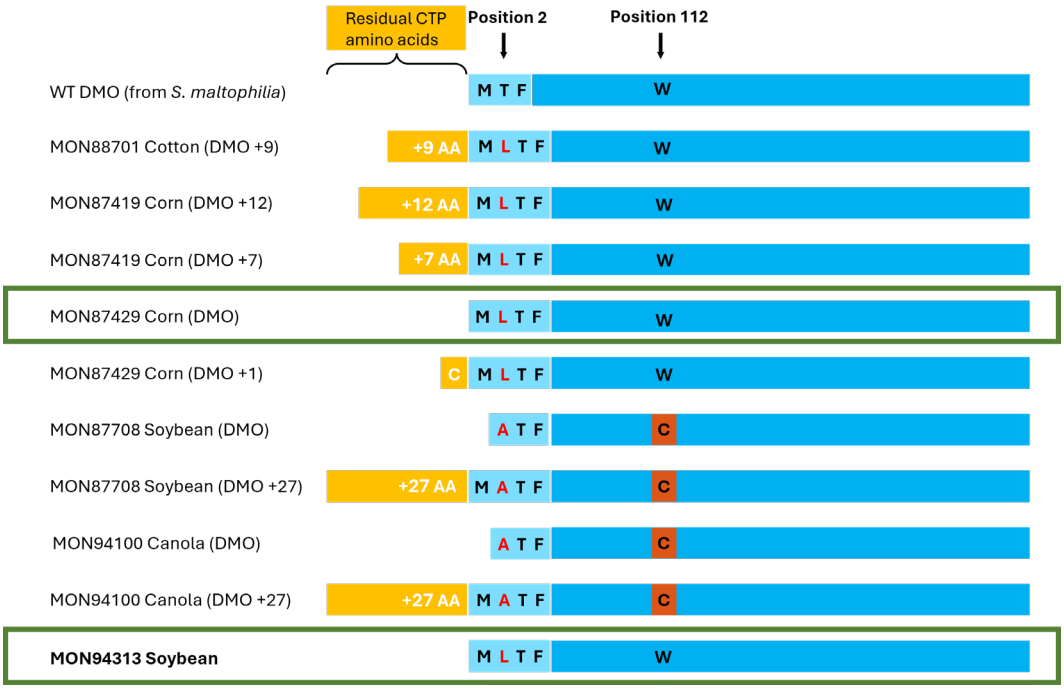


Figure 4: *Forms of DMO protein expressed in different GM commodities compared with wildtype DMO derived from S. maltophilia. Red text/boxes denote amino acid (aa) differences. Blue regions indicate areas of 100% amino acid identity. Some commodities contain a mixture of alternatively processed DMO proteins. The green boxes around MON94313 soybean and one of the MON87429 corn isoforms indicate that these are identical in sequence.*

4.1.1 Safety of the introduced DMO

The DMO protein has been considered in 5 previous FSANZ safety assessments[12]. The detailed safety assessment reports for each of these applications are available on the

[12] A1063 – soybean line MON87708 (FSANZ 2012); A1080 – cotton line MON88701 (FSANZ 2013a); A1118 – corn line MON87419 (FSANZ 2016); A1192 – corn line MON87429 (FSANZ 2020); A1216 – canola line MON94100 (FSANZ 2021).

FSANZ website[13]. In each of these previous assessments, studies on potential allergenicity and toxicity were submitted and assessed. These previous assessments did not raise any safety concerns. Results in the published literature also support the safety of DMO (EFSA 2011; Delaney et al. 2008; Behrens et al. 2007; Chakraborty et al. 2005; Duke 2005; Schmidt & Shaw 2001).

Since the sequence of the protein expressed in MON94313 is identical to a previous DMO sequence assessed by FSANZ (A1192 – corn line MON87429; see Figure 4), no further safety evaluation is required other than the examination of updated bioinformatics searches.

The applicant has submitted updated bioinformatic studies for DMO that looked for amino acid sequence similarity to known protein allergens and toxins (December 2022). FSANZ has assessed the data submitted by the applicant and the results do not alter conclusions reached in previous assessments.

4.1.2 Conclusion

The data provided by the applicant confirms the DMO expressed in MON94313 is identical to a previously assessed DMO protein. Updated bioinformatic analyses confirm that DMO has no similarity with known allergens or toxins that is of significance or concern.

4.2 PAT

The *pat* gene in MON94313 encodes the protein phosphinothricin N-acetyltransferase (PAT), which enzymatically inhibits phosphinothricin (PPT) (Strauch et al. 1988; Wohlleben et al. 1988). PPT is the active constituent of glufosinate ammonium herbicides and acts by irreversibly inhibiting the endogenous plant enzyme glutamine synthetase. This enzyme is involved in amino acid biosynthesis in plant cells and its inhibition causes accumulation of ammonia, leading to plant death. In glufosinate-tolerant GM plants, the introduced PAT enzyme chemically inactivates PPT by acetylation of the free ammonia group to produce N-acetyl glufosinate, allowing plants to continue amino acid biosynthesis in the presence of the herbicide (Hérouet et al. 2005).

The wildtype PAT protein encoded by the *pat* gene from *S. viridochromogenes* consists of 183 amino acids and shares 85% amino acid identity with the PAT protein encoded by the *bar* gene from *S. hygroscopicus*. The two PAT proteins are regarded as equivalent (OECD 1999): both exhibit a high degree of enzyme specificity, recognising and detoxifying only PPT (Wehrmann et al. 1996).

The PAT enzyme has been used to confer glufosinate-tolerance in crops for approximately 25 years (ILSI 2016). FSANZ has assessed and approved numerous events with both *pat*- and *bar*-encoded glufosinate-tolerance. There have been no credible reports of adverse effects on human health since it was introduced into food.

The PAT protein produced in MON94313 is identical to the wildtype *S. viridochromogenes* enzyme except that the N-terminal methionine has been removed co-translationally. This results in a protein comprised of 182 amino acids with an apparent molecular weight of ~25.5 kilodalton (kDa). Identical PAT proteins lacking the N-terminal methionine are found in a number of commercially available glufosinate-tolerant plants.

4.2.1 Safety of the introduced PAT

The PAT protein has been considered in numerous FSANZ safety assessments, including 6

[13] https://www.foodstandards.gov.au/consumer/gmfood/applications/Pages/default.aspx

in soybean[14]. These assessments, together with the published literature, have firmly established the safety of PAT and confirm that it does not raise toxicity or food allergenicity concerns in humans (ILSI 2016; Hammond et al. 2011; Delaney et al. 2008; Hérouet et al. 2005).

In previous FSANZ assessments, studies on potential allergenicity and toxicity were submitted and assessed. These previous assessments did not raise any safety concerns and there have been no credible reports of adverse health effects in humans. Since the sequence of the protein expressed in MON94313 is identical to the previous PAT sequences assessed by FSANZ, no further safety evaluation is required other than the examination of updated bioinformatics searches.

The applicant has submitted updated bioinformatic studies for PAT that looked for amino acid sequence similarity to known protein allergens and toxins (March 2022). FSANZ has assessed the data submitted by the applicant and the results do not alter conclusions reached in previous assessments.

4.2.2 Conclusion

The data provided by the applicant confirms the PAT expressed in MON94313 is identical to previously assessed PAT proteins. Updated bioinformatic analyses confirm that PAT has no similarity with known allergens or toxins that is of significance or concern.

4.3 FT_T.1

FT_T.1 is a modified version of the enzyme FT_T, which has been previously assessed by FSANZ in MON87429 corn (A1192; FSANZ 2020). FT_T is in turn a modified version of RdpA, an Fe(II)/alpha-ketoglutarate-dependent dioxygenase from the soil bacterium *Sphingobium herbicidovorans.* RdpA, and its modified derivatives, belong to the Fe(II)/alpha-ketoglutarate dependent hydroxylase superfamily, a diverse superfamily which is widely distributed throughout the animal, plant, and bacterial kingdoms (Hausinger et al. 2004). Humans have been exposed to many different members of this superfamily, each with their own protein sequence but similarities in structure, without adverse effects.

All three proteins (RdpA, FT_T, and FT_T.1) have the same function and carry out the same enzymatic reaction: the degradation of the synthetic auxin herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) to form the non-herbicidal compound 2,4-dichlorophenol (2,4-DCP) and glyoxylic acid (Figure 5). RdpA and the FT enzymes also catalyse the inactivation of the aryloxyphenoxypropionate (FOP) herbicides (Figure 5). However, given the natural tolerance of soybeans to FOPs, the inactivation of 2,4-D is of most relevance for MON94313.

[14] The PAT protein has been assessed in soybean lines IND-00410-5 (A1264; *bar* gene); DAS-81419-2 (A1087; *pat* gene); SYHT0H2 (A1081; *pat* gene); DAS-44406-6 (A1073; *pat* gene); DAS-68416-4 (A1046; *pat* gene); A2704-12 and A5547-127 (A481; *pat* gene).

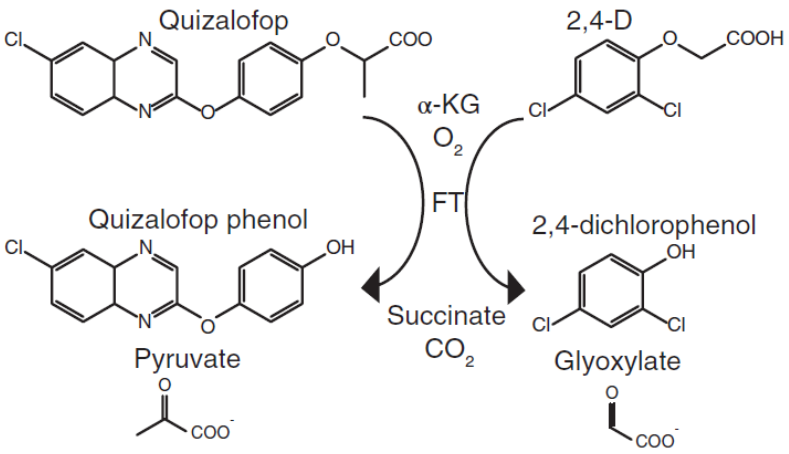


Figure 5: *Substrates and metabolites of the reactions catalysed by FT and RdpA enzymes. FOP herbicides, represented here by the compound quizalofop, are metabolised to an inactive quizalofop phenol. 2,4-D is degraded to the herbicidally inactive compound 2,4-DCP.*

The FT_T.1 protein sequence in MON94313 differs by three amino acids from the sequence of FT_T (Figure 6). FT_T in turn contains 30 amino acid modifications compared to the wildtype RdpA enzyme from *S. herbicidovorans*. The modifications introduced into FT_T and FT_T.1 have not altered their substrate specificity relative to RdpA. However, the modifications made to RdpA to create FT_T have improved its activity on herbicide substrates and its stability at higher temperatures (Larue et al. 2019). FT_T.1, compared to FT_T, has an even stronger activity on 2,4-D, while maintaining the same level of heat stability (Larue et al. 2019). Soybean is highly sensitive to 2,4-D compared to corn and therefore requires a higher level of FT_T activity to provide in-crop herbicide tolerance.

The *ft_t.1* gene prepared by the applicant encodes a protein of 295 amino acids, with an apparent molecular weight of 34 kDa.

4.3.1 Characterisation of FT_T.1 expressed in MON94313 and equivalence to a bacterially-produced form

The equivalence of the MON94313- and *E. coli*-produced FT_T.1 proteins must be established before the safety data generated using *E. coli*-produced FT_T.1 can be applied to MON94313-produced FT_T.1 protein.

The plant-produced FT.T.1 protein was purified from the seed of MON94313 using a combination of anion-exchange and immunoaffinity chromatography. *E. coli*-produced FT_T.1 protein was generated following the fermentation of *E. coli* containing a plasmid expressing the protein.

In order to confirm the identity and equivalence for the MON94313- and *E. coli*-produced FT_T.1 proteins, a series of analytical tests were done, the results of which are summarised below.

4.3.1.1 Molecular weight analysis

Aliquots of purified MON94313- and *E. coli*-produced FT_T.1 proteins were subjected to SDS-PAGE then visualised with Brilliant Blue G-Colloidal stain. The MON94313- and *E. coli*-produced FT_T.1 migrated with apparent molecular weights of 34.0 and 33.6 kDa, respectively, which are within the acceptance limits for equivalence. The average purity of the MON94313-produced FT_T.1 was determined to be 98%.

4.3.1.2 N-terminal sequencing

N-terminal sequencing confirmed that the first 15 amino acids of MON94313- and *E. coli*-produced FT_T.1 proteins were as expected based on the *ft_t.1* gene.

4.3.1.3 Peptide mass fingerprint analysis

MON94313-derived and *E. coli*-derived FT_T.1 were digested with a combination of trypsin and Asp-N and analysed by LC-MS/MS. For MON94313-derived FT_T.1, 57 unique peptides were identified, covering 100% of the expected protein sequence (295 of 295 amino acids). For *E. coli*-derived FT_T.1, 44 unique peptides, covering 100% of the expected protein sequence, were identified. These results further confirm that the protein being expressed in MON94313 and *E. coli* is FT_T.1.

4.3.1.4 Western blot and immunoreactivity analysis

Western blot analysis with an FT_T.1-specific antibody showed that the protein being expressed in MON94313 and *E. coli* was indeed FT_T.1 and they have equivalent immunoreactivity.

4.3.1.5 Glycosylation analysis

An ECL glycoprotein detection procedure showed the FT_T.1 proteins from both MON94313 and *E. coli* were equivalent and that neither is glycosylated.

4.3.1.6 Functional activity analysis

A colourimetric assay that measures the degradation of 2,4-D to 2,4-DCP catalysed by FT_T.1 showed that the MON94313- and *E. coli*-produced FT_T.1 proteins had equivalent functional activity.

4.3.1.7 Conclusion

These data demonstrated that the bacterially-produced FT_T.1 protein is a suitable:

- positive control for the Western blot experiment used to characterise the phenotypic stability of FT_T.1 in Section 3.4.4.2
- standard for the immunoassay study used to detect FT_T.1 protein in plant tissues as discussed in Section 4.5
- surrogate for use in the safety assessment experiments described in Section 4.3.2.

4.3.2 Safety of the introduced FT_T.1

The FT_T protein, which is 98% identical to FT_T.1 (see Figure 6), has been assessed previously by FSANZ in application A1192 (corn line MON97429; FSANZ 2020). The studies on potential allergencity and toxicity that were submitted for this previous assessment did not raise any safety concerns.

The applicant submitted additional safety studies specific to the FT_T.1 protein with this application. The results of a bioinformatics analysis (December 2022) that looked for amino acid sequence similarity to known protein toxins and allergens in publicly available databases. The analysis did not identify any known allergens or toxins with homology to FT_T.1. The applicant also provided *in vitro* digestibility (pepsin and pancreatin) and thermolability studies for the FT_T.1 protein. These studies confirmed conclusions from the assessment of the FT_T protein for Application A1192: FT_T.1 would be rapidly degraded in the stomach following ingestion, and that the protein, while remaining largely intact at

elevated temperatures, is functionally inactivated by heating.

4.3.3 Conclusion

A range of characterisation studies were performed on MON94313-derived FT_T.1 protein confirming its identity and functionality as well as equivalence to the corresponding protein produced in a bacterial expression system. Bioinformatic analyses showed FT_T.1 had no amino acid sequence similarity to known toxins or allergens. The protein was shown to be inactivated by heating and susceptible to pepsin and pancreatin digestion. Taken together this indicates that the FT_T.1 protein is unlikely to be toxic or allergenic to humans.

4.4 TDO

The TDO protein expressed in MON94313 confers tolerance to the herbicide mesotrione. Mesotrione is a potent inhibitor of 4-hydroxyphenylpyruvate dioxygenase (HPPD), which is responsible for the conversion of hydroxyphenylpyruvic acid (HPPA) into homogentisic acid (HGA) in plants. Inhibition of HPPD and the subsequent reduction in HGA levels leads to the depletion of a number of downstream metabolites, including plastoquinone and tocopherols, which are crucial for carotenoid biosynthesis. Given the role of carotenoids in protecting the photosynthetic machinery from oxidative damage, application of mesotrione to susceptible plants causes leaf bleaching and eventually plant death (Mitchell et al. 2001).

TDO is the product of the *TDO* gene – a version of the *HPPD INHIBITOR SENSITIVE 1* (*HIS1*) gene from rice (*Oryza sativa*) that has been codon-optimised for expression in soybean. In rice, HIS1 confers natural tolerance to β-triketone herbicides, which include mesotrione, by catalysing their hydroxylation to herbicidally inactive metabolites (Maeda et al. 2019). Soybean plants are highly sensitive to mesotrione (Mitchell et al. 2001). When expressed in soybean, TDO confers tolerance to mesotrione by the same mechanism in rice, i.e. directly inactivating mesotrione by sequentially oxidising it first to hydroxy-mesotrione, then oxy-mesotrione (Figure 6) (Dai et al. 2022).

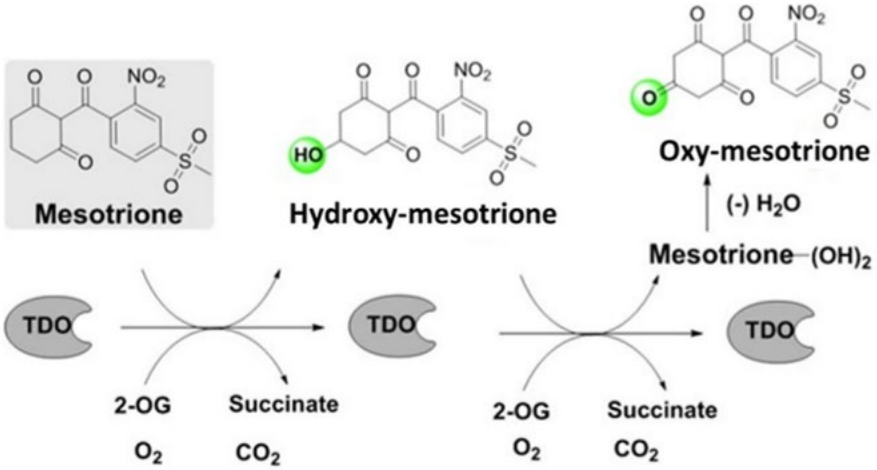


Figure 6: *TDO mode of action. TDO oxidises mesotrione in a two-step reaction, yielding hydroxy-mesotrione then oxy-mesotrione, which are herbicidally inactive.*

Like FT_T.1, TDO is an Fe(II)/alpha-ketoglutarate-dependent dioxygenase, and belongs to the Fe(II)/alpha-ketoglutarate dependent hydroxylase superfamily. To determine its specificity, an *in silico* screen was used to identify any endogenous soybean small molecules that could potentially act as TDO substrates. *In vitro* testing of 32 of these molecules alongside mesotrione and other β-triketone herbicides showed that TDO had no significant (p

<0.05) enzymatic activity with any of the endogenous molecules, and was specifically active towards the β-triketone herbicides.

The *TDO* gene prepared by the applicant encodes a protein of 351 amino acids (including the N-terminal methionine), with an apparent molecular weight of ~37 kDa. Other than lacking the N-terminal methionine residue as a result of protein processing (See Section 4.5.1.2), TDO expressed in MON94313 is identical in amino acid sequence to rice HIS1.

4.4.1 Characterisation of TDO expressed in MON94313 and equivalence to a bacterially-produced form

The equivalence of the MON94313- and *E. coli*-produced TDO proteins must be established before the safety data generated using *E. coli*-produced TDO can be applied to MON94313-produced TDO protein.

The plant-produced TDO protein was purified from the seed of MON94313 using a combination of anion-exchange, immunoaffinity, and size-exclusion chromatography. *E. coli*-produced TDO protein was generated following the fermentation of *E. coli* containing a plasmid that expresses the protein.

In order to confirm the identity and equivalence for the MON94313- and *E. coli*-produced TDO proteins, a series of analytical tests were done, the results of which are summarised below.

4.4.1.1 Molecular weight analysis

Aliquots of purified MON94313- and *E. coli*-produced TDO proteins were subjected to SDS-PAGE then visualised with Brilliant Blue G-Colloidal stain. The MON94313- and *E. coli*-produced TDO migrated with apparent molecular weights of 36.9 and 37.0 kDa, respectively, which are within the acceptance limits for equivalence. The average purity of the MON94313-produced TDO was determined to be 71%.

4.4.1.2 N-terminal sequencing

N-terminal sequencing confirmed that the first 15 amino acids of MON94313- and *E. coli*-produced TDO proteins were as expected based on the *TDO* gene. The N-terminal methionine had been cleaved in the MON94313-produced TDO (see Figure 8), which is a common process in many organisms (Wingfield 2017). The *E. coli*-produced TDO protein was produced with alanine as the N-terminal residue, meaning that both proteins possessed an identical N-terminal sequence. The results of N-terminal sequencing confirms the identity of the MON94313-produced TDO protein.

4.4.1.3 Peptide mass fingerprint analysis

MON94313-derived and *E. coli*-derived TDO were digested with trypsin and analysed by LC-MS/MS. For MON94313-derived TDO, 47 unique peptides were identified, covering 88% of the expected protein sequence (309 of 351 amino acids). For *E. coli*-derived TDO, 79 unique peptides, covering 100% of the expected protein sequence (350 of 350 amino acids), were identified. These results further confirm that the protein being expressed in MON94313 and *E. coli* is TDO.

4.4.1.4 Western blot and immunoreactivity analysis

Western blot analysis with a TDO-specific antibody showed that the protein being expressed in MON94313 and *E. coli* was indeed TDO and they have equivalent immunoreactivity.

4.4.1.5 Glycosylation analysis

An ECL glycoprotein detection procedure showed the TDO proteins from both MON94313 and *E. coli* were equivalent and that neither is glycosylated.

4.4.1.6 Functional activity analysis

An assay measuring the amount of mesotrione converted to hydroxy-mesotrione and keto-mesotrione over a 30 minute period showed that the MON94313- and *E. coli*-produced TDO proteins had equivalent functional activity.

4.4.1.7 Conclusion

These data demonstrated that the bacterially-produced TDO protein is a suitable:

- positive control for the Western blot experiment used to characterise the phenotypic stability of TDO in Section 3.4.4.2
- standard for the immunoassay study used to detect TDO protein in plant tissues as discussed in Section 4.5
- surrogate for use in the safety assessment experiments described in Section 4.4.2.

4.4.2 Safety of the introduced TDO

The TDO protein has not been previously assessed by FSANZ. Data were provided to assess the potential toxicity and allergenicity of the TDO protein expressed in MON94313.

4.4.2.1 Bioinformatic analyses of TDO

Bioinformatic analyses, as described in Section 3.4.5.1, were performed to compare the TDO amino acid sequence to known allergenic proteins in the COMPARE allergen database (AD_2022). The search did not identify any known allergens with homology to TDO. No alignments had an E-score of $\leq 1\times10^{-5}$ or met or exceeded the threshold of greater than 35% similarity over ≥ 80 amino acids, and no eight amino acid peptide matches were shared between the TDO sequence and proteins in the allergen database.

The applicant also provided the results of *in silico* analyses comparing the amino acid sequence of TDO to proteins identified as “toxins” in the TOX_2022 database, as described in Section 3.4.5.2. The search did not identify any known toxins with homology to TDO.

4.4.2.2 Susceptibility of TDO to digestion with pepsin and pancreatin

E. coli-produced TDO (test protein) was mixed with pepsin (10U enzyme/µg protein), then incubated for 0-60 min at 37°C. Reactions occurred under acidic conditions in simulated gastric fluid (Thomas et al. 2004). A control mixture lacking pepsin, and a separate control lacking TDO, were also incubated for 60 min and analysed at 0 min and 60 min.

The extent of digestion was visualised by Brilliant Blue G-colloidal stained SDS-PAGE and Western blotting with an anti-TDO antibody. Concurrently, a serial dilution of the reaction mix (test protein plus enzyme) without incubation was used to determine the limit of detection (LOD) for the TDO protein after gel staining and Western blotting. In the protein gel staining analysis, 1 µg test protein was loaded per lane and the LOD was calculated to be ~12.5 ng. In the Western blotting experiments, 5 ng test protein was loaded and the LOD was ~0.08 ng.

Visual inspection of both the stained gel and the Western blot showed that by 0.5 min of incubation, the amount of intact TDO remaining in the reaction mix was below the LOD.

Based on the LOD, it was calculated that >98.8% of the intact TDO was digested within 0.5 min when analysed by SDS-PAGE and >98.4% when analysed by Western blot. A single peptide fragment of ~3.5 kDa was observed in the SDS-PAGE until 5 min, but this fragment disappeared after 10 min of digestion. No smaller peptides were detected in the Western blot at any time beyond 0.5 min. TDO in the control mixture lacking pepsin remained intact after 60 min of incubation, indicating that the rapid loss of TDO protein in the reaction mixes is due to proteolytic digestion of TDO.

To assess the susceptibility of TDO to pancreatin[15] digestion, *E. coli*-derived TDO was incubated with pancreatin (~55 µg pancreatin/µg TDO) at 37°C for 0-24 h, in a simulated intestinal fluid system at a neutral pH range. A control mixture lacking pancreatin, and a separate control lacking TDO, were also incubated for 24 h and analysed at 0 min and 24 h. The extent of digestion was visualised by Western blot. A serial dilution of the reaction mixture without incubation was used to determine the LOD for the protein, which was approximately 0.16 ng. For the digestibility analysis, ~5 ng of protein was loaded per lane.

Visual inspection of the Western blot showed that after 5 min of incubation, the level of intact TDO remaining in the reaction mix was below the LOD. Based on the LOD, it was calculated that >96.8% of the intact TDO protein was degraded within 5 min. TDO in the control mixture lacking pancreatin remained intact after 24 h of incubation, indicating that the loss of TDO in the reaction mix is indeed due to proteolytic digestion of TDO by pancreatin.

Taken together, these results indicate that TDO will be fully degraded by gastric and intestinal enzymes in the human digestive system.

4.4.2.3 Structural stability and bioactivity of TDO after exposure to heat

E. coli-produced TDO protein was heated for 15 or 30 min at temperatures ranging from 25-95°C. A control sample was kept on wet ice (~0°C). Aliquots of the control and heated protein samples were run on SDS-PAGE and stained with Brilliant Blue G-Colloidal stain to detect the extent of protein degradation. No visible degradation or decrease in band intensity was observed for the TDO protein in the control, 25, 37, 55, 75 or 95°C treated samples at either 15 or 30 min.

While heat treatment did not degrade the TDO protein, heat may impact its structure and functionality. To assess its functional activity, heated and unheated TDO protein was tested in an assay that measures the conversion of mesotrione to hydroxy-mesotrione and keto-mesotrione (Table 4). When TDO was incubated at 25°C for either 15 or 30 min it retained full activity relative to the control. Increasing the temperature to 37°C for either 15 or 30 min reduced the activity of the TDO protein to 82% and 42%, respectively. At temperatures of 55°C or above for either time period, the functional activity of TDO was reduced to below the LOD. These data indicate that at elevated temperatures the TDO protein becomes denatured and loses functional activity.

[15] Pancreatin is a mixture of proteolytic enzymes

Table 4: Triketone dioxygenase activity of heat-treated TDO

Temperature	15 min heat treatment		30 min heat treatment	
	Specific activity (nmol x min^{-1} x mg^{-1})	Relative activity (% of control sample)[1]	Specific activity (nmol x min^{-1} x mg^{-1})	Relative activity (% of control sample)[1]
0°C (control)	197	100	197	100
25°C	204	104	205	104
37°C	162	82	83	42
55°C	≤LOD[2]	0	≤LOD[2]	0
75°C	≤LOD[2]	0	≤LOD[2]	0
95°C	≤LOD[2]	0	≤LOD[2]	0

1. The activity of the TDO protein in the control sample was assigned 100% active, with other samples calculated relative to this.
2. The limit of detection (LOD) of this assay was 22.8 nmol x min^{-1} x mg^{-1}

4.4.3 Conclusion

A range of characterisation studies were performed on MON94313-derived TDO protein confirming its identity and functionality as well as equivalence to the corresponding protein produced in a bacterial expression system. Bioinformatic analyses showed TDO had no amino acid sequence similarity to known toxins or allergens. The protein was shown to be inactivated by heating and susceptible to pepsin and pancreatin digestion. Taken together this indicates that the TDO protein is unlikely to be toxic or allergenic to humans.

4.5 Expression levels of novel proteins

For analysis of the expression levels of the DMO, PAT, FT_T.1 and TDO proteins in MON94313, tissues were collected from four replicate plots at each of five field-trial sites in representative soybean-producing regions of the United States during the 2020 growing season[16]. MON94313 was treated with dicamba, glufosinate, 2,4-D, and mesotrione throughout the growing period. Forage, grain, leaf and root tissue samples were collected from each plot at specified growth stages (see Figure 7 for a summary of soybean growth stages and the stage at which each tissue type was collected).

[16] Field sites for testing protein expression levels were in the following states – Iowa, Missouri, Nebraska, Illinois and Wisconsin

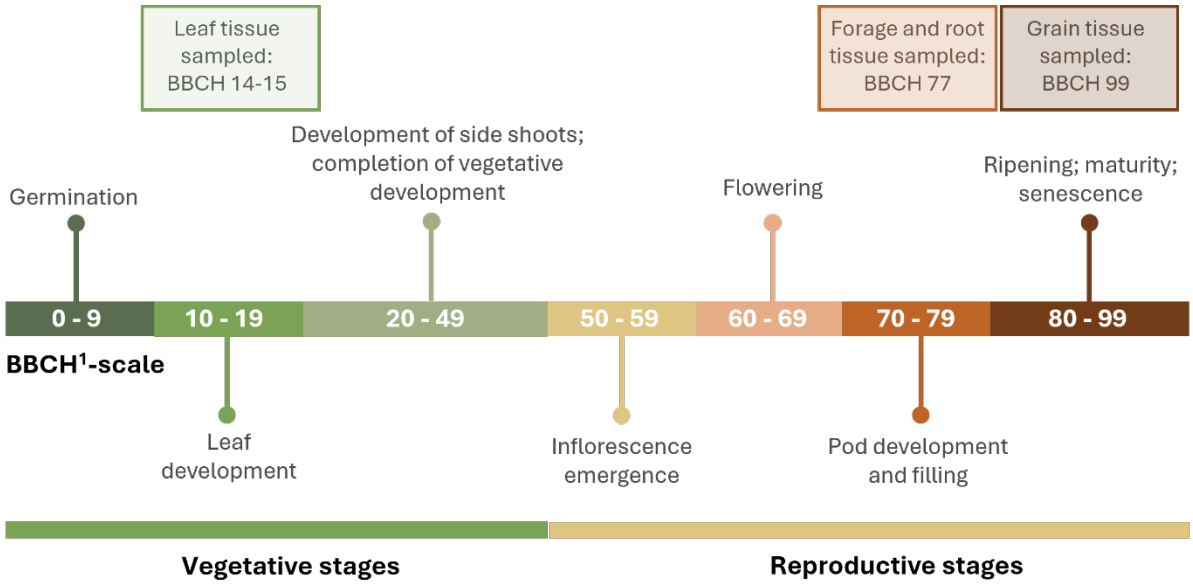


Figure 7: *Growth stages of soybean. The stages at which the leaf, forage, root and grain tissue for protein expression analysis were sampled are indicated.*
1. BBCH = Biologische Bundesanstalt, Bundessortment und Chemische Industrie (Meier 2001)

DMO, PAT, FT_T.1 and TDO were extracted from tissues using standard methods and their expression levels were quantified in each tissue using a quantitative enzyme-linked immunosorbent assay (ELISA). The *E. coli*-derived versions of each protein were used as analytical references for the respective plant-derived proteins.

The mean level of each protein in each tissue type determined by ELISA is shown in Figure 8. Of the four proteins, DMO had the highest expression levels across all four tissue types. The mean DMO expression in herbicide-treated MON94313 was highest on a dry weight (dw) basis in forage (150 µg/g dw) and lowest in root at 20 µg/g dw. The mean DMO protein level in MON94313 grain was 40 µg/g dw.

For PAT, FT_T.1 and TDO, the mean expression levels were highest in leaf and lowest in root (for TDO, the expression level in root was below the LOQ of 0.50 µg/g dw). The mean PAT protein level in MON94313 grain was 3.8 µg/g dw. The mean FT_T.1 protein level in herbicide-treated MON94313 grain was 6.1 µg/g dw. The mean TDO protein level in MON94313 grain was 5.0 µg/g dw.

For the full set of expression data, including standard errors and ranges, refer to the Application dossier[17] (pages 132 – 132).

[17] https://www.foodstandards.gov.au/code/applications/Pages/A1276---Food-derived-from-herbicide-tolerant-soybean-line-MON94313.aspx

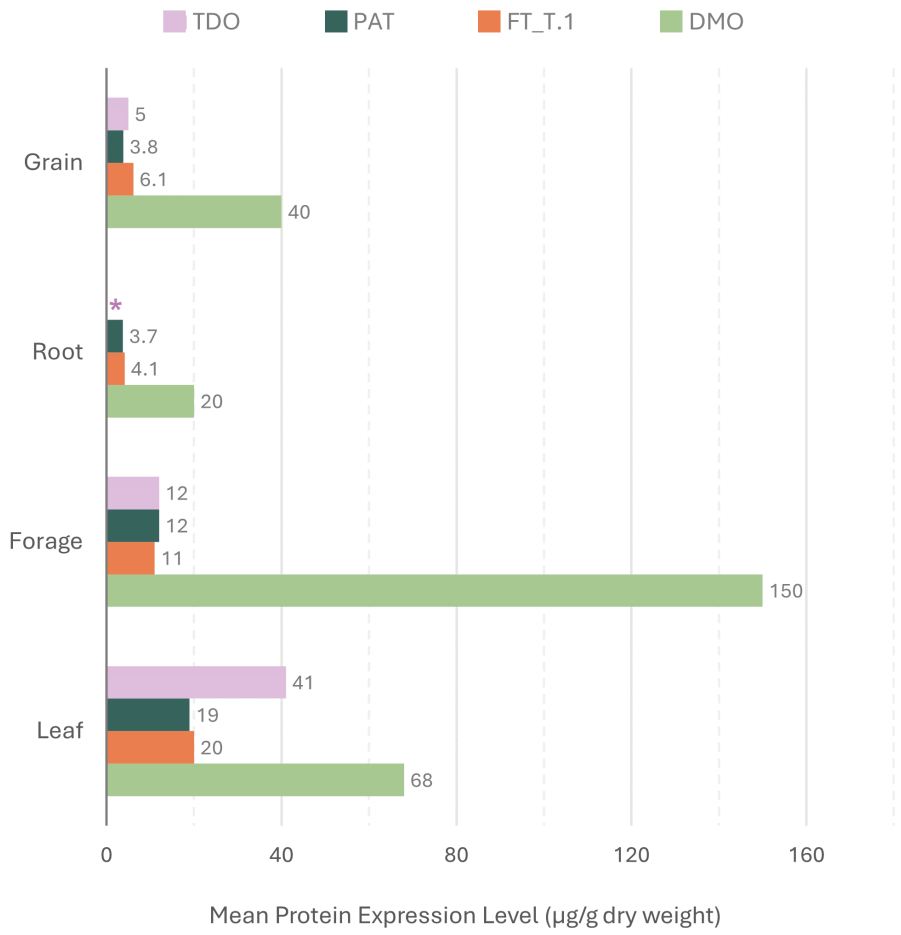


Figure 8: *Mean expression levels of the DMO, PAT, FT_T.1 and TDO proteins in four tissue types from herbicide-treated MON94313. The asterisk denotes that the level of TDO in root tissue was below the LOQ (0.50 µg/g dw).*

4.6 Novel herbicide metabolites in GM herbicide-tolerant plants

As part of the safety assessment it is important to establish whether the expression of novel proteins for conferring herbicide tolerance is likely to result in the accumulation of any novel herbicide metabolites. Novel metabolites are those not normally found in non-GM crops sprayed with the same herbicide. If such substances are found to occur as a result of the genetic modification, then it is important to determine their potential toxicity. The detoxification of herbicides to inactive compounds, which is the mechanism of action of DMO, PAT, FT_T.1 and TDO, may result in the production of novel metabolites.

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

For PAT and DMO, the metabolic profiles resulting from the novel protein x herbicide interaction have been established. The glufosinate-tolerance (PAT) and dicamba-tolerance (DMO) traits have been examined in multiple previous applications to FSANZ. These previous assessments indicate the spraying of MON94313 with glufosinate ammonium or dicamba would result in the same metabolites that are produced in non-GM soybean sprayed

with the same herbicides. As no new glufosinate or dicamba metabolites would be generated in soybean line MON94313, further assessment is not required.

The 2,4-D-tolerance trait conferred by FT_T has been previously assessed by FSANZ in corn line MON87429 (A1192; FSANZ 2020). This previous assessment found that no herbicide metabolites are produced in MON87429 that are not also produced in conventional crops treated with 2,4-D herbicide. FSANZ notes that the 3 amino acid difference between FT_T.1 and FT_T is unlikely to make a significant difference to its functional activity against 2,4-D. The applicant has also specified that no novel metabolites are produced in 2,4-D-treated MON94313 as a result of FT_T.1 compared to 2,4-D-treated MON87429 as a result of FT_T. They have proposed that 2,4-D undergoes a similar metabolic pathway in both crops. As no new 2,4-D metabolites would be generated in soybean line MON94313, further assessment is not required.

4.6.1 Metabolism of mesotrione

The metabolic profile resulting from the TDO protein x mesotrione interaction has not been previously assessed by FSANZ. A mesotrione-tolerant soybean line, SYHT0H2, has been previously assessed by FSANZ (A1081; FSANZ 2013b). However, the AvHPPD-03 protein expressed in SYHT0H2 confers mesotrione-tolerance by acting as a HPPD enzyme with reduced affinity for mesotrione, not by inactivating mesotrione directly. It is therefore necessary to establish whether the application of mesotrione to MON94313 results in the accumulation of novel herbicide metabolites.

A metabolism study was conducted in which radiolabelled [phenyl-U-^{14}C]mesotrione or [cyclohexane-2-^{14}C]mesotrione were applied to the leaves of MON94313 at the R1 growth stage (equivalent to BBCH stage ~60; see Figure 7). The purpose of using radiolabelled mesotrione was to determine the nature of residues found in or on agricultural commodities of mesotrione-tolerant soybean, following pre-emergence or combined pre- and post-emergence treatments with the herbicide. Soybean forage and hay were harvested at the R1-R3 growth stage and soybean seed was harvested at maturity (BBCH stage 89; see Figure 7).

The parent mesotrione molecule was identified as the major residue in the soybean forage and hay following pre-emergence and post-emergence application, suggesting that mesotrione is not metabolised extensively in MON94313 following combined treatment. A conjugate of 4-methylsulfonyl-2-nitrobenzoic acid (MNBA) was observed as the major residue in the soybean forage and hay following pre-emergence application only. In soybean seed, mesotrione and its metabolites were not detected, suggesting that mesotrione residues are not transported out of the treated tissues at appreciable levels.

The mesotrione metabolite profile observed in MON94313 is consistent with what has been observed with other mesotrione-tolerant plants as well as non-GM plants. Hydrolysis of mesotrione to form MNBA, followed by conjugation, has been observed in cranberries, soybeans, maize, rice and peanut (Figure 9; JMPR 2014). Other pathways of mesotrione metabolism, such as hydroxylation to 4- or 5-hydroxy mesotrione, and reduction of MNDA to form AMBA and MBA (Figure 9) were not observed in MON94313. Taken together, the results of the labelled mesotrione metabolism study show that no herbicide metabolites are produced in MON94313 that are not also produced in conventional crops treated with mesotrione.

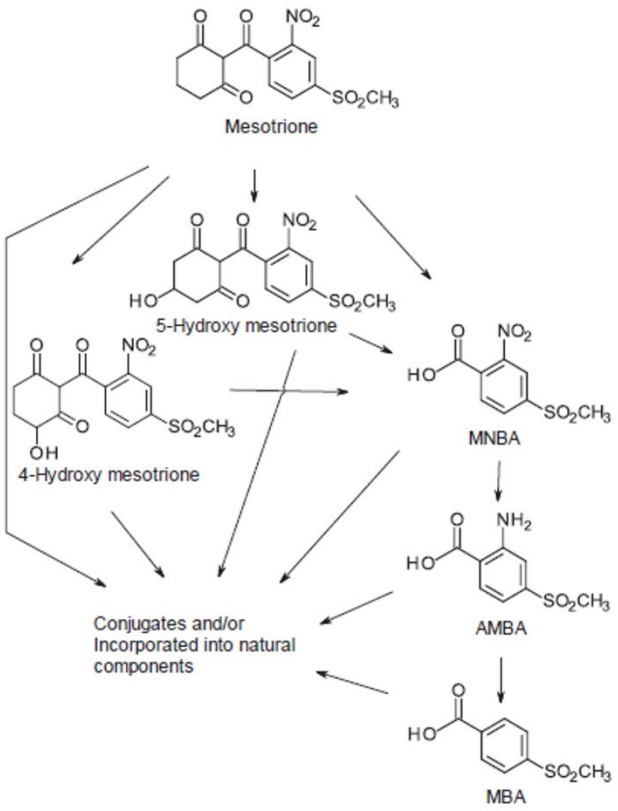


Figure 9: *Proposed metabolism of mesotrione in plants, based on metabolism studies with labelled mesotrione in cranberries, soybeans, maize, rice and peanut (JMPR 2014). Only the mesotrione parent and conjugates of MNBA were detected following labelled mesotrione treatment of MON94313.*

5 Compositional analysis

The main purpose of compositional analysis is to determine if, as a result of the genetic modification, an unexpected change has 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 analysis 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.

5.1 Key components

The key components to be analysed for the comparison of GM and conventional soybean are outlined in the OECD Consensus Document on Compositional Considerations for New Varieties of soybean (OECD 2012), and include: proximates, fibre, amino acids, fatty acids, minerals, vitamins, anti-nutrients and isoflavones.

5.2 Study design

MON95275 (R5 generation) and a non-GM control of similar genetic background (A3555) were grown and harvested from five field trial sites in the United States during the 2020 growing season[18]. The sites were representative of soybean growing regions suitable for commercial production. The field sites were established in a randomised complete block design with four replicates per site. Plants were grown under agronomic field conditions typical for each of growing region. MON94313 plots were treated with glufosinate, dicamba, 2,4-D and mesotrione.

At maturity, grain was harvested from all plots and shipped to an analytical laboratory at ambient temperature, before being frozen at -20°C until analysis. Compositional analyses were based on internationally recognised procedures including official methods specified by the Association of Official Analytical Chemists (AOAC), the Analytical Oil Chemists' Society (AOCS), methods specified by the manufacturer of the equipment used for analysis, or other published scientific methods.

58 different analytes were measured in grain (see Figure 10 for a complete list). Moisture was also measured and used to convert the analyte values from fresh to dry weight. Analytes were expressed as either percent dry weight (% dw), µg/g dw or as a percentage of total fatty acids (% total FA), as shown in Figure 13. Of the 58 components measured, 9 had more than 50% of observations below the LOQ (listed in grey in Figure 10) and were excluded from the statistical analyses, leaving a total of 49 components that were fully analysed in grain.

A linear mixed model analysis of variance was applied on data combined across the five replicated field trial sites. Statistical analyses were performed using SAS 9.4 (SAS Institute, Cary, North Carolina 2012). For each analyte, ‘descriptive statistics’ (mean, standard error (SE), and range) were generated.

In assessing the statistical significance of any difference between MON94313 and the conventional control, a *p*-value of 0.05 was used. Any statistically significant differences were evaluated further to assess whether they were likely to be biologically meaningful. The magnitude of differences in mean values between MON94313 and the control were determined, and this difference was compared to the variation observed within the control grown at multiple sites.

In addition, the natural variation of analytes from the literature and from the AFSI Crop Composition Database (AFSI CCDB) was also considered (Lundry et al. 2008; Berman et al. 2009; Bellaloui et al. 2011; Thompson et al. 2016; Breeze et al. 2015; AFSI 2023). The ranges derived from these values account for variability present in non-GM soybean varieties due to a wider range of agronomic and environment conditions, as well as different genetic backgrounds.

Key analyte levels were also analysed in forage but the results are not included in this report. It is noted however that, in the combined site analysis, none of the analyte levels in MON94313 differed significantly from those of the control.

[18] The states in which the five field trial sites were located: Iowa, Missouri, Nebraska, Illinois and Wisconsin

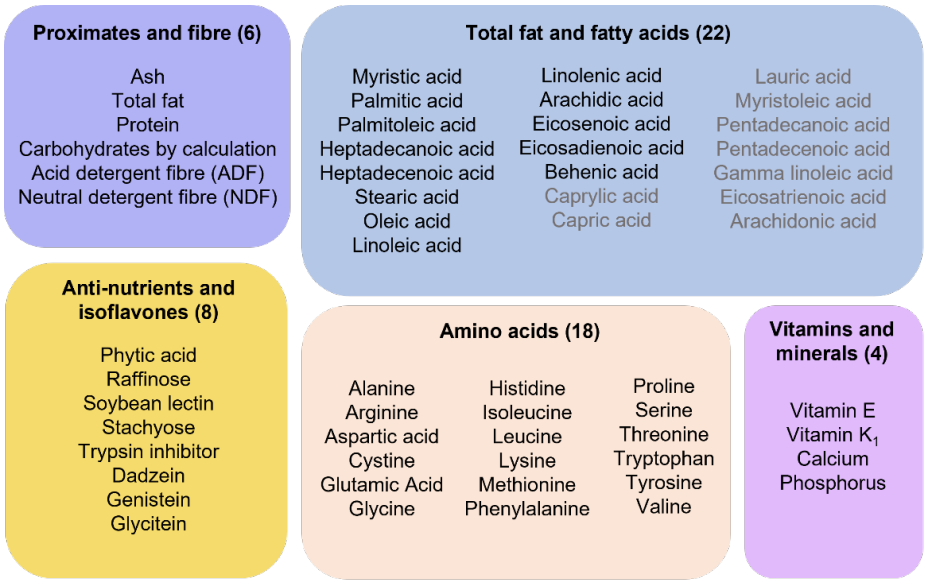


Figure 10: *Analytes measured in MON94313 grain samples. Analytes listed in grey text had >50% of samples below the LOQ and were excluded from statistical analysis. The 49 analytes listed in black text were analysed fully.*

5.3 Analyses of key components in grains

Of the 49 analytes for which mean values were provided (Figure 10), there were 7 for which there was a statistically significant difference ($p < 0.05$) between herbicide-treated MON94313 and the control: cystine, tryptophan, palmitic acid, linolenic acid, carbohydrates by calculation, vitamin K_1, and glycitein. A summary of these 7 analytes is provided in Figure 11. For the complete data set, including values for the analytes for which no statistically significant differences were found, refer to the Application dossier[19] pages 199 – 212).

For all 7 analytes where a statistical significant difference was found, the deviation of the MON94313 mean from the control mean was less than 8% (Figure 11a). However, as can be observed in Figure 11 (panels b-h), the MON94313 mean for each of these components was within the control range value, indicating that MON94313 has a smaller impact on the levels of these analytes than does natural variation within the conventional control. In addition, the observed MON94313 means fall well within the natural variability seen in the publicly available data ranges (grey bars, Figure 11, b-h). The differences reported here are therefore consistent with the normal biological variability that exists in soybean.

Overall, the compositional data are consistent with the conclusion that there are no biologically significant differences in the levels of key constituents in MON94313 when compared with conventional non-GM soybean varieties already available in agricultural markets. Grain from MON94313 can therefore be regarded as equivalent in composition to grain from conventional non-GM soybean.

[19] https://www.foodstandards.gov.au/code/applications/Pages/A1276---Food-derived-from-herbicide-tolerant-soybean-line-MON94313.aspx

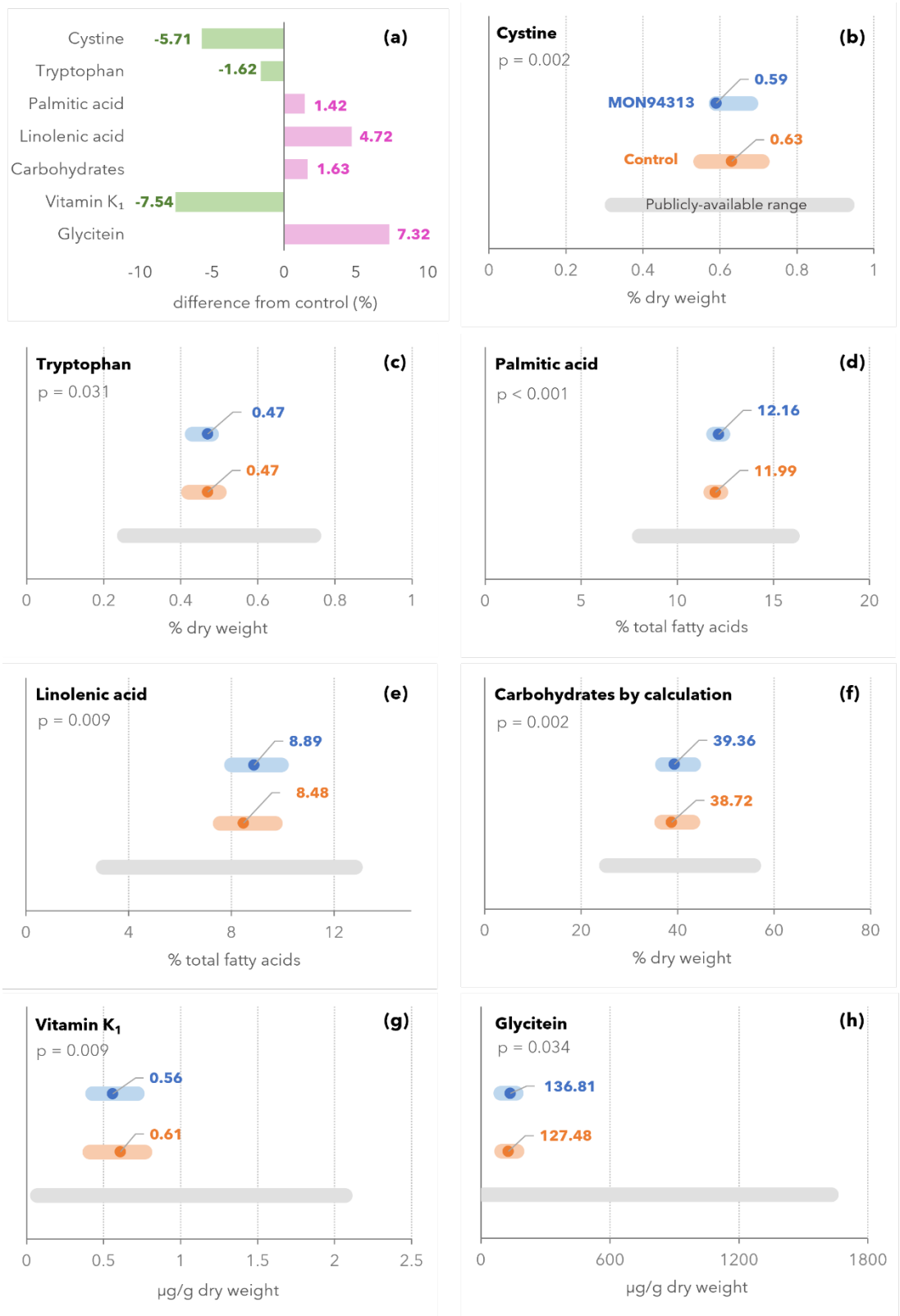


Figure 11: *Visual summary of statistically significantly analyte differences in MON94313 compared to the conventional control.* ***(a)*** *Percentage deviation of the mean MON94313 value from the mean control value for each of the 7 analytes for which significant differences were found.* ***(b) - (h)*** *Measured means (dots) and ranges for MON94313 (blue bars) and the control (orange bars) for the 7 analytes as labelled. The grey bars represent the publicly-available ranges for each analyte. Note that the x-axes vary in scale and unit for each analyte.*

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 wellbeing. In most cases, this can be achieved through a detailed understanding of the genetic modification and its consequences, together with an extensive compositional analysis of the food, such as that presented in Section 5 of this report.

Where a GM food has been shown to be compositionally equivalent to conventional varieties, the evidence to date indicates that feeding studies using target livestock or other animal species will add little to the safety assessment (OECD 2003; Bartholomaeus et al. 2013). If the compositional analysis indicates biologically significant changes, either intended or unintended, to the levels of certain nutrients in the GM food, additional nutritional studies should be undertaken to assess the potential impact of the changes on the whole diet.

MON94313 is the result of genetic modifications to confer tolerance to the herbicides dicamba, glufosinate, 2,4-D, and mesotrione, with no intention to significantly alter nutritional parameters in the food. The compositional analyses have demonstrated that the genetic modifications have not altered the nutrient composition of MON94313 compared with that of conventional non-GM soybean varieties. The introduction of food derived from MON94313 into the food supply is therefore expected to have negligible nutritional impact.

7 References

AFSI (2023) Crop composition database, version 9.1. Agriculture & Food Systems Institute. https://www.cropcomposition.org/ Accessed September 2023

Anderson JE, Michno JM, Kono TJ, et al. (2016) Genomic variation and DNA repair associated with soybean transgenesis: a comparison to cultivars and mutagenized plants. BMC Biotechnol 16(1):41

ASA (2023) SoyStats 2023. International: world soybean production. American Soybean Association, St. Louis, Missouri. http://soystats.com/international-world-soybean-production/ Accessed September 2023

Axelos M, Bardet C, Liboz T, Le Van Thai A, Curie C, Lescure B (1989) The gene family encoding the *Arabidopsis thaliana* translation elongation factor EF-1α: Molecular cloning, characterization and expression. Mol Gen Genet 219:106-112

Barker RF, Idler KB, Thompson DV, Kemp JD (1983) Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. Plant Mol Biol 2:335-350

Bartholomaeus A, Parrott W, Bondy G, Walker K, ILSI (2013) The use of whole food animal studies in the safety assessment of genetically modified crops: limitations and recommendations. Crit Rev Toxicol. 43(Suppl 2):1-24

Bäumlein H, Boerjan W, Nagy I, Bassüner R, Van Montagu M, Inzé D, Wobus U (1991) A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and *Arabidopsis* plants. Mol Gen Genet 225:459-467

Bauer R, Bekker JP, van Wyk N, du Toit C, Dicks LMT, Kossmann J (2009) Exopolysaccharide production by lactose-hydrolyzing bacteria isolated from traditionally fermented milk. Int J Food Microbiol. 131(2-3):260–264

Bautista-Zapanta J-n, Suzuki K, Yoshida K (2002) Characterization of four ribosomal RNA operons in the genome of *Agrobacterium tumefaciens* MAFF301001. Nucleic Acids Res S2:91-92

Beck E, Ludwig G, Auerswald EA, Reiss B, Schaller H (1982) Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene 19:327-336

Bellaloui N, Smith JR, Gillen AM, Ray JD (2011) Effects of maturity, genotypic background, and temperature on seed mineral composition in near-isogenic soybean lines in the early soybean production system. Crop Sci 51:1161-1171

Behrens MR, Mutlu N, Chakraborty S, Dumitru R, Jiang WZ, LaVallee BJ, Herman PL, Clemente TE, Weeks DP (2007) Dicamba resistance: Enlarging and preserving biotechnology-based weed management strategies. Science 316:1185-1188

Berman KH, Harrigan GG, Riordan SG, Nemeth MA, Hanson C, Smith M, Sorbet R, Zhu E, Ridley WE (2009) Compositions of seed, forage, and processed fractions from insect-protected soybean MON 87701 are equivalent to those of conventional soybean. J Agric Food Chem 57:11360-11369

Bevan M, Barnes WM, Chilton M-D (1983) Structure and transcription of the nopaline synthase gene region of T-DNA. Nucleic Acids Res 11:369-385

Breeze ML, Meng C, Harrison JM, George C, Colyer JD (2015) Relationship between composition of oilseed processed fractions and the whole oilseeds. J Am Oil Chem Soc 92:203-213

Chakraborty S, Behrens M, Herman PL, Arendsen AF, Hagen WR, Carlson DL, Wang X-Z, Weeks DP (2005) A three-component dicamba O-demethylase from *Pseudomonas maltophilia*, strain DI-6: Purification and characterization. Arch Biochem Biophys 437:20-28

Chaudhardy DK, Jeong SW, Kim J (2017) *Sphingobium naphthae* sp. nov., ith the ability to degrade aliphatic hydrocarbons, isolated from oil-contaminated soil. Int J Syst Evol Microbiol 67:2986-2993

Coruzzi G, Broglie R, Edwards C, Chua N-H (1984) Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1, 5-bisphosphate carboxylase. EMBO J 3:1671-1679

D'Ordine RL, Rydel TJ, Storek MJ, Sturman EJ, Moshiri F, Bartlett RK, Brown GR, Eilers RJ, Dart C, Qi Y, Flasinski S Franklin SJ (2009) Dicamba monooxygenase: Structural insights into a dynamic Rieske oxygenase that catalyzes an exocyclic monooxygenation. J Mol Biol 392:481-497

Dai S, Georgelis N, Bedair M, Hong Y-J, et al. (2022) Ectopic expression of a rice triketone dioxygenase gene confers mesotrione tolerance in soybean. Pest Manag Sci 78:2816-2827

Delaney B, Astwood JD, Cunny H, et al. (2008) Evaluation of protein safety in the context of agricultural biotechnology. Food Chem Toxicol 46 Suppl 2:S71-97

Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM (1982) Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet. 1(6):561-573

Duke S (2005) Taking stock of herbicide-resistant crops ten years after introduction. Pest Manag Sci 61: 211-218

Dumitru R, Jiang WZ, Weeks DP, Wilson MA (2009) Crystal structure of dicamba monooxygenase: A Rieske nonheme oxygenase that catalyzes oxidative demethylation. J Mol Biol 392:498-510

EFSA (2011) Conclusion on the peer review of the pesticide risk assessment of the active substance dicamba. The EFSA Journal 9(1):1965

Enya J, Shinohara H, Yoshida S, Tsukiboshi T, Negishi H, Suyama K, Tsushima S (2007) Culturable Leaf-Associated Bacteria on Tomato Plants and Their Potential as Biological Control Agents. Microb Ecol 53(4):524–536

FAOSTAT (2023) FAOSTAT Statistics Division Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/#home. Accessed September 2023

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

Fraley RT, Rogers SG, Horsch RB, et al. (1983) Expression of bacterial genes in plant cells. PNAS 80:4803-4807

FSANZ (2012) Application A1063 – Food derived from herbicide-tolerant soybean MON87708. Report prepared by Food Standards Australia New Zealand. https://www.foodstandards.gov.au/code/applications/Pages/applicationa1063food5198.aspx

FSANZ (2013a) Application A1080 – Food derived from herbicide-tolerant cotton line MON88701. Report prepared by Food Standards Australia New Zealand. https://www.foodstandards.gov.au/code/applications/Pages/A1080-Food-derived-from-Herbicide-tolerant-Cotton-Line-MON88701.aspx

FSANZ (2013b) Application A1081 – Food derived from herbicide-tolerant soybean event SYHT0H2. Report prepared by Food Standards Australia New Zealand. https://www.foodstandards.gov.au/code/applications/Pages/a1081foodderivedfrom5825.aspx

FSANZ (2016) Application A1118 – Food derived from herbicide-tolerant corn line MON87419. Report prepared by Food Standards Australia New Zealand. https://www.foodstandards.gov.au/code/applications/Pages/A1118GM-CornLineMON87419.aspx

FSANZ (2020) Application A1192 – Food derived from herbicide-tolerant corn line MON87429. Report

prepared by Food Standards Australia New Zealand.
https://www.foodstandards.gov.au/code/applications/Pages/A1192.aspx

FSANZ (2021) Application A1216 – Food derived from herbicide-tolerant canola line MON94100. Report prepared by Food Standards Australia New Zealand.
https://www.foodstandards.gov.au/code/applications/Pages/A1216.aspx

Giza PE, Huang RC (1989) A self-inducing runaway-replication plasmid expression system utilizing the Rop protein. Gene 78:73-84

Gomes IB, Simões LC, Simões M (2018) The effects of emerging environmental contaminants on *Stenotrophomonas maltophilia* isolated from drinking water in planktonic and sessile states. Sci Total Environ 643:1348-1356

Hammond BG, Jez JM (2011) Impact of food processing on the safety assessment for proteins introduced into biotechnology-derived soybean and corn crops. Food Chem Toxicol. 49(4):711-721

Hausinger RP (2004) Fe(II)/α-Ketoglutarate-dependent hydroxylases and related enzymes. Crit Rev Biochem Mol Biol 39:21-68

Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. PNAS 89(22):10915-10919

Herman PL, Behrens M, Chakraborty S, Chrastil BM, Barycki J, Weeks DP (2005) A three-component dicamba O-demethylase from *Pseudomonas maltophilia*, strain DI-6: Gene isolation, characterization, and heterologous expression. J Biol Chem 280:24759-24767

Herouet C, Esdaile DJ, Mallyon BA, et al. (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. Regul Toxicol Pharmacol 41(2):134-149

Herrmann KM (1995) The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. The Plant Cell 7:907-919

Hunt AG (1994) Messenger RNA 3′ end formation in plants. Ann Rev Plant Physiol Plant Mol Biol 45:47-60.

ILSI Research Foundation (2016) A Review of the Food and Feed Safety of the PAT Protein. Washington, D.C. USA. International Life Sciences Institute, Center for Environmental Risk Assessment.

JMPR (2014) Report of the joint meeting of the FAO panel of experts of pesticide residues in food and the environment and the WHO core assessment group on pesticide residues. Rome, Italy.

Khush (1997) Origin, dispersal, cultivation and variation of rice. Plant Mol Biol 35:25-34

Klee HJ, Muskopf YM, Gasser CS (1987) Cloning of an *Arabidopsis thaliana* gene encoding 5-enolpyruvylshikimate-3-phosphate synthase: Sequence analysis and manipulation to obtain glyphosate-tolerant plants. Mol Gen Genet 210:437-442

Krueger JP, Butz RG, Atallah YH, Cork DJ (1989) Isolation and identification of microorganisms for the degradation of dicamba. J Agric and Food Chem 37:534-538

Larue CT, Goley M, Shi L, et al. (2019) Development of enzymes for robust aryloxyphenoxypropionate and synthetic auxin herbicide tolerance traits in maize and soybean crops. Pest Manag Sci 75:2086-2094

Lira F, Berg G, Martínez JL (2017) Double-face meets the bacterial world: The opportunistic pathogen *Stenotrophomonas maltophilia*. Front Microbiol 8:2190

Liu KS (2004) Soybeans as a powerhouse of nutrients and phytochemicals. In: Soybeans as Functional Foods and Ingredients (pp 1-22). AOCS Press, Champaign, Illinois.

Lundry DR, Ridley WP, Meyer JJ, Riordan SG, Nemeth MA, Trujillo WA, Breeze ML, Sorbet R (2008) Composition of grain, forage, and processed fractions from second-generation glyphosate-tolerant soybean, MON 89788, is equivalent to that of conventional soybean (Glycine max L.). J Agric Food Chem 56:4611-4622

Maeda H, Murata K, Sakuma N, et al. (2019) A rice gene that confers broad-spectrum resistance to β-triketone herbicides. Science 365:393-396

Mazodier P, Cossart P, Giraud E, Gasser F (1985) Completion of the nucleotide sequence of the central region of Tn5 confirms the presence of three resistance genes. Nucleic Acids Res 13:195-205

Meier U (2001) Growth stages of mono- and dicotyledonous plants. BBCH Monograph. Second Edition. Federal Biological Research Centre for Agriculture and Forestry, Grossbeeren, Germany

Mitchell G, Bartlett DW, Fraser TEM, Hawkes TR, Holt DC, Townson JK, Wichert RA (2001) Mesotrione: A new selective herbicide for use in maize. Pest Manag Sci 57:120-128

Mukherjee P, Roy P (2016) Genomic potential of *Stenotrophomonas maltophilia* in bioremediation with an assessment of its multifaceted role in our environment. Front Microbiol 7:967.

Munsch-Alatossava P, Alatossava T. (2006) Phenotypic characterization of raw milk-associated psychrotrophic bacteria. Microbiol Res 161(4):334-46

Myouga F, Motohashi R, Kuromori T, Nagata N, Shinozaki K (2006) An *Arabidopsis* chloroplast-targeted Hsp101 homologue, APG6, has an essential role in chloroplast development as well as heat-stress response. Plant J 48:249-260

Norris SR, Meyer SE, Callis J (1993) The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. Plant Mol Biol 21:895-906

OECD (1999) Consensus document on general information concerning the genes and their enzymes that confer tolerance to phosphinothricin herbicide. ENV/JM/MONO(99)13. Series on Harmonization of Regulatory Oversight in Biotechnology No.11. Organisation for Economic Co-operation and Development, Paris.

OECD (2003) Considerations for the Safety Assessment of Animal Feedstuffs Derived from Genetically Modified Plants: Series on the Safety of Novel Foods and Feeds. Report No: 9. Development, OFEC-OA. Paris.

OECD (2006) Soybean (GLYCINE MAX (L.) MARR.). Organisation for Economic Co-operation and Development. Paris.

OECD (2012) Revised Consensus Document on Compositional Considerations for New Varieties of Soybean [Glycine Max (L.) Merr.]: Key Food and Feed Nutrients, Antinutrients, Toxicants and Allergens: Series on the Safety of Novel Foods and Feeds. Report No: 25. Organisation for Economic Cooperation and Development. Paris.

OECD. 2016. Revised consensus document on compositional considerations for new varieties of rice (*Oryza sativa*): Key food and feed nutrients, anti-nutrients and other constituents. Organisation for Economic Co-operation and Development, Paris.

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. PNAS 85(8):2444-2448

Piper KR, von Bodman SB, Hwang I, Farrand SK (1999) Hierarchical gene regulatory systems arising from fortuitous gene associations: Controlling quorum sensing by the opine regulon in *Agrobacterium*.

Mol Microbiol 32:1077-1089

Richins RD, Scholthof HB, Shepherd RJ (1987) Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Res 15:8451-8466

Rijavec T, Lapanje A, Dermastia M, Rupnik M (2007) Isolation of bacterial endophytes from germinated maize kernels. Can J Microbiol 53(6):802–808

Rogers SG (2000) Promoter for transgenic plants. Patent 6,018,100, U.S. Patent Office, Washington, D.C.

Ryan RP, Monchy S, Cardinale M, Taghavi S, Crossman L, Avison MB, Berg G, van der Lelie D, Dow JM (2009) The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. Nat Rev Microbiol 7:514-525

Salomon S, Puchta H (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J 17(20):6086-6095

Schmidt CL, Shaw L (2001) A comprehensive phylogenetic analysis of Rieske and Rieske-type iron-sulphur proteins. J Bioenerg Biomembr 33:9-26

Strauch E, Wohlleben W, Pühler A (1988) Cloning of a phosphinothricin N-acetyltransferase gene from *Streptomyces viridochromogenes* Tü494 and its expression in *Streptomyces lividans* and *Escherichia coli*. Gene 63(1):65-74

Sutcliffe JG (1979) Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symposia on Quantitative Biology 43:77-90

Thomas P, Kumari S, Swarna GK, Gowda TKS (2007) Papaya shoot tip associated endophytic bacteria isolated from in vitro cultures and host-endophyte interaction *in vitro* and *in vivo*. Can J Microbiol 53(3):380–390

Thomas K, Aalbers M, Bannon GA, et al. (2004) A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. Regul Toxicol Pharmacol 39:87-98

Thompson CJ, Movva NR, Tizard R, et al. (1987) Characterization of the herbicide-resistance gene bar from *Streptomyces hygroscopicus*. EMBO J 6(9):2519-2523

Thompson MM, Niemuth A, Sabbatini J, Levin D, Breeze ML, Li X, Perez T, Taylor M, Harrigan GG (2016) Analysis of vitamin K_1 in soybean seed: Assessing levels in a lineage representing over 35 years of breeding. J Am Oil Chem Soc 93:587-594

To JPC, Davis IR, Marengo MS, et al. (2021) Expression elements derived from plant sequences provide effective gene expression regulation and new opportunities for plant biotechnology traits. Front Plant Sci 12:712179

Todaro M, Francesca N, Reale S, Moschetti G, Vitale F, Settanni L (2011) Effect of different salting technologies on the chemical and microbiological characteristics of PDO Pecorino Siciliano cheese. Eur Food Res Technol 233:931-940.

Videira SS, Simoes de Araujo JL, da Silva Rodrigues L, Baldani VLD, Baldini JI (2010) Occurrence and diversity of nitrogen-fixing *Sphingomonas* bacteria associated with rice plants grown in Brazil. FEMS Microbiol Lett 293(1):11–19

Wang X-Z, Li B, Herman PL, Weeks DP (1997) A three-component enzyme system catalyzes the O demethylation of the herbicide dicamba in *Pseudomonas maltophilia* DI-6. Appl Environ Microbiol 63:1623-1626.

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. Nat Biotechnol 14(10):1274-1278

Wingfield P (2017) N-Terminal Methionine Processing. Curr Protoc Protein Sci 3(88):6.14.1-6.14.3

Wohlleben W, Arnold W, Broer I, Hillemann D, Strauch E, Punier A (1988) Nucleotide sequence of the phosphinothricin N-acetyltransferase gene from *Streptomyces viridochromogenes* Tü494 and its expression in *Nicotiana tabacum*. Gene 70:25-37

Yang L, Wang C, Holst-Jensen A, Morisset D, Lin Y, Zhang D (2013) Characterization of GM events by insert knowledge adapted re-sequencing approaches. Sci Rep 3:2839

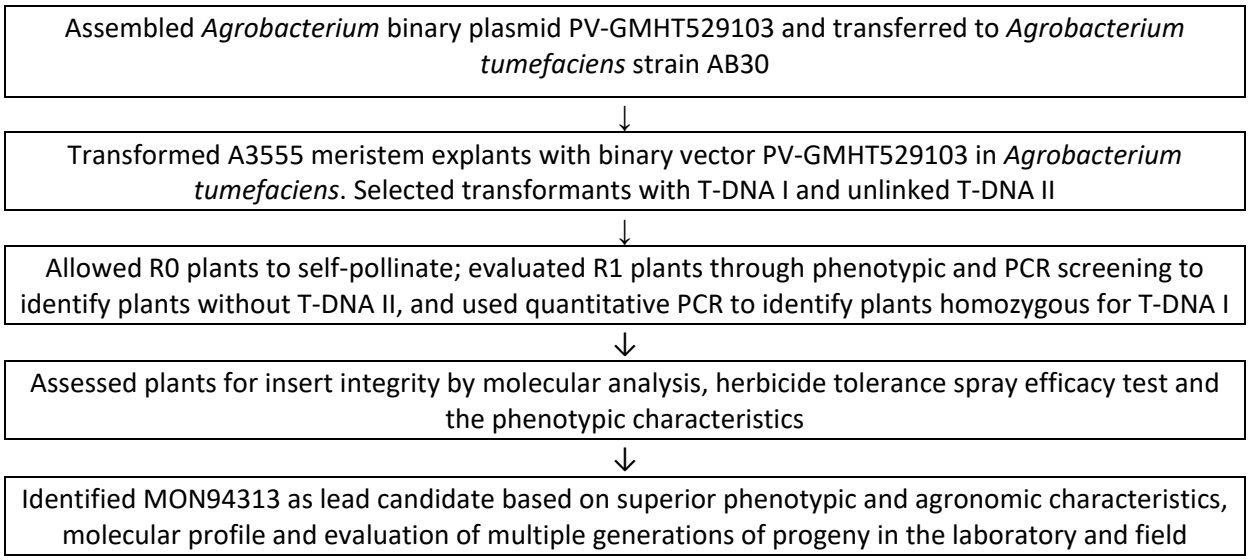
Ye X, Williams EJ, Shen J, Johnson S, Lowe B, Radke S, Strickland S, Esser JA, Petersen MW, Gilbertson MA (2011) Enhanced production of single copy backbone-free transgenic plants in multiple crop species using binary vectors with a pRi replication origin in *Agrobacterium tumefaciens*. Transgenic Res 20:773-786.

Zambryski P, Depicker A, Kruger K, Goodman HM (1982) Tumor induction by *Agrobacterium tumefaciens*: analysis of the boundaries of T-DNA. J Mol Appl Genet 1:361-370

Zastrow-Hayes GM, Lin H, Sigmund AL, et al. (2015) Southern-by-Sequencing: a robust screening approach for molecular characterization of genetically modified crops. Plant Genome 8(1): e2014.08.0037

Appendix 1

Development of MON94313



Appendix 2

Genetic elements present in PV-GMHT529103

Genetic element	Relative position	Description, Source & Reference
Right Border Region	1-285	DNA region from *Agrobacterium tumefaciens* containing the right border sequence used for transfer of the T–DNA (Depicker et al. 1982; Zambryski et al. 1982)
Intervening Sequence	286-324	DNA sequence used for cloning (synthetic)
P-*ubq3-At1*	325-1332	Promoter, leader and intron from *Arabidopsis thaliana* of the polyubiquitin gene *ubq3* (Norris et al. 1993), which directs transcription in plant cells
TS-*apg6-At1*	1333-1536	Targeting sequence of the *APG6* gene from *Arabidopsis thaliana* encoding a HSP101 (heat shock protein) homologue and acts as a transit peptide that directs transport of the protein to the chloroplast (Myouga et al. 2006)
dmo	1537-2559	Codon optimized coding sequence for the dicamba mono-oxygenase (DMO) protein of *Stenotrophomonas maltophilia* that confers dicamba resistance (Herman et al. 2005; Wang et al. 1997)
Intervening Sequence	2560-2578	Sequence used in DNA cloning
T-s*ali3-2-Mt1*	2579-3078	3' UTR sequence from *Medicago truncatula* (barrel medic) of an aluminium-induced Sali3-2 protein that directs polyadenylation of the mRNA (Hunt 1994)
Intervening Sequence	3079-3156	DNA sequence used for cloning (synthetic)
P-*GSP579*	3157-3656	A promoter and 5' UTR that has been developed from multiple promoter and 5' UTR sequences from *Arabidopsis thaliana* (To et al. 2021) that directs transcription in plant cells
I-*GSI102*	3657-3966	An intron that has been developed from multiple intron sequences from *Arabidopsis thaliana* (To et al. 2021) that is involved in regulating gene expression
Intervening Sequence	3967-3972	Sequence used in DNA cloning
pat	3973-4524	Codon optimized coding sequence for the phosphinothricin N–acetyltransferase (PAT) protein of *Streptomyces viridochromogenes* that confers tolerance to glufosinate (Wehrmann et al. 1996; Wohlleben et al. 1988)
Intervening Sequence	4525-4532	Sequence used in DNA cloning
T-*Hsp20-Mt1*	4533-5032	3' UTR sequence from *Medicago truncatula* (barrel medic) of a putative *Hsp20* gene encoding a heat shock protein that directs polyadenylation of the mRNA (Hunt 1994)
Intervening Sequence	5033-5115	DNA sequence used for cloning (synthetic)
P-*ubq10-At1*	5116-6317	Promoter, leader and intron from *Arabidopsis thaliana* of the polyubiquitin gene *ubq10* (Norris et al. 1993)
Intervening Sequence	6318-6323	Sequence used in DNA cloning
ft_t.1	6324-7211	Modified version of the R-2,4 dichlorophenoxypropionate dioxygenase (*RdpA*) gene of *Sphingobium herbicidovorans* that expresses a modified FOPs and 2,4-D dioxygenase protein (FT_T.1) that confers tolerance to 2,4-D herbicide in soybean (Larue et al. 2019).

Genetic element	Relative position	Description, Source & Reference
Intervening Sequence	7212-7219	Sequence used in DNA cloning
T-*guf-Mt2*	7220-7719	3' UTR from an expressed gene of *Medicago truncatula* of unknown function that directs polyadenylation of mRNA (Hunt 1994)
Intervening Sequence	7720-7857	DNA sequence used for cloning (synthetic)
P-*GSP576*	7858-8357	A promoter and 5' UTR that has been developed from multiple promoter and 5' UTR sequences from *Arabidopsis thaliana* (To et al. 2021) that directs transcription in plant cells
I-*GSI17*	8358-8657	An intron that has been developed from multiple intron sequences from *Arabidopsis thaliana* (To et al. 2021) that is involved in regulating gene expression
Intervening Sequence	8658-8692	Sequence used in DNA cloning
TDO	8693-9748	Codon optimized coding sequence for the triketone dioxygenase (TDO) protein of *Oryza sativa* that confers tolerance to mesotrione (Maeda et al. 2019)
Intervening Sequence	9749-9778	Sequence used in DNA cloning
T-*GST7*	9779-10078	A 3' UTR that has been developed from multiple 3' UTR sequences from *Zea mays* (maize) (To et al. 2021) that directs polyadenylation of the mRNA.
Intervening Sequence	10079-10178	Sequence used in DNA cloning
Left Border Region	10179-10620	DNA region from *Agrobacterium tumefaciens* containing the left border sequence used for transfer of the T–DNA (Barker et al. 1983)
Intervening Sequence	10621-10657	Sequence used in DNA cloning
ble1	10658-10809	Partial coding sequence of the bleomycin resistance gene from transposon Tn5 that confers antibiotic resistance (Mazodier et al. 1985)
Intervening Sequence	10810-10829	Sequence used in DNA cloning
nptII	10830-11624	Coding sequence of the *neo* gene from transposon Tn5 of *E. coli* encoding neomycin phosphotransferase II (NPT II) (Beck et al.1982) that confers neomycin and kanamycin resistance (Fraley et al. 1983)
P-*rrn*	11625-11849	Promoter of the ribosomal RNA operon from *Agrobacterium tumefaciens* (Bautista-Zapanta et al. 2002) that drives transcription in plant cells
Intervening Sequence	11850-11925	Sequence used in DNA cloning
OR-*ori-pBR322*	11926-12514	Origin of replication from plasmid pBR322 for maintenance of plasmid in *E. coli* (Sutcliffe 1979)
Intervening Sequence	12515-12941	Sequence used in DNA cloning
rop	12942-13133	Coding sequence for repressor of primer protein from the ColE1 plasmid for maintenance of plasmid copy number in *E. coli* (Giza and Huang 1989)
Intervening Sequence	13134-13321	Sequence used in DNA cloning
OR-*ori-pRi*	13322-17435	Origin of replication from plasmid pRi for maintenance of plasmid in *Agrobacterium* (Ye et al. 2011)

Genetic element	Relative position	Description, Source & Reference
Intervening Sequence	17436-17442	Sequence used in DNA cloning
Left Border Region	17443-17761	DNA region from *Agrobacterium tumefaciens* containing the left border sequence used for transfer of the T-DNA (Barker et al. 1983)
Intervening Sequence	17762-17793	Sequence used in DNA cloning
T-*nos*	17794-18046	3' UTR sequence of the *nopaline synthase (nos)* gene from *Agrobacterium tumefaciens* pTi encoding NOS that directs polyadenylation (Bevan et al. 1983; Fraley et al. 1983)
Intervening Sequence	18047-18062	Sequence used in DNA cloning
splA	18063-19520	Coding sequence of the *splA* gene from *Agrobacterium tumefaciens* strain C58 encoding the sucrose phosphorylase protein that catalyzes the conversion of sucrose to fructose and glucose–1–phosphate (Piper et al. 1999)
Intervening Sequence	19521-19532	Sequence used in DNA cloning
P-*Usp*	19533-20711	5' UTR, promoter, and enhancer sequence of an unknown seed protein gene from *Vicia faba* (broad bean) encoding an unknown seed protein that is involved in regulating gene expression (Bäumlein et al. 1991)
Intervening Sequence	20712-20762	DNA sequence used for cloning (synthetic)
T-*E9*	20763-21405	3' UTR sequence from *Pisum sativum* (pea) rbcS gene family encoding the small subunit of ribulose bisphosphate carboxylase protein (Coruzzi et al. 1984) that directs polyadenylation of the mRNA
Intervening Sequence	21406-21420	Sequence used in DNA cloning
aadA	21421-22212	Coding sequence for an aminoglycosidemodifying enzyme, 3'(9)-*O*-nucleotidyltransferase from the transposon Tn7 (Fling et al. 1985) that confers spectinomycin and streptomycin resistance
TS-*CTP2*	22213-22440	Targeting sequence of the *ShkG* gene from *Arabidopsis thaliana* encoding the EPSPS transit peptide region that directs transport of the protein to the chloroplast (Herrmann 1995; Klee et al. 1987)
Intervening Sequence	22441-22449	Sequence used in DNA cloning
P-*EF-1*α	22450-23597	Promoter, leader , and intron sequences of the *EF–1*α gene from *Arabidopsis thaliana* encoding elongation factor EF–1α (Axelos et al. 1989) that directs transcription in plant cells
Intervening Sequence	23598-23620	Sequence used in DNA cloning
E-*FMV*	23621-24157	Enhancer from the 35S RNA of figwort mosaic virus (FMV) (Richins et al. 1987) that enhances transcription in most plant cells (Rogers 2000)
Intervening Sequence	24158-24203	Sequence used in DNA cloning
B-Right Border Region	24204-24534	DNA region from *Agrobacterium tumefaciens* containing the right border sequence used for transfer of the T–DNA (Depicker et al. 1982; Zambryski et al. 1982)
Intervening Sequence	24535-24549	Sequence used in DNA cloning